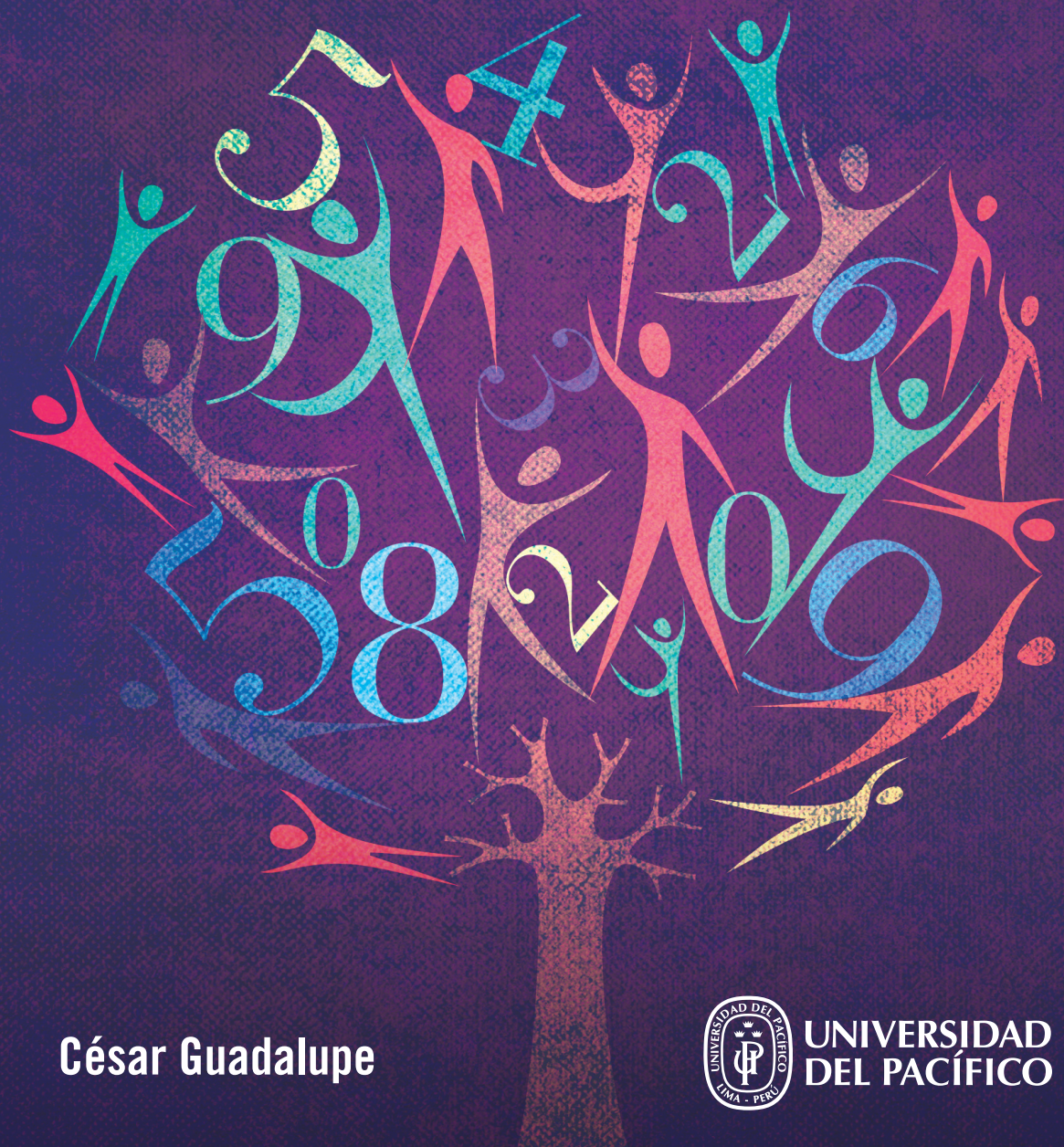


Contar para que cuente:

una introducción general a los sistemas de información educativa



César Guadalupe



UNIVERSIDAD
DEL PACÍFICO

Contar para que cuenta:

una introducción general a los sistemas
de información educativa

César Guadalupe



**UNIVERSIDAD
DEL PACÍFICO**

© César Guadalupe, 2015

De esta edición:

© Universidad del Pacífico
Av. Salaverry 2020
Lima 11, Perú
www.up.edu.pe

**CONTAR PARA QUE CUENTE:
UNA INTRODUCCIÓN GENERAL A LOS
SISTEMAS DE INFORMACIÓN EDUCATIVA**

César Guadalupe

1ª edición versión *e-book*: noviembre de 2015
Diseño de la carátula: Icono Comunicadores
ISBN: 978-9972-57-344-6

BUP

Guadalupe Mendizábal, César.

Contar para que cuente : una introducción general a los sistemas de información educativa / César Guadalupe. -- 1a edición versión e-book. -- Lima : Universidad del Pacífico, 2015.

391 p.

1. Educación -- Sistemas de información
2. Educación -- Estadística
3. Indicadores educativos
4. Evaluación educativa
- I. Universidad del Pacífico (Lima)

370.21 (SCDD)

Miembro de la Asociación Peruana de Editoriales Universitarias y de Escuelas Superiores (Ape-su) y miembro de la Asociación de Editoriales Universitarias de América Latina y el Caribe (Eulac).

La Universidad del Pacífico no se solidariza necesariamente con el contenido de los trabajos que publica. Prohibida la reproducción total o parcial de este texto por cualquier medio sin permiso de la Universidad del Pacífico.

Derechos reservados conforme a Ley.

ÍNDICE

Siglas usadas	9
Introducción	13
Agradecimientos.....	21
Sección 1. Los cimientos.....	25
1.1 La información tiene un propósito.....	26
1.2 Las dimensiones que conforman un sistema de información	31
1.2.1 Los destinatarios del sistema.....	31
1.2.2 Los elementos de información que han de componer el sistema	33
1.2.3 La dinámica institucional existente.....	34
1.2.4 Dimensiones y atributos básicos.....	34
1.2.4.1 Arreglo institucional	36
1.2.4.2 Elementos técnico-científicos	38
1.2.4.3 Funcionalidad	39
1.2.5 La construcción de las herramientas con las que opera el sistema.....	41
1.3 ¿Qué contar para que la información cuente?.....	43
1.3.1 Definir y diferenciar conceptos	44
1.4 Directorios y tasas de respuesta	49
1.5 Sistemas de captura de datos	53
1.5.1 Estadísticas y registros	53
1.5.2 Fechas de cierre del levantamiento (convivir con la omisión pero conocerla)	54

1.5.3	Períodos de referencia y cálculo de edades	56
1.5.4	El problema de la «oportunidad» de los datos y cómo abordarlo	60
1.6	La base tecnológica	64
1.7	¿Registros administrativos o información estadística?	66
1.8	La estadística educativa como parte de un Sistema Estadístico Nacional.....	69
1.9	La diseminación de la información	75
1.10	Niveles de observación y análisis	87
1.10.1	El propósito de los sistemas nacionales de información y del desarrollo de sus instrumentos y herramientas.....	89
1.10.2	Los vínculos entre los sistemas de información y la gestión en sus distintos ámbitos.....	93
Sección 2. Sistemas de indicadores.....		95
2.1	¿Por qué un sistema de indicadores y no un listado?.....	99
2.2	El sistema de indicadores	101
2.3	Calidad de la educación	108
2.3.1	El concepto de calidad de la educación	109
2.3.2	Dimensiones de la calidad de los sistemas educativos	119
2.3.3	La correspondencia entre el enfoque propuesto y los modelos de análisis más extendidos sobre la medición de la calidad de la educación	124
2.4	Construcción de indicadores sobre algunos temas centrales del debate educativo	128
2.4.1	Cobertura y perfiles de matrícula	128
2.4.2	Eficiencia	147
2.4.3	Nivel de escolaridad de la población.....	175
2.4.4	Ingreso.....	176
2.4.5	Conclusión de un nivel educativo	179
2.4.6	Gasto o inversión y esfuerzo.....	190
2.4.7	Equidad	202
2.4.8	Medición de aprendizajes (con consideraciones sobre el caso del alfabetismo adulto)	205
2.4.8.1	Temas fundamentales vinculados a la medición de aprendizajes ...	211
2.4.8.2	Aspectos adicionales de la medición de aprendizajes en la escuela: características socioeconómicas y culturales de los estudiantes y desarrollo de actitudes.....	235
2.4.9	La medición de los contenidos de la educación	239
2.4.10	Los docentes	241

Sección 3. Mejorando el sistema.....	257
3.1 Consistencia	258
3.1.1 Consistencia interna.....	258
3.1.2 Consistencia temporal.....	259
3.1.3 Consistencia entre diversas fuentes.....	261
3.2 Omisiones y estimaciones; proyecciones	263
3.2.1 Omisiones y estimaciones	263
3.2.2 Proyecciones	266
3.2.2.1 ¿Cómo proyectar una tendencia histórica?.....	266
3.2.2.2 ¿Cómo proyectar en base a un modelo de simulación?	277
3.3 Metadatos e interpretación.....	279
3.4 Censos, muestras y periodicidad	282
3.5 Datos agregados y microdatos	285
3.6 Cartografía educativa	287
3.7 Educación no formal	289
3.7.1 Los conceptos de la CINE 1997.....	289
3.7.2 Los conceptos de Eurostat.....	293
3.7.3 La revisión de conceptos de la CINE 2011.....	296
3.7.4 Efectos potenciales de la ampliación del alcance de los sistemas de información	297
 Sección 4. Análisis.....	 299
4.1 Modelos de análisis.....	301
4.2 Herramientas y procedimientos de análisis.....	304
4.3 Alcance.....	307
4.4 El análisis comparado	309
4.5 Índices	312
4.6 Comentario final	315
 Reflexiones finales	 317
 Anexos	 323
 Anexo 1. Conceptos estadísticos básicos	 323
1. Variables y niveles de medición	325
2. Elementos de estadística descriptiva	327
2.1 Descripciones basadas en una sola variable (univariabes)	328
2.2 Descripciones basdas en más de una variable (multivariabes).....	337

ÍNDICE

3.	Elementos de estadística inferencial.....	345
3.1	Muestreo.....	346
3.1.1	El tamaño de la muestra.....	348
3.1.2	Formas de obtener una muestra	351
3.1.3	El cálculo de los errores de estimación.....	353
3.2	Estimación de parámetros (medias y proporciones).....	354
3.3	Pruebas de hipótesis (diferencias de medias y proporciones).....	355
3.3.1	Pasos en la prueba de hipótesis.....	356
3.3.2	Usos de la prueba de hipótesis.....	357
3.4	Pruebas de significancia (asociación bivariable)	359
4.	Escalas actitudinales.....	361
4.1	Escala de Likert.....	362
4.2	Escala de Thurstone	365
4.3	Escala de Guttman.....	367
	Anexo 2. Cronología del desarrollo internacional de la información educativa	370
	Referencias	375

SIGLAS USADAS

ABS	Sistema de freno antibloqueo (por sus siglas en inglés)
AED	Academy for Educational Development
AIMS	Assessment, Information Systems, Monitoring and Statistics-Unesco Bangkok
ALL	Adult Literacy and Lifeskills Survey
Bicse	Board on International Comparative Studies in Education
CE	Centro educativo
Celade	Centro Latinoamericano y Caribeño de Demografía de la Cepal
Cepal	Comisión Económica para América Latina y el Caribe de las Naciones Unidas
CINE	Clasificación Internacional Normalizada de Educación
CIPP	Contexto-insumos-procesos-producto
Cived	Estudio sobre Educación Cívica (por sus siglas en inglés)
Civics	Estudio Internacional sobre Educación Cívica y Ciudadana (por sus siglas en inglés)
Comped	Estudio sobre Computadoras en la Educación (por sus siglas en inglés)
Confemen	Conferencia de Ministros de Educación de los Países de Habla Francesa (por sus siglas en francés)
DQAF	Data Quality Assessment Framework
EGRA	Early Grade Reading Assessment
EPT	Educación para Todos
FIMS	Estudio Internacional sobre Matemáticas (por sus siglas en inglés)
FMI	Fondo Monetario Internacional
GAPE	Grado aprobado promedio esperado
GMR	Global Monitoring Report

SIGLAS USADAS

GPS	Sistema de posicionamiento global (por sus siglas en inglés)
IALS	International Adult Literacy Survey
IBE	Oficina Internacional de Educación (por sus siglas en inglés)
ICCS	Estudio Internacional sobre Educación Cívica y Ciudadana (por sus siglas en inglés)
IEA	International Association for the Evaluation of Educational Achievement
IIEP	Instituto Internacional de Planificación Educativa de la Unesco (por sus siglas en inglés)
IMF	Fondo Monetario Internacional (por sus siglas en inglés)
INES	Indicators of Education Systems, OECD
ISO	Organización Internacional de Estándares (por sus siglas en inglés)
LAMP	Programa de Evaluación y Monitoreo del Alfabetismo (por sus siglas en inglés)
Llece	Laboratorio Latinoamericano de Evaluación de la Calidad Educativa
NCES	Centro Nacional de Estadísticas Educativas de los Estados Unidos (por sus siglas en inglés)
OECD	Organización para la Cooperación y el Desarrollo Económico (por sus siglas en inglés)
ONE	Oficina nacional de estadística
ONG	Organización no gubernamental
Orealc	Oficina Regional de Educación para América Latina y el Caribe de la Unesco
Pasec	Programa de Análisis de los Sistemas Educativos de la Confemen
Piaac	Programme for International Assessment of Adult Competencies
PIB	Producto interno bruto
Pirls	Estudio Internacional de Progreso en Comprensión Lectora (por sus siglas en inglés)
PISA	Programme for International Student Assessment
PNUD	Programa de las Naciones Unidas para el Desarrollo
PPP	Paridad de poder de compra (por sus siglas en inglés)
Preal	Programa de Promoción de la Reforma Educativa en América Latina y el Caribe
Prelac	Proyecto Regional de Educación para América Latina y el Caribe
PRIE	Proyecto Regional de Indicadores Educativos
RECH	Registro de Estudiantes de Chile
RLS	Estudio en Alfabetismo Lector (por sus siglas en inglés)
Sacmeq	Consortio de África Meridional y Oriental para la Supervisión de la Calidad de la Educación (por sus siglas en inglés)
SDG	Objetivos de Desarrollo Sostenible (por sus siglas en inglés)

SEN	Sistema Estadístico Nacional
Simce	Sistema de Medición de la Calidad de la Educación, Chile
SIMS	Estudio Internacional sobre Matemáticas (por sus siglas en inglés)
SIRI	Sistema Regional de Información-Unesco
SISS	Estudio Internacional sobre Ciencias (por sus siglas en inglés)
Siteal	Sistema de Información de Tendencias Educativas de América Latina
Sites	Segundo Estudio sobre Tecnologías de la Información en Educación (por sus siglas en inglés)
SRS	Muestreo simple al azar (por sus siglas en inglés)
TEDS-M	Estudio Internacional sobre la Formación Inicial del Profesorado de Matemáticas (por sus siglas en inglés)
Timms	Estudio Internacional de Tendencias en Matemáticas y Ciencias (por sus siglas en inglés)
UEE	Unidad de Estadística Educativa, Perú
UIL	Instituto de la Unesco para el Aprendizaje a lo Largo de Toda la Vida (por sus siglas en inglés)
UIS	Instituto de Estadística de la Unesco (por sus siglas en inglés)
Unesco	Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura
Unicef	Fondo de las Naciones Unidas para la Infancia
UNSD	División de Estadística de las Naciones Unidas (por sus siglas en inglés)
WEI	World Education Indicators Programme

INTRODUCCIÓN

A finales de 2007 e inicios de 2008, luego de un importante cambio en mis responsabilidades laborales, nació la idea de compilar un conjunto de aprendizajes en un volumen único y coherente. Este es el punto de partida desde el que se empezó a producir las notas que derivaron en este texto.

En aquel momento, tras más de una década trabajando en el terreno de la información educativa, sentí la necesidad de poner por escrito lo que había aprendido desde las primeras reuniones que tuve para discutir los términos de mi primer contrato con el Ministerio de Educación en el Perú a inicios de 1998 hasta mis vivencias como funcionario de la Unesco, que me han llevado a tener contacto con los sistemas de información de prácticamente todos los países de la América Latina de habla hispana, Brasil, varios del Caribe anglófono y unos más en otras latitudes y asimismo a participar de dos grupos de trabajo internacional de primer nivel encargados de proponer mejoras en los indicadores a utilizar para dar cuenta de (y mejorar) la calidad de los aprendizajes. También, debo confesar, el recuerdo de mi primer encuentro con los formularios de acopio de datos estadísticos del Ministerio de Educación en 1988 en una provincia del sur andino peruano es algo que me ha asaltado muchas veces, presentándome los rostros y las inquietudes no solo de aquellos que diseñan los cuestionarios usados para el acopio de datos, sino también los de quienes, en cierta medida, **padecen** su administración.

Así, este texto, si bien escrito por un individuo, da cuenta de un largo proceso de aprendizaje desarrollado a través de múltiples experiencias con otras personas en el interior del Perú, en su capital y a nivel internacional. La gran fortuna que he tenido en estos años de mi vida profesional se torna para mí en una obligación: poner por escrito y entregar lo aprendido a aquellos que facilitaron que yo aprendiera.

La experiencia nacional es absolutamente invaluable en este terreno. Compartir con aquellos involucrados en la operación directa de un sistema de información es una

fuerza inmensa de aprendizajes. Siempre recuerdo con afecto la arriesgada idea que, con el equipo de la Unidad de Estadística Educativa (UEE) del Ministerio de Educación del Perú, concretamos a inicios del año 2000: en vez de organizar, como todos los años, como lanzamiento del ciclo de producción estadística una docena de talleres regionales con las oficinas del interior que después harían el «efecto multiplicador» con las oficinas locales de todo el país, decidimos juntar a los más de trescientos responsables de la tarea estadística en un único lugar durante una semana, combinando algunos momentos con todos los participantes y, luego, una estructura de diez talleres en paralelo. La idea funcionó no solo en términos de la organización del ciclo anual de producción de datos, sino y fundamentalmente en términos de galvanizar un equipo nacional a cargo de la tarea. Compartir las sesiones de trabajo y los ratos de ocio con esta multitud compuesta por más de trescientos «estadísticos» fue un terreno de aprendizaje profesional y humano difícilmente replicable. Debo agradecer a quienes en el ministerio me dejaron ir adelante con esa idea, a la cooperación británica que nos apoyó financieramente, a cada uno de los participantes y, en especial, al equipo de la UEE en Lima por el enorme esfuerzo que tuvo que desplegar para que esto funcione (y para luego sostener este trabajo como parte de la rutina anual). En años posteriores, una de las cosas que más he lamentado es ver que muchos «expertos» no llegarían a tener la menor idea de por qué lo descrito en estas líneas posee algún valor.

Los sistemas de información educativa son un área del conocimiento humano y, al mismo tiempo, algo que se encuentra profundamente entroncado en la función pública, con todo lo que esto último implica. Por eso, a veces es algo distante de la idea de una disciplina científica y termina perdiéndose en oscuros corredores o sótanos de la burocracia (nacional o internacional), sin que esta tenga ninguna capacidad para aprender, ni para ayudar a que otros aprendan.

Sin embargo, el hecho de que los sistemas de información educativa formen parte de la función pública es también una fuente posible de fortalezas, ya que puede y debe ayudar a que su desarrollo se nutra de la reflexión intelectual, pero sin perder de vista un propósito específico: **los sistemas de información educativa existen para contribuir a hacer realidad el derecho de las personas a la educación.** Y nada como trabajar en un país teniendo contactos regulares con estudiantes, madres, padres, docentes y gestores de la educación para recordar esto de modo permanente.

Esa es la preocupación central desde la que se ha escrito este texto: contribuir a esa tarea prestando un servicio a todos aquellos que me ayudaron a que de alguna forma yo tenga algún conocimiento de este tema y también a los que han venido después y seguirán viniendo a continuar el trabajo en este terreno.

El aprendizaje adquirido en el país cobra un carácter distinto cuando uno tiene la oportunidad de contrastarlo con las penurias, travesías, anhelos, logros, frustraciones y miserias que se experimentan en otras latitudes. Cada experiencia nacional está teñida

por características propias derivadas de las historias nacionales e institucionales, de la cultura y de todo aquello que entretiene la mente de los científicos sociales y hace la vida humana tan rica a la vez que tan desafiante. Cada equipo nacional es una historia distinta y, a pesar de existir retos comunes, aporta cosas nuevas, sufre sus propios reveses y alcanza sus propios éxitos de manera que revelan nuevos elementos que enriquecen el aprendizaje. Personalmente, he tenido una profunda fortuna al poder compartir a lo largo de los últimos años con muchas personas en diversos lugares.

Pero las experiencias no son suficientes, necesitamos elementos que las ordenen, que les den sentido, que nos permitan capitalizar en aprendizaje lo observado y experimentado, porque, de lo contrario, se corre el riesgo de que devengan solamente en un acervo de anécdotas. Hoy, contamos con algunas de esas herramientas, pero nos hacen falta otras. Justamente esa es una razón más (y de primera importancia) para escribir este texto: sistematizar y brindar un marco de referencia que ayude a capitalizar las experiencias en aprendizajes sostenibles. Y documentar de una forma sistemática y reflexiva con un sentido crítico profundo nos ayuda a crecer. Espero que este texto sea una contribución en dicho sentido.

Durante los años 2001 a 2003, desde la Oficina Regional de Educación para América Latina y el Caribe (Orealc) de la Unesco en Santiago de Chile, iniciamos un proceso de visitas a países que, dado el alcance del proyecto que las hacía posible, se concentraba en una revisión sistemática y documentada del proceso implicado en reportar datos a nivel internacional¹. Claro está que dicho reporte de datos depende, en último término, de las debilidades y fortalezas del sistema de información que hay detrás y que la única forma sensata de contar con información internacional robusta y sostenida pasa por: (i) contar con las herramientas metodológicas adecuadas y (ii) que los países tengan sistemas de información sólidos respecto de los cuales la información reportada a nivel internacional es una pequeña parte. Debo anotar, sin embargo, que esa pequeña parte puede tener un efecto enorme (positivo o negativo) cuando se traduce en publicaciones o acciones internacionales de gran visibilidad.

Así, dichas visitas llevaron a que el equipo del Sistema Regional de Información (SIRI; es decir, la unidad de estadística educativa de la Oficina Regional de la Unesco en Santiago) reflexionara sobre la necesidad de hacer observaciones sistemáticas de los sistemas de información educativa de los países de la región. Para eso, elaboramos una pequeña guía

¹ En el Programa WEI (World Education Indicators Programme), del que participaron alrededor de dieciséis países de ingreso medio entre 1998 y 2013 y que constituyó el marco en el cual estos reportaban datos educativos a nivel internacional, se desarrollaron algunas versiones preliminares de estas revisiones, llamadas **planes de datos**. En 1998 y 1999 una versión general fue producida a solicitud y con recursos del Perú por un consultor del programa a efectos de facilitar la incorporación de este país a dicho programa; simultáneamente otros países tuvieron una experiencia similar, aunque concentrada en los aspectos financieros, ya que se correspondían con el estudio sobre financiamiento realizado en esos años (Unesco-UIS y OECD 2000).

de observación al tiempo que empezábamos a buscar materiales que ayudaran en esta tarea. Encontrar que en el terreno de las estadísticas económicas existía una herramienta llamada Data Quality Assessment Framework (DQAF) fue muy importante y ella se convirtió en un importante elemento para organizar el trabajo². Esto derivó en más de una docena de diagnósticos de los sistemas de información nacional que tenían como propósito brindar un punto de partida sólido para desarrollar acciones de capacitación y fortalecimiento de los sistemas de información a escala regional³.

Al mismo tiempo, la creación del Instituto de Estadística de la Unesco (UIS) en 1999 permitió contar con espacios regulares de encuentro entre los funcionarios responsables de las estadísticas educativas en las diversas regiones del mundo. En el caso de América Latina y el Caribe, los talleres regionales iniciados en 2000 (La Habana y Kingston) y posteriormente sostenidos en 2001 (Ciudad de Panamá y Bridgetown), 2002 (Lima y Port of Spain), 2003 (Quito y St. George's), 2005 (Bogotá), 2006 (Bridgetown), 2007 (Cancún), 2009 (Cancún), 2011 (Viña del Mar) y 2013 (Antigua) han sido espacios invaluable para el intercambio entre los países y para el encuentro entre las personas. Esto fortaleció la capacidad del SIRI –creado en 1984– para operar de modo efectivo como una red que aglutina a las unidades de estadística educativa de los ministerios de Educación tanto en América Latina como en el Caribe. Tuve la fortuna de participar de este proceso primero como representante de país (entre 1999 y 2001) y luego como integrante y responsable del SIRI en la Unesco hasta mayo de 2007.

Simultáneamente, la región ha experimentado desde mediados de los años 1980, pero con especial énfasis a partir de los años 1990 –dado el empuje brindado por los procesos de reforma inspirados y sostenidos mediante operaciones de crédito, principalmente, con el Banco Mundial y el Banco Interamericano de Desarrollo–, el desarrollo de sistemas de medición de logro académico de los estudiantes, usualmente conocidos como sistemas de medición o evaluación de la calidad educativa⁴.

Estos sistemas de medición de logro académico de los estudiantes se han desarrollado en paralelo a los sistemas estadísticos «tradicionales» de los ministerios y, con mucha frecuencia, en medio de situaciones de falta de coordinación y comunicación y/o de desconfianza y conflicto abierto. De alguna forma, muchos de estos sistemas fueron deliberadamente creados en paralelo como una forma de invertir importantes recursos en ellos y de desarrollar capacidades profesionales que no se veían vinculadas a la

² El DQAF fue desarrollado por el Fondo Monetario Internacional (IMF 2012) y es un marco genérico que ayuda a observar los sistemas de información. Por su parte, el UIS-Unesco y el Banco Mundial produjeron una (no muy sólida) versión para el caso de las estadísticas educativas (Unesco-UIS y World Bank 2003).

³ Los resultados de estos diagnósticos se presentan sumariamente en: Unesco Santiago y México-Secretaría de Educación Pública (2008).

⁴ La identificación entre calidad educativa y logro académico de los estudiantes tal y como es revelado en pruebas cognitivas estandarizadas es un tema que se discute en la segunda sección de este documento.

estadística tradicional tenida por poco útil, dado el énfasis del discurso en la «calidad» (logro académico).

Esto ha provocado dos tipos de resultados, por una parte, una deseable consolidación de sistemas de medición de logro académico de los estudiantes y, por otra, una no deseable fractura en los sistemas de información que es potencial fuente de conflictos, desperdicio de recursos y otras debilidades.

En 1994, la Oficina Regional de la Unesco recibió el encargo de coordinar una red de colaboración entre las unidades de medición de logro académico de los estudiantes de América Latina mediante la creación del Laboratorio Latinoamericano de Evaluación de la Calidad Educativa (Llece), nombre que revela la antes aludida identificación conceptual entre calidad y logros académicos expresados como resultados de pruebas estandarizadas. El Llece ha sido capaz de conducir tres estudios regionales comparativos, el primero en 1997, el segundo en 2006 y el tercero en 2013, en base a un esquema de colaboración entre los países que ha permitido, por un lado, establecer un estilo de trabajo que consagra el rol protagónico de estos y, por otro, conducir dichos estudios, a pesar, incluso, de las serias limitaciones observadas en varios momentos en la coordinación de los procesos por parte de la Unesco. Haber podido colaborar con el laboratorio, especialmente en momentos muy difíciles de su devenir, me permitió tener un primer acercamiento sistemático a este otro componente de los sistemas de información educativa.

Posteriormente, conducir el Programa de Evaluación y Monitoreo del Alfabetismo (LAMP) del UIS y tener la responsabilidad de establecer las bases de las actividades del UIS en el terreno de la evaluación de los logros de aprendizaje de los estudiantes, me brindó una experiencia invaluable para mejor entender el lugar de este tipo de esfuerzos en el marco de un sistema de información educativa, así como la utilidad y los importantes límites que tienen las pruebas estandarizadas.

Al mismo tiempo, coordinar el Grupo de Trabajo Internacional sobre Evaluación y Mejora de la Calidad de los Aprendizajes, que operó entre 2008 y 2010, fue un espacio importante para contrastar mis ideas y preocupaciones con las de un grupo clave de personas que abordan temas similares en otras latitudes y desde distintas agendas. Lo mismo se aplica a la Learning Metrics Task Force, que fue establecida como un esfuerzo conjunto entre el UIS y The Brookings Institution en 2012 y donde tuve la oportunidad de actuar, primero, desde el secretariado y, luego, como coordinador del grupo de trabajo responsable de los temas de medición.

Todos los anteriormente mencionados han sido espacios de intercambio y aprendizaje. Una parte de lo aprendido se encuentra ya en diferentes publicaciones⁵ o en notas y

⁵ Unesco-UIS (2009b); Unesco Santiago (2007, 2008); Guadalupe y Taccari (2004); Guadalupe y Louzano (2003); Guadalupe (2002); Guadalupe *et al.* (2002); Guadalupe y Cardoso (2011); Learning Metrics Task

borradores no publicados, pero cada uno de estos documentos obedece a un propósito específico y a un momento particular. Así, me pareció pertinente abocarme a acopiar, reelaborar y redactar por vez primera un conjunto amplio de ideas en un solo volumen sistemático y estructurado que cubriera un grupo significativo de los temas abordados o sugeridos por reuniones, visitas a países e intercambios informales.

Este documento no busca sustituir o desconocer principios, normas y buenas prácticas que se encuentran documentadas, son accesibles y deberían ser de lectura obligatoria para quienes trabajan en el terreno de los sistemas de información del sector público⁶. Al contrario, busca dar cuenta de muchos de estos principios a través de sus manifestaciones en las tareas cotidianas de los sistemas de información de acuerdo a lo que ha sido mi observación vivencial de los mismos.

De esta manera, si bien la redacción de este volumen es un esfuerzo reciente, su contenido se ha ido inesperadamente construyendo a través de todas esas experiencias y, por lo mismo, debe muchos a muchos y, a la vez, honra la máxima de Nietzsche:

Ich will keinen Autor mehr lesen, dem man anmerkt, er wollte ein Buch machen: sondern nur jene, deren Gedanken unversehens ein Buch wurden⁷ (Nietzsche 1880: 121).

A efectos de organizar las ideas, el presente volumen se encuentra estructurado de la siguiente manera:

La **primera sección** se aboca al tratamiento de los aspectos más generales de la conformación de un sistema de información, poniendo particular énfasis en las cuestiones vinculadas a la clara identificación de propósitos y usuarios, así como a los aspectos institucionales centrales vinculados a su operación. Con estos elementos en mente, se define los grandes ámbitos de acción que abordan los sistemas de información educativa nacionales gestionados desde el sector público.

La **segunda sección** se detiene en la especificación de la información que generan o habrían de generar dichos sistemas de información. Para esto se enfatiza la necesidad de que la información sea parte de un conjunto coherente que le dé orden y significado, al tiempo que se presenta reflexiones generales sobre lo que es la calidad de la educación y se aborda en mayor detalle algunos temas centrales (cobertura y perfiles de matrícula, eficiencia, nivel de escolaridad de la población, alfabetismo, ingreso, conclusión, gasto,

Force (2013a). Dos textos recientes (Guadalupe 2010, 2013) fueron elaboradas sobre la base de borradores de algunos acápite del presente volumen.

⁶ En particular UNSD (2005, 2013) y las diversas reflexiones contenidas en Kisilevsky y Roca (2013). Asimismo, puede consultarse la propuesta de «Código de buenas prácticas de la Conferencia Estadística de las Américas» (Cepal 2011) y el código peruano al respecto (Perú-INEI 2012).

⁷ «No quiero volver a leer a un autor que se haya propuesto escribir un libro, solo a aquellos cuyas ideas se han tornado inadvertidamente en uno» (traducción de María Victoria Guadalupe).

equidad, logros de aprendizaje, contenidos de la educación y docentes) respecto de los cuales se expone consideraciones conceptuales y de política, así como detalles técnicos sobre indicadores que se producen o podrían o deberían producirse de modo regular.

La **tercera sección** trata sobre algunos aspectos de singular importancia para la mejora permanente de los sistemas de información. Especialmente toca temas vinculados a la consistencia, el manejo de las omisiones, estimaciones y proyecciones, la importancia de los llamados metadatos para la interpretación y uso de la información, la vinculación con la gestión, el uso de eventos censales y muestrales, la agregación de datos y el uso de microdatos y la cartografía educativa; finalmente, ofrece algunas reflexiones sobre la educación no formal.

La **cuarta sección** presenta elementos que se vinculan a las tareas de análisis de la información, contribuyendo a establecer los límites de los sistemas de generación de información y a identificar la necesidad de que la información se entronque con la investigación educativa.

Posteriormente, se presentan algunos **comentarios y reflexiones finales**.

Por último, dos **anexos** se abocan a proporcionar lo siguiente:

- El anexo 1 presenta una revisión general de conceptos y herramientas estadísticas básicas. El propósito de este anexo es atender al hecho de que una fracción importante del personal que trabaja en los sistemas de información educativa no cuenta con un entrenamiento formal en estadística y una breve introducción como la aquí presentada puede serle de utilidad para esclarecer y enriquecer algunos aspectos de su trabajo.
- El anexo 2 intenta reseñar algunos elementos relativos a cómo se ha abordado el tema de las estadísticas educativas a lo largo del tiempo a nivel internacional. Esto es importante tanto a efectos de identificar materiales de referencia y el contexto en el que fueron producidos, como para tomar en cuenta el momento en el que ciertos indicadores comúnmente usados fueron creados, de modo que se tenga mayor claridad sobre sus propiedades y limitaciones.

Bellagio, 2014

AGRADECIMIENTOS

Como se anota en la introducción, este documento es resultado de un largo itinerario en el que he aprendido mucho con y de diversas personas, aunque ninguna de ellas puede ser considerada responsable de las limitaciones de lo aquí escrito.

La primera mención que debo hacer corresponde a Ana Ponce. Ella fue mi profesora de Estadística durante mis estudios de pregrado en Sociología y, por alguna razón que aún sigo sin entender del todo, me invitó a ser su asistente en 1983. Al hacerlo, me abrió una puerta de desarrollo profesional que yo no esperaba ni sospechaba y que, de alguna forma, llega hasta el día de hoy cuando cometo el atrevimiento de componer este texto.

Quince años después de eso, Juan Fernando Vega recordó con cierta imprudencia esta opción de Ana y me invitó no solo a ser parte del equipo de la Oficina de Planificación Estratégica del Ministerio de Educación en el Perú, sino a que me haga cargo de actividades de análisis, con un importante componente estadístico, y luego de la propia UEE del ministerio. Este hecho es un hito central en el camino que llega a este documento. Ahí tuve la inmensa fortuna de trabajar con, entre otros, Fredy Injoque, Luis Paz, José Rodríguez, Juan Pablo Silva, Ramiro Velaochaga y, por supuesto, el propio Juan Fernando. Por su parte, durante mi tiempo en la UEE del ministerio tuve la suerte de trabajar con un equipo singular de individuos interesados en hacer bien las cosas. Lo logrado en ese tiempo, como lo que yo aprendí, debe mucho a todos ellos, pero siento una particular gratitud por María Luisa Quintana, Juana Huilcamisa, Cristian Pastor, Liliana Miranda y Diego Silva.

Los años en la Orealc de la Unesco Santiago fueron de especial importancia por la presencia de Ivan Castro de Almeida, Rosa Blanco, Marcela Copetta, Ana María Corvalán, Ricardo Hevia, Paula Louzano, Ana Luiza Machado, Lilia Toranzos y Mami Umayahara. Asimismo, trabajar en la región con Paula Darville, Hilda González, Liliana

Mascardi, Mara Pérez, Joel Sherman y Dalila Zarza ha sido una gratísima fuente de aprendizajes.

Juan Cruz Perusia, Ernesto Fernández-Polcuch y Manuel Cardoso fueron excelentes compañeros en la sede del UIS de la Unesco en Montreal. La magnitud de los desafíos que tuve que abordar, especialmente durante mis primeros dos años en Montreal, ha sido una fuente de importantes aprendizajes posibles gracias a la confianza que depositó en mí Hendrik van der Pol.

Las actividades del International Working Group on Assessing and Improving Quality Learning también han sido una fuente importante de *food for thought*. En particular quiero expresar mi gratitud hacia Jean Bernard y Lene Buchert, que iniciaron este proceso desde la sede de la Unesco, aunque ahora ya no laboren para esta institución. Ellas iniciaron esta iniciativa bajo la conducción de Ana Luiza Machado. Debo expresar mi gratitud asimismo a otros miembros del grupo, como Luis Crouch, Margaret Meagher, Alejandro Tiana y Hans Wagemaker.

El trabajo desarrollado a través de la Learning Metrics Task Force ha sido también una importante fuente de debates y de problematización de los temas aquí presentados, de cara a impactar en la formulación de la agenda educativa global para el periodo que se inicia en 2015 con el cierre de los Objetivos de Desarrollo del Milenio y del Plan de Acción de Dakar. Mi gratitud a Rebecca Winthrop, Maribel Soliván, Kate Anderson Simons y Maya Prince.

Hay otros a quienes debo agradecer por su ejemplo contrastante con lo dado por las personas anteriormente mencionadas. Si bien existe una tendencia muy difundida a compilar «buenas prácticas», también sería de mucha utilidad revisar las otras, ya que pueden brindar muchas pistas para mejorar. Errores garrafales, despropósitos descomunales o simples exposiciones de arrogancia, solo comparables a la inopia intelectual que la acompaña, han sido también fuentes de aprendizaje (al menos para mí). Sin embargo, en este caso no vale la pena individualizar a los personajes. En el texto he recurrido en varias ocasiones a argumentos polémicos y a críticas abiertas mencionando ejemplos de situaciones presentes en el quehacer de la generación de información con el propósito de ilustrar mejor los argumentos y conjurar algunos fantasmas que recorren el mundo de la información educativa. En ningún caso, el recurso dialéctico implica un juicio sobre personas e instituciones, pero sí sobre procedimientos que, si bien muchas veces son bienintencionados, considero que no hacen sino debilitar los esfuerzos por construir información útil y relevante.

Hace varios siglos, Tomás de Aquino habló de la necesidad de combinar la intención correcta con la razón cierta, mostrando que no basta con conocer, sino que también es necesario buscar un propósito noble. Hoy en día, sabemos que tener absoluta certeza sobre el conocimiento es algo particularmente peligroso y que es mejor lograr convicciones firmes acompañadas por la apertura a reconocer sus potenciales debilidades

y, por lo mismo, entender que la permanente búsqueda de la verdad y el bien pasan por el diálogo con otros. La búsqueda del entendimiento mutuo por la vía del diálogo es un aprendizaje por el que la humanidad pagó un precio muy alto en el siglo pasado; sin embargo, hay muchos para los que esto todavía es *terra incognita*.

A lo largo de estos años he encontrado demasiados ejemplos muy claros acerca de cómo cierto saber técnico, al verse acompañado de una absoluta imposibilidad para el diálogo y de una desaforada vocación por el control político y/o mercantil de lo que se considera coto privado, lleva ineludiblemente a incluso tornar falso el conocimiento de partida.

Más allá de las experiencias (las buenas y las otras), debo hacer una mención especial a aquellos colegas y amigos que con generosidad me brindaron su valioso tiempo para leer borradores de este texto y hacerme llegar comentarios que, sin duda, lo han enriquecido de manera significativa. Mi gratitud a Juan Cruz Perusia, Manuel Cardoso, Liliana Miranda y Fredy Injoque. Ellos, debo insistir, no son responsables de ninguna de las limitaciones de este documento.

Deseo también expresar mi gratitud a la Fundación Rockefeller, que me apoyó otorgándome la oportunidad de residir en su Centro Bellagio desde finales de febrero hasta finales de marzo de 2014. En la paz y tranquilidad de este espacio de encuentro con académicos, artistas y profesionales de diversas partes del mundo es donde pude terminar de redactar este texto. Mi agradecimiento a los colegas con los que coincidí durante la residencia y que hicieron de los espacios de encuentro momentos personal y profesionalmente muy valiosos.

Unas palabras de profunda y muy especial gratitud a la Universidad del Pacífico, a Felipe Portocarrero (rector hasta junio de 2014) y a Elsa del Castillo (rectora); a David Mayorga (vicerrector), a la directora del Centro de Investigación, Cynthia Sanborn, y a Liuba Kogan, jefa del Departamento Académico de Ciencias Sociales y Políticas. Desde mi regreso al Perú en agosto de 2012, ellos, junto a muchos otros colegas y amigos (en el marco del gran ambiente de trabajo que han creado en la universidad), me han brindado todo su apoyo (incluida una licencia para poder desarrollar la residencia en el Centro Bellagio) y han hecho que desde el primer día sienta a la Universidad del Pacífico como la casa profesional a la que pertenezco con mucha alegría.

Este texto fue originalmente concebido en un momento muy particular de mi vida personal y familiar. Las primeras ideas acerca de cómo estructurar su contenido fueron esbozadas en un tiempo en el que lidiaba con el final de mi tesis doctoral, con responsabilidades particularmente complejas en el trabajo y con una situación familiar marcada por dos cambios sucesivos de país de residencia en un periodo de tiempo muy breve. Como cualquiera puede imaginar, la contribución emocional de mi familia a que hayamos logrado llegar a buen término en todos estos terrenos ha sido inconmensurable. Es difícil encontrar las palabras adecuadas para expresar mi reconocimiento por todo

AGRADECIMIENTOS

esto; sin embargo, lo intento afirmando que no soy capaz de dimensionar adecuadamente lo mucho que han puesto y que, como en el caso de mi tesis doctoral, encuentro que lo bueno que hay en este texto pertenece íntegramente a María del Carmen y a María Victoria, como les pertenece también aquella vocación, sembrada por mis padres, que orienta mis acciones incluyendo escribir este volumen.

Finalmente, mi agradecimiento al Fondo Editorial de la Universidad del Pacífico, por las sugerencias hechas por los revisores anónimos y por creer que la publicación de este texto es una empresa con sentido. Espero sinceramente que la vida que este volumen tenga a partir de su salida de mis manos justifique esa apuesta.

SECCIÓN 1

LOS CIMIENTOS

Los sistemas de información educativa se encuentran en un momento de cambios marcados por las manifiestas insuficiencias de la información estadística «tradicional», así como por las tensiones introducidas por la preocupación creciente por la «calidad», que al parecer solo pueden ser atendidas con cuerpos de evidencia que son producidos por fuera de esa estadística «tradicional». Asimismo, otras fuentes de información usualmente catalogadas como «cualitativas» también juegan un papel creciente como evidencia en los debates públicos sobre la educación.

Abordar este momento de cambios de modo exitoso pasa por revisar los elementos conceptuales, técnicos y analíticos que se encuentran a la base de la configuración de los sistemas de información: la principal tarea es hacerlos explícitos y mejorarlos a la luz de los debates y urgencias del momento actual.

1.1 LA INFORMACIÓN TIENE UN PROPÓSITO

La forma como los seres humanos vemos las cosas es el resultado de múltiples aspectos que incluyen nuestra formación y la manera como vemos nuestras responsabilidades y nuestras preferencias, creencias, prejuicios y conocimientos de la vida diaria.

En el terreno del hablar diario (e incluso en algunos espacios universitarios de no muy buen nivel académico) persiste una de las herencias del racionalismo ilustrado tal y como se concretó en las versiones más simplistas del positivismo en el siglo XIX: estas afirman que la información debe ser «objetiva» (completamente independiente de los sujetos) y, por lo mismo, capaz de hablar por sí misma.

Esta idea tan fuertemente entroncada en nuestras visiones desconoce algo que empezó a ser subrayado con particular fuerza en la filosofía de inicios del siglo XX⁸: **el rol activo del sujeto que conoce en la tarea de conocimiento**. En efecto, las personas conocemos a partir de intenciones, conceptos, categorías y demás herramientas simbólicas profundamente entroncadas en la cultura. La propia idea «dato» es, en primer lugar, una idea humana construida de una determinada manera en un momento histórico dado y no una realidad «natural» independiente de las personas.

Lo anterior no implica la imposibilidad de un conocimiento válido o incluso objetivo, pero la validez y la objetividad no pueden entenderse según dicho esquema simplista propio del siglo XIX. Si bien esta discusión escapa largamente a los límites de este texto⁹, considero necesario mencionar dos cosas que se derivan directamente de tal reflexión: (i) una tarea científica que no se pregunta por sus supuestos metodológicos más profundos

⁸ Claro está que existen innumerables antecedentes de esto en siglos precedentes (como las primeras de las *Tesis sobre Feuerbach* de Marx). Lo que quiero subrayar aquí es el contraste entre la idea comúnmente repetida y originada en el positivismo del siglo XIX y la crítica del mismo, al menos como la inició Husserl y fue desarrollada de modo profundo a lo largo del siglo XX.

⁹ Siguiendo las reflexiones del **realismo crítico** (Bhaskar 1975, 1998) es posible distinguir tres diferentes acepciones de la idea de «objetividad»: (i) afirmaciones o conocimientos que son independientes de los valores o visiones de las personas vis a vis aquellos que son «sesgados» por las personas –de hecho, el acto de conocer es un acto humano intencional y, por lo tanto, está siempre «sesgado», por lo cual este tipo de objetividad es un contrasentido–; (ii) afirmaciones verdaderas vis a vis opiniones (falsas) –esto tiene que ver con la posibilidad de establecer la veracidad o falsedad de una afirmación y este es un tema epistemológico central que, por cierto, va mucho más allá de cuestionar opiniones (una opinión puede ser verdadera o falsa e, independientemente de ello, ser legítima, ya que se corresponde con el ejercicio del derecho al libre pensamiento y expresión)–; y (iii) afirmar la existencia del universo como algo diferente de la experiencia subjetiva del mismo vis a vis visiones relativistas del mundo.

La primera acepción tiene que ver con la ética, la segunda con la epistemología y la tercera corresponde al terreno de la ontología. Sin embargo, el uso corriente confunde estos tres planos en una sola afirmación general.

(incluidos los aspectos ontológicos y epistemológicos y los sistemas de valores y creencias en juego)¹⁰ no merece llamarse científica; y (ii) la contribución de corrientes filosóficas contemporáneas brinda luces nuevas y mucho más profundas para abordar los temas básicos de la tarea cognoscitiva (temas ontológicos y epistemológicos, la posibilidad del naturalismo y la objetividad)¹¹.

La construcción de información pasa por la construcción de un objeto a observar y respecto del cual se ha de producir alguna evidencia o información. Asimismo, supone un proceso de generación de información que implica la acción humana desde diversos planos que incluyen dinámicas institucionales y relaciones de poder material o simbólico.

Lo anterior implica que una primera tarea a asumir a efectos de construir un sistema de información robusto consiste en preguntarse por los propósitos de la información, es decir:

- ¿Cuáles son los **fin**es a los que el sistema de información ha de servir?
- ¿A **quién** ha de servir?

Estas no son preguntas poco trascendentes, todo lo contrario, la propia validez de la información, así como diferentes problemas operativos, pasa por considerar estos factores. Unos ejemplos hipotéticos muy puntuales deben servir al propósito de ilustrar este punto:

- Diálogo entre el ministro de Educación de un país hipotético¹² y la jefa de la UEE de dicho ministerio:

Ministro: ¿Me podría decir cuál es la cobertura actual en educación primaria?

Jefa de la UEE: La cobertura bruta en 2012 [estamos en 2015] era de 117% y la neta de 98%

Ministro: ¿? ... Bueno, ¿y cuántos terminan la primaria?

Jefa de la UEE: Usando la metodología que nos indicó la misión del Banco Mundial en 2009 tenemos que la tasa de conclusión supera el 110% en dieciséis de nuestras dieciocho provincias y en las otras dos se encuentra por encima del 90%.

Ministro: ¿?

¹⁰ Sobre entender la metodología como un complejo de interacciones entre estos factores, véase Dunne *et al.* (2005).

¹¹ Una de las corrientes contemporáneas que con mayor énfasis ha abordado estos temas es el **realismo crítico**. Al respecto, véase Archer *et al.* (2013); Archer y Outhwaite (2004); Bhaskar (1975, 1998).

¹² A lo largo del texto se combina referencias a situaciones e información reales con situaciones o información hipotéticas.

Este ejemplo muestra un problema de desajuste entre las preguntas y las respuestas. Lamentablemente, algunos individuos en los equipos técnicos tienden a minimizarlo argumentando que el ministro «no entiende los temas técnicos». Sin embargo, el problema ilustrado en el ejemplo es más complejo y central a la operación general de los sistemas de información: la información con la que se cuenta no se corresponde con la preocupación de política y, por lo tanto, más allá de cómo se explique, no contribuye a responder las preguntas planteadas, las mismas que, en ambos casos, corresponden a preocupaciones centrales no solo del ministro de este país hipotético, sino de la agenda pública vinculada al ejercicio del derecho a la educación. Más adelante se explicará cuál es el origen de ese desajuste y cómo puede abordarse, quede aquí solo el subrayar el impacto de este en la credibilidad y pertinencia del sistema de información.

Tres ejemplos adicionales permiten ilustrar otros problemas que también subyacen al diálogo entre la información y la política:

- Diálogo entre el ministro de Educación del mismo país y el jefe de la Unidad de Medición de Logros de Aprendizaje de dicho ministerio:

Ministro: ¿Me podría decir cuánto aprenden los niños en educación primaria?

Jefe de la Unidad de Medición de Logros de Aprendizaje: Habiendo corrido 254 modelos matemáticos de alta complejidad podemos concluir que en Lenguaje y Comunicación de sexto grado hay una importante dispersión, de manera tal que, por ejemplo, los estudiantes de las escuelas rurales logran en promedio 354 puntos y los de las escuelas urbanas 573 en la escala con promedio 500 y desviación estándar de 100.

Ministro: ¿?

- La jefa de la UEE explica a la alta dirección del ministerio parte de los problemas de los datos señalando: «Los colegios privados no quieren reportar datos y no tenemos cómo obligarlos».
- En la misma oportunidad, la jefa de la Unidad de Análisis también señala: «Dado que la cobertura es casi universal en primaria y que la población en edad de ingresar a la escuela está disminuyendo debido a cambios demográficos, la matrícula debería estar decreciendo, pero los datos de matrícula no muestran eso, ya que los profesores y directores tienden a exagerar el reporte de matrícula para defender sus puestos de trabajo y el bono salarial que reciben dependiendo del tamaño de la escuela».

Estos dos últimos ejemplos adicionales ilustran un problema vinculado a la interacción entre los sistemas de información y los arreglos institucionales en los que operan. En el primer caso, se sugiere que existen incentivos institucionales para (no) reportar la información (que pueden vincularse a mecanismos de control) y, en el segundo, la existencia de incentivos institucionales que afectan la veracidad de la información reportada. Si los sistemas de información no toman en cuenta los arreglos institucionales de los que hacen parte, difícilmente pueden ser exitosos en su tarea.

De estas reflexiones iniciales derivan los siguientes elementos fundamentales para la robusta operación de los sistemas de información:

- El conocimiento se construye con conceptos y visiones, por lo tanto, es preciso tener claridad sobre los mismos. Un marco conceptual bien desarrollado y explícito es imprescindible para definir qué información es preciso generar, con qué frecuencia y para qué fines. Asumir que un marco conceptual no es necesario solo sirve para ocultar, por la vía del prejuicio (es decir, asumir el punto de vista propio como si fuera «natural» u «obvio» que no requiere demostración), visiones subyacentes a una determinada perspectiva. La necesidad de que los conceptos sean explícitos se vincula a temas básicos relativos a: (i) la solidez académica; (ii) la transparencia y (iii) la posibilidad de comunicar la información.
- Las definiciones básicas del sistema de información corresponden a propósitos. En el caso que nos ocupa, esto tiene que ver en primer lugar con los **finés públicos**. Por lo tanto, los esquemas conceptuales y las herramientas que de ellos se derivan requieren ajustarse a las preguntas de la política a efectos de brindar información pertinente.
- Los propósitos pueden ser diversos, por eso no es posible establecer recetas únicas válidas en todo tiempo y lugar, así como para cualquier agenda de preocupaciones.

Moraleja 1: la labor técnica implicada en el diseño, implementación y gestión de los sistemas de información no podrá ser exitosa si es que no se desarrolla en diálogo con quienes definen las cosas que son relevantes; es decir, en diálogo con la política. Asimismo, es necesario entender que la política no se restringe al mundo de la burocracia estatal o de las autoridades que circunstancialmente ocupan un cargo, sino también al debate público y a la definición social de los asuntos que son de interés público.

Moraleja 2: el desarrollo de los sistemas de información, por tanto, no es un tema exclusivamente «técnico», ni mucho menos uno que se puede abordar solo desde una preocupación centrada exclusivamente en uno de los aspectos de su componente técnico: la informática. Esta moraleja es particularmente importante justamente porque durante los últimos años una visión simplista consistente en afirmar que los sistemas de información se reducen a formularios, computadoras, *software* y redes de comunicación

(en el sentido de redes informáticas) se ha extendido como una plaga y, de hecho, se ha comportado en muchos casos como tal, ya que insume ingentes recursos sin producir nada sostenible (excepto una referencia falaz en el *curriculum vitae* de algún consultor o una buena oportunidad de negocio para proveedores de equipo, *software* o consultorías «especializadas»). Estas historias se vinculan directamente a la explosión de las tecnologías asociadas a la información, lo cual, en algunos casos no muy bien informados, ha tendido a velar la naturaleza sustantiva de los sistemas de información: **estos no son un paquete tecnológico, sino más bien un arreglo institucional.**

1.2 LAS DIMENSIONES QUE CONFORMAN UN SISTEMA DE INFORMACIÓN

Plantear que los sistemas de información pertenecen al universo de los arreglos institucionales pone el énfasis de la reflexión que aquí se presenta en las acciones humanas y en los propósitos asociados a estas. Es desde la comprensión de dichas acciones que un sistema de información puede ser adecuadamente diseñado y exitosamente implementado y desarrollado.

El foco en el accionar humano suscita un primer conjunto de preguntas que deben ser abordadas a la hora de diseñar un sistema de información en educación. Estas preguntas se refieren a las siguientes dimensiones.

1.2.1 LOS DESTINATARIOS DEL SISTEMA

Identificar a los destinatarios del sistema es un paso clave para determinar cuáles son las necesidades de información que es preciso satisfacer.

Quién necesita **qué** información y **para qué** propósitos es el primer agrupamiento de preguntas que se debe abordar. Un sistema de información educativa no se puede construir copiando la estructura de uno que se considere como modelo –por ejemplo, el producido por la Organización para la Cooperación y el Desarrollo Económico (OECD 2013a: 17 y ss.) y publicaciones conexas– ya que este puede corresponder a un conjunto diferente de necesidades de información.

Los agentes que demandan (o que podrían o deberían demandar) información educativa al Estado son diversos y tienen un conjunto diversificado de necesidades y propósitos. La primera tentación que es necesario evitar es la de pretender satisfacer las necesidades de información de todos de la misma manera.

Un sistema **estatal** de información educativa tiene una finalidad básica que toca a su **naturaleza pública**, es decir, vinculada a intereses generales de la colectividad y de la acción pública en particular: **brindar información general sobre el desempeño del sistema educativo tanto a la ciudadanía como a los responsables de conducir las acciones públicas en el terreno educativo.**

Dicho de otra manera, es profundamente deseable que los agentes políticos basen sus opciones y acciones en una discusión sistemática y rigurosa de la evidencia, pero esto es algo que un sistema de información no puede garantizar. Sin embargo, tal sistema sí debe asegurar que dicha evidencia exista y sea accesible a los agentes políticos, incluida la ciudadanía en general, de manera tal que se permita conocer el estado de la educación.

Así, es necesario establecer cuál es el conjunto de información que satisface el ejercicio del derecho ciudadano a conocer el estado de los sistemas educativos en el país. Asimismo, dicho conjunto ha de brindar luces sobre los temas centrales de la acción pública en educación. **La agenda pública educativa es la fuente clave de orientaciones acerca de lo que un sistema de información educativa debe hacer.**

Si bien esta agenda pública tiene que ver con lo que una determinada gestión busca, va más allá de eso. Por lo mismo, si bien muchas veces las personas involucradas en la operación de los sistemas de información tratan de prestar atención a las demandas de quienes formulan la política, estas no agotan el tema. La opinión y los temas planteados por organizaciones civiles son igualmente importantes.

De esta manera, los temas centrales de la política educativa son aquellos que deben ser abordados por el sistema de información en primer lugar. Esto, por ejemplo en América Latina y el Caribe, significa brindar atención a las necesidades de información vinculadas, entre otras cosas: a la expansión de los servicios educativos a la primera infancia, a la conclusión universal de la educación primaria, al acceso creciente y la conclusión de la educación secundaria, a que las personas transiten por el sistema educativo logrando los aprendizajes a los que tienen derecho, a las características de la expansión de la educación terciaria, a que el sistema educativo no reproduzca desigualdades sociales, etc. Es decir, se debe abordar primordialmente temas que se consideren claves desde el punto de vista de la indagación sobre en qué medida se garantizan los derechos a la educación de las personas.

Esto, por otra parte, pone el foco de la acción de los sistemas de información en las políticas y el debate educativo nacional y no en la evaluación o en la gestión de programas de intervención en particular, cuya lógica es diferente y, por lo tanto, plantea otras necesidades de información. Las necesidades de información vinculadas, por ejemplo, a la identificación de brechas de equidad en la conclusión de la educación primaria son diferentes a las requeridas para diseñar o evaluar la efectividad de un programa de distribución de materiales educativos: al ser preguntas de naturaleza diferente, las respuestas también lo son. Por ejemplo, para lo primero podría no ser necesario contar con información con periodicidad anual, mientras que para lo segundo sí lo es (incluso con una periodicidad menor).

Los sistemas de información educativa enfrentan hoy un conjunto de presiones destinadas a que atiendan crecientemente este segundo conjunto de necesidades, aunque no logran responder satisfactoriamente al primero. No se trata, sin embargo, de afirmar que solo algunos propósitos son válidos y otros no. Nadie puede negar la importancia de contar con información para el seguimiento y la evaluación de las intervenciones, pero eso no puede traducirse en desconocer que la necesidad de atender fines diversos puede (si no va acompañada de recursos proporcionales a las tareas) comprometer toda la operación.

Los temas de la agenda de política educativa, por otra parte, también se vinculan a compromisos internacionales de largo aliento que son también responsabilidad de los Estados y se vinculan a principios relacionados a los derechos fundamentales de las personas y a las perspectivas de desarrollo humano¹³. En este sentido, el reporte internacional de datos es también una fuente importante de definición de las necesidades que debe satisfacer un sistema de información educativa.

Al mismo tiempo, existe también una creciente convicción de que los sistemas de información no solo deben atender las necesidades nacionales de información, sino también aquellas que se presentan a nivel subnacional o local, dependiendo del grado efectivo de descentralización de las decisiones. Este es otro elemento clave, ya que los sistemas de información vinculados a las políticas públicas requieren ser sensibles a la estructura según la cual se distribuyen las facultades de acción pública. Dicho de otra forma: (i) brindar información inapropiada (por escasa, tardía u otro criterio que abordaremos más adelante –acápite 1.2.4–) a los gestores hace que el sistema de información sea considerado irrelevante; y (ii) brindar un volumen elevado de información a agentes que no están facultados a tomar decisiones, o siquiera opinar sobre temas sustantivos, puede devenir en cuotas elevadas de frustración o, en casos marcados por el autoritarismo, en una deseable rebeldía.

1.2.2 LOS ELEMENTOS DE INFORMACIÓN QUE HAN DE COMPONER EL SISTEMA

Lo anterior supone una reflexión sobre los elementos de información que han de componer el sistema. Empecemos esta reflexión haciendo una breve referencia a las estadísticas y a los registros administrativos.

En efecto, los sistemas educativos son el objeto central de la acción de los sistemas de información educativa y, como tales, tienen atributos propios de su organización y generan recurrentemente información. Esto lleva a preguntarse acerca de la necesidad de identificar qué tipo de información es requerida en términos de la forma como se produce: ¿se requiere eventos estadísticos específicos cuyo propósito sea generar información? ¿o es posible/deseable/necesario usar el sistema de registros administrativos que hace parte de la operación regular del sistema? Por ejemplo, ¿por qué preguntar (mediante un cuestionario) cuántos alumnos están matriculados en una jurisdicción dada si se supone que existe un registro de matrícula del que se puede extraer esa información?

¹³ Tales compromisos incluyen documentos de carácter vinculante, como el artículo 26 de la *Declaración universal de los derechos humanos* (Naciones Unidas 1948), la *Convención contra la discriminación en la educación* (Unesco 1960) y la *Convención sobre los derechos del niño* (Naciones Unidas 1989), así como declaraciones políticas, como las emanadas de conferencias internacionales, por ejemplo el Foro Mundial de la Educación para Todos (Unesco 2000) o la Conferencia Internacional de Educación de Adultos (Unesco 2009) y los Objetivos de Desarrollo Sostenible (SDG) actualmente en elaboración (<<https://sustainabledevelopment.un.org/topics/sustainabledevelopmentgoals>>).

Los sistemas de información pueden basarse en una combinación eficiente de estos dos tipos de información en tanto y en cuanto su diseño identifique con claridad qué se necesita y cuál es el dispositivo más adecuado para **satisfacer dicha necesidad dada la estructura organizacional e institucional vigente.**

1.2.3 LA DINÁMICA INSTITUCIONAL EXISTENTE

Lo anterior implica preguntarse por la dinámica institucional existente y por los vínculos en el interior del Sistema Estadístico Nacional (SEN; v. gr. la relación entre los organismos sectoriales y las oficinas nacionales de estadística).

Los sistemas de información educativa no son idénticos a la labor que en mayor o menor medida desarrollan las unidades de estadística y de medición de logros de aprendizaje de los ministerios de Educación. Existe un conjunto de aparatos públicos vinculados a esta tarea y, también, principios y normas que favorecen su operación como un SEN. Un eficiente uso de los recursos públicos (es decir, uno que honre el esfuerzo que la colectividad destina a este fin) supone conjugar armónicamente diferentes agentes que generan diversos tipos de información clave para la atención de las necesidades de información educativa. El diseño del sistema de información ha de considerar, por ejemplo, que los ministerios de Economía o Hacienda cuentan con cuentas consolidadas sobre la asignación (presupuesto) y uso (ejecución) de recursos públicos en la educación, que las oficinas nacionales de estadística tienen información de las personas y los hogares que incluyen atributos educativos y que los ministerios de Educación cuentan con el tejido institucional que abarca a las instituciones y programas educativos.

Por otra parte, es preciso considerar las necesidades de los usuarios en los diferentes niveles de toma de decisiones (centro educativo, autoridad local, autoridad subnacional, autoridad nacional) vis a vis sus competencias y facultades.

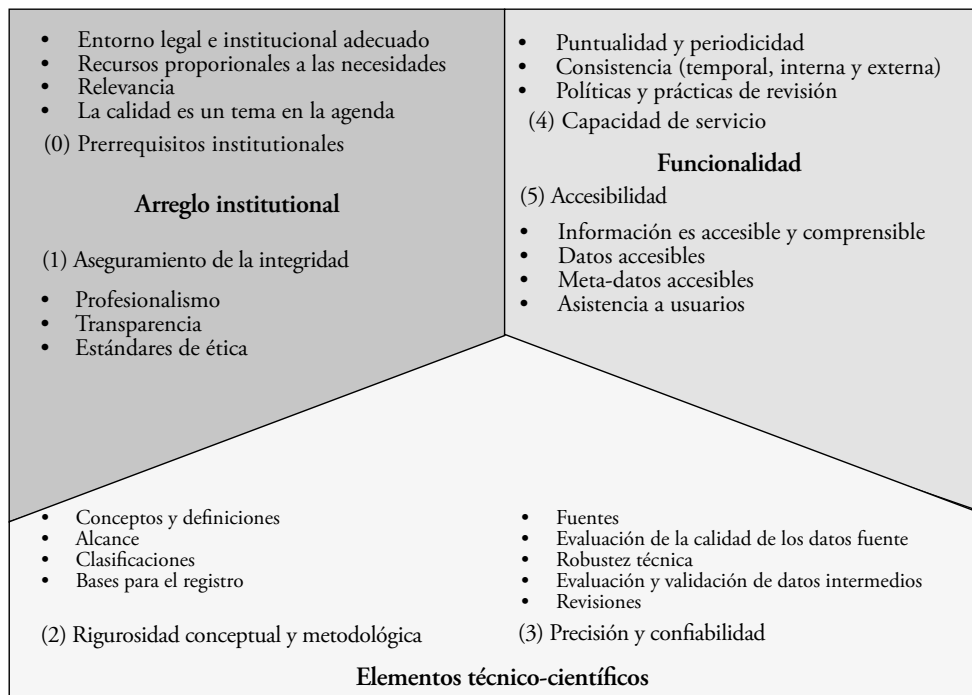
En efecto, este es un tema vital que está a la base de la organización y de la prestación de servicios por parte del sistema de información. Ello supone identificar que las necesidades de información a cada nivel no solo no son idénticas sino que, como ya se anotó, dependen de modo directo de las competencias y facultades de decisión de las que goza cada agente. Así, un desbalance entre la información proporcionada y las competencias del agente puede derivar en la pérdida de legitimidad del sistema de información, sea por exceso (genera expectativas que no pueden ser abordadas, provocando frustración) o por defecto (genera una sensación de insuficiencia).

1.2.4 DIMENSIONES Y ATRIBUTOS BÁSICOS

La calidad de un sistema de información tiene que ver con el conjunto complejo de factores o dimensiones que se detallarán en este acápite. La información es de calidad en la medida en

que la operación del sistema sea adecuada en dichas dimensiones, las cuales son claramente recogidas en el ya mencionado DQAF (IMF 2012)¹⁴, como se muestra en el gráfico 1.

Gráfico 1. Dimensiones de la calidad de los sistemas de información



Nota: adaptación del autor a partir del DQAF. Nótese que el DQAF habla de seis dimensiones que aquí hemos colapsado en tres. Las dimensiones originales del DQAF se corresponden con las subdivisiones numeradas (de 0 a 5) que se han hecho en el interior de cada una de las tres aquí consideradas¹⁵.

Estas dimensiones expresan la interacción entre los ámbitos centrales en los que se construye y desarrolla el sistema de información, a saber: (i) los elementos institucionales que hacen posible la operación del sistema (recursos, normativa, pautas de transparencia, etc.); (ii) los elementos técnico-científicos que dan robustez y confiabilidad al sistema; y (iii) la funcionalidad o capacidad para brindar el servicio.

¹⁴ Acerca del programa de calidad de la información del Fondo Monetario Internacional, resulta conveniente revisar la información disponible en el sitio web especialmente destinado a él (<<http://dsbb.imf.org/Pages/DQRS/Home.aspx>>).

¹⁵ Un uso de tres macrodimensiones coincidente con lo aquí propuesto se puede encontrar en Laliberté *et al.* (2004). Este esquema fue adoptado en 2009 por el UIS a efectos de dar seguimiento a su propio desempeño (Unesco-UIS 2010a).

1.2.4.1 Arreglo institucional

De acuerdo al DQAF, los aspectos vinculados al arreglo institucional hacen referencia tanto a «prerrequisitos» como a ciertas salvaguardas de la integridad.

Los primeros incluyen cuestiones como: (i) la existencia de una normativa adecuada respecto de los procesos de información; (ii) una adecuada dotación de recursos proporcional a las necesidades de los sistemas de información; (iii) la permanente revisión de la relevancia de la información que se genera; y (iv) el lugar que ocupa la preocupación por la calidad de la información en la gestión del sistema de información.

Por su parte, las segundas hacen referencia: (i) al profesionalismo como principio rector de las prácticas; (ii) a la transparencia en la gestión del sistema de información; y (iii) a normas éticas que rigen las políticas y prácticas.

Todos estos son elementos básicos y centrales para asegurar la adecuada operación de los sistemas de información. Ahora bien, para que su lectura y evaluación sea un ejercicio significativo, en el desarrollo de los sistemas de información es preciso tomar en cuenta algunos aspectos específicos adicionales que dotan de mayor contenido a estos principios:

- La necesidad de enraizar los procesos de producción de información en las estructuras organizativas existentes y potenciar las fortalezas de estas: es decir, el anclaje institucional. Esto supone conocer las estructuras formales y la cultura organizacional que está en la base de las acciones de los agentes educativos a efectos de determinar qué es lo posible y cuál podría ser un esquema de operación óptimo, lo cual incluye conocer los calendarios y ciclos de matrícula y otros procesos educativos, además de los esquemas de reporte de información ya existentes.

Por ejemplo, los «resultados a fin de año» suponen definir qué momento debe ser considerado «fin de año» y esto pasa por tener claridad sobre el cierre del año regular y, por ejemplo, la existencia de ciclos de recuperación vacacional.

Asimismo, lo anterior requiere conocer y eventualmente crear o potenciar, en operación conjunta con los responsables, los sistemas de registro administrativo existentes (matrícula, resultados académicos, docentes, directores, instituciones educativas, locales escolares). ¿Cómo mantener un directorio de instituciones y programas educativos si las instancias que autorizan su operación y registran su cierre no logran sostener dichas tareas?

- Las personas y grupos de personas tienen un conjunto no necesariamente armónico de visiones e intereses que contribuyen a definir su accionar y los arreglos normativos e institucionales brindan información a los agentes, los mismos que procesan esa información desde tales visiones e intereses. Si esto no es tomado

en cuenta, un diseño técnicamente «perfecto» puede fracasar en la práctica. A fin de cuentas, un arreglo social **es condición y resultado de lo que la gente hace de modo recursivo en el tiempo** (Giddens 1984).

Las personas tenderán a reportar mejor en la medida en que dicho reporte tenga sentido para ellas, por ejemplo al ser percibido como una oportunidad para obtener un resultado que les resulte valioso. De modo inverso, un proceso de producción de información que resulte una carga onerosa para las personas tiende a afectar de modo negativo el reporte de datos. Asimismo, es preciso prestar atención a la existencia de «incentivos» explícitos o implícitos que puedan tener efectos perversos (usualmente no buscados) sobre el reporte de datos.

Por ejemplo, más allá de la normativa sobre información, ¿existen otras regulaciones que, dadas las características de los agentes y de las instituciones, pueden afectar la información? Un caso típico es el vinculado a la sobredeclaración de matrícula usualmente explicada por un tipo de normativa diferente a la referida y también por prácticas institucionales que incluyen la existencia de bonos salariales por tamaño de matrícula (y este es determinado para dichos efectos a partir del registro administrativo o del reporte estadístico) u otros aspectos que si bien pueden ser deseables en otro plano de acción (en este caso las políticas salariales) tienen un efecto sobre los sistemas de información que estos (si no cuentan con los recursos adecuados para ello) no pueden controlar.

- Idealmente, se requiere precisar tanto las funciones como los perfiles y entorno de trabajo que son demandados a cada nivel de operación del sistema (nivel central, subnacional, local, etc.). Del mismo modo, los elementos de entrenamiento y desarrollo de personal han de ser considerados prestando particular atención a la vinculación entre el quehacer de los sistemas de información tanto con la formulación de políticas como con la investigación educativa.
- Asimismo, es necesario establecer con precisión la relación entre el trabajo técnico asociado a la definición, operación y mejora continua de los sistemas de información y los elementos de soporte de dicho trabajo técnico. En este terreno, resulta crucial establecer con claridad la posición subordinada del soporte informático y evitar la tendencia a definir elementos sustantivos a partir de condicionamientos tecnológicos reales o aparentes.
- Es preciso tener siempre presente que un propósito básico consiste en construir una cadena estable de flujo de información. Los procesos de levantamiento ad hoc de datos que se tornan recurrentes pero asistemáticos terminan minando la credibilidad general del sistema. Particular atención en este terreno merecen las iniciativas de levantamiento de datos asociadas a proyectos financiados por cooperantes o instituciones financieras externas, ya que suelen operar independientemente de la institucionalidad existente y por períodos limitados de tiempo:

se termina el financiamiento y entonces se termina el proyecto y, por tanto, la producción de datos, dejando una situación peor que la inicial, ya que se afecta la credibilidad general.

- Considerar que la responsabilidad del Estado consiste en brindar información sobre el sistema educativo en su conjunto y no solo sobre las instituciones y programas educativos que dependen del ministerio o de otras agencias oficiales.

1.2.4.2 Elementos técnico-científicos

De acuerdo al DQAF, esta dimensión de la calidad de los sistemas de información hace referencia a los siguientes aspectos: (i) los conceptos y definiciones; (ii) el alcance del sistema, es decir, si aborda los aspectos que le corresponden siendo **suficiente** y, por lo tanto, no deja de lado aspectos centrales; (iii) los sistemas de clasificación y definición de sectores; (iv) las bases que se usan para el registro de la información; (v) la adecuación de las fuentes de información; (vi) la revisión periódica de las propiedades de los datos fuente en términos de coberturas y tasas de respuesta, errores de muestreo o no, etc.; (vii) la robustez de las técnicas estadísticas; (viii) la evaluación y validación de los datos incluyendo el análisis de discrepancias con otras fuentes de datos; y (ix) el desarrollo constante de revisiones que permitan mejorar el sistema de modo continuo.

El DQAF discute los primeros cuatro aspectos (a los que agrupa bajo la categoría «rigurosidad conceptual y metodológica») partiendo por asumir como deseable la estandarización internacional de la información. Es decir, define estos aspectos con relación a su grado de ajuste a normas internacionales¹⁶.

Asimismo, torna evidente que la estandarización pasa por la existencia de marcos conceptuales comunes, lo que es siempre una tarea cuya factibilidad e incluso deseabilidad es susceptible de ser debatida o, por lo menos, circunscrita a aspectos específicos. Sin embargo, como ya hemos anotado, la información es construida por personas con visiones, creencias, propósitos e intereses que son disímiles; por lo tanto, asumir un único marco de referencia como universalmente válido es potencialmente muy problemático, a menos que se trate de un marco consensuado y que corresponda a atributos básicos de la condición humana, como es el caso de un enfoque basado en los derechos humanos universales. Esta discusión será retomada y desarrollada en la segunda sección de este volumen.

¹⁶ Lo que es una preocupación fundamental en el caso de la información de cuentas nacionales, que es el terreno primario para el que se creó el DQAF. Sin embargo, esto no tiene por qué funcionar de la misma manera en otros ámbitos. Por ejemplo, la estructura del sistema educativo corresponde con opciones de política, tradiciones, etc. y no tiene por qué esperarse que se ajuste a la Clasificación Internacional Normalizada de Educación (CINE; Unesco-UIS 2013b), por lo que la generación de información ha de seguir la estructura nacional y, luego, ver (para efectos específicos) cómo se traduce en términos de la CINE. (No notar este tipo de situaciones es lo que hace problemática la antes mencionada adaptación que el Banco Mundial y el UIS hicieron del DQAF al sector educación).

1.2.4.3 Funcionalidad

De acuerdo al DQAF, la funcionalidad del sistema de información se vincula a los siguientes aspectos: (i) periodicidad y oportunidad; (ii) consistencia interna y con otros conjuntos de información; (iii) periodicidad y transparencia de las políticas y procedimientos de revisión; (iv) accesibilidad e inteligibilidad de la información; (v) información publicada junto con los detalles técnicos implicados en su producción: conceptos, procedimientos, fuentes, etc. (metadatos); y (vi) asistencia a los usuarios.

Junto a estos elementos básicos, es preciso tomar en cuenta algunos aspectos específicos adicionales que dotan de mayor contenido a los principios que se está considerando:

- Para que la información sea útil tiene que ser relevante. La identificación de las necesidades de los usuarios principales es un primer elemento central para el diseño de un sistema de información.

El diálogo hipotético mencionado al inicio de esta sección (cuando se trata de responder una pregunta de cobertura con una ratio bruta de matrícula) es solo un ejemplo del desajuste entre la información brindada y la necesidad de información. Esto tiene potenciales impactos negativos de muy distinta forma, que incluyen: (i) la merma de la credibilidad de los sistemas de información; o, peor aun, (ii) la toma de decisiones sobre evidencia inapropiada.

- Tan importante como producir la información es hacerla accesible. Mientras la información no esté a disposición de los usuarios es como si no existiese; incluso las consecuencias pueden ser peores si se suscitan sospechas acerca de por qué no es accesible.

El sentido básico del sistema de información es informar, prestar un servicio. Por lo tanto, las herramientas de diseminación de la información deben ser diseñadas con tanto cuidado como las destinadas a producir la información.

Tradicionalmente, las oficinas de estadística han organizado sus labores como un largo ciclo de producción de datos que finalmente son compilados en anuarios estadísticos. Del mismo modo, las unidades de medición de logros académicos suelen organizar sus ciclos culminando con la producción de informes de resultados e informes técnicos de los estudios que conducen. Estos documentos constituyen materiales de referencia de gran utilidad para investigadores y otros usuarios que requieren consultar información; sin embargo, tienen un carácter de referencia y, por lo mismo, si bien su lugar está asegurado en bibliotecas de referencia, su utilidad en los procesos de toma de decisiones resulta, por decir lo menos, dudosa. Asimismo, la existencia de internet hace que la mejor manera de diseminar información de referencia entre estudiantes, investigadores y otros que hacen consultas puntuales de datos esté dada por la existencia de bases de

datos en línea y no mediante la acumulación de papel impreso en estantes, lo que, por otra parte, tiene implicancias negativas sobre el medio ambiente.

La producción de información para **apoyar los procesos de toma de decisión** pone sobre el tapete la necesidad de que las áreas de información dejen de centrar su labor en **contar magnitudes numéricas** (contenido típico de los anuarios estadísticos y de muchos informes de estudios) y pasen a **contar una historia o construir un relato** que encuentra en la evidencia estadística su sustento. De esta manera, los números podrán **contar** en los procesos de toma de decisiones¹⁷.

En efecto, el análisis, incluso el consistente en descripciones básicas de la información, constituye una potente arma de diseminación entre los tomadores de decisiones que requieren del apoyo de los sistemas de información para identificar problemas que deben ser objeto de la política educativa o para medir progresos logrados (o no) en un período dado.

Si bien la cultura anglosajona, muy marcada por el empirismo, ha tendido a mirar la información de un modo más «inocuo» (la expresión «dejar que los datos hablen por sí mismos» es siempre una falacia), ese no es el caso de otras culturas, como la latinoamericana, donde se espera de modo explícito que la información porte un mensaje y, por tanto, tenga un «significado».

Cabe añadir que, de acuerdo a la frase atribuida a Ronald Coase, «si uno tortura a los datos lo suficiente, ellos siempre van a confesar». Es decir, el tratamiento de los datos puede dar lugar tanto a relatos significativos como a arbitrariedades mayúsculas, lo que hace completamente indispensable que se explicita la forma como se trata a la información, dando cuenta, de modo transparente, de los conceptos, supuestos, propósitos y procedimientos que están detrás de cada presentación de información. Esto permitirá a los usuarios reconocer la legitimidad, el valor y el alcance de la información.

Para terminar con este punto relativo a los principios, dimensiones y atributos de un sistema de información, se requiere mencionar los aspectos siguientes.

- Junto a los elementos de transparencia propios del sistema, es importante mantener una relación constante con usuarios clave a efectos de asegurar el ajuste entre las necesidades de información y la operación del sistema, así como entre esta y las expectativas de los usuarios. La falta de balance en cualquiera de estos

¹⁷ Se atribuye a Albert Einstein haber dicho: «Everything that can be counted does not necessarily count; everything that counts cannot necessarily be counted» («Todo lo que puede ser contado [numéricamente] no necesariamente cuenta [importa]; todo lo que cuenta [importa] no necesariamente puede ser contado [numéricamente]» (las palabras entre corchetes han sido añadidas por el autor, ya que la frase implica un juego de palabras relativo al significado del verbo *count* en inglés (enumerar o ser de importancia), lo cual en español resulta más complejo, pues ese término tiene un significado adicional (relatar).

dos elementos atenta contra la credibilidad del sistema y consiguientemente contra su operación regular y sus posibilidades de sostenimiento en el tiempo.

- La entrega de información requiere ir más allá de la sola disseminación de datos y de la inclusión de elementos sustantivos de identificación del potencial y las limitaciones de la información para satisfacer las necesidades de los usuarios. Una labor permanente de «educación de los usuarios» (incluidos los periodistas) para un mejor uso de la información contribuye a mejorar la credibilidad del sistema así como a afinar las demandas que este debe atender.
- Por otra parte y finalmente, el acceso a la información requiere estar claramente normado, de manera que los usuarios no perciban arbitrariedad en él. Para ello, es imprescindible contar con calendarios claros de publicación de información así como con clara documentación sobre la misma (períodos de referencia, definiciones, etc.) y sobre las restricciones al acceso (por ejemplo, las derivadas de normas nacionales sobre el secreto estadístico).

1.2.5 LA CONSTRUCCIÓN DE LAS HERRAMIENTAS CON LAS QUE OPERA EL SISTEMA

Todos los elementos anteriores son la parte sustantiva de una reflexión que, finalmente, derivará en la construcción de los mecanismos o herramientas con los que opera el sistema y a los que se abocarán las siguientes secciones de este documento: tanto sistemas de acopio, carga y recuperación de datos, como un sistema de datos e indicadores que permita reportar adecuadamente la información a los usuarios y, por lo tanto, alcanzar la finalidad para la que existe tal sistema.

Los elementos mencionados (sistemas de acopio, carga y recuperación de datos y sistema de datos e indicadores) son puntos de llegada de la reflexión, y el fracaso de muchos esfuerzos radica en que se consideran el punto de partida. Por ejemplo, prácticamente todas las personas que tienen alguna vinculación con los sistemas de información educativa considerarían que hay que producir ratios brutas de matrícula (¿quién cuestionaría esto?)¹⁸ y, por lo tanto, se produce información que permite su cálculo y se incluye en los reportes de datos. Sin embargo, ¿cuántas pueden definir con precisión cuáles son las necesidades de información y, consiguientemente, los temas de política educativa que encuentran en las ratios brutas de matrícula un elemento de evidencia relevante, clave o imprescindible?

¹⁸ Debe notarse que la Unesco Santiago (2008) es un ejemplo de descripción de la situación educativa que no recurre a una sola ratio bruta de matrícula. En el único caso en el que hubiese cabido usarlas (matrícula en educación terciaria), los autores optaron por calcular ratios de matrícula sobre el total de la población (por cada cien mil habitantes) que dan cuenta del mismo fenómeno (tamaño relativo de la matrícula) sin inducir a ningún equívoco en términos de la discusión sobre «cobertura» (Unesco Santiago 2008: 93 y ss.).

La creación de instrumentos para producir información debe atender, al menos, a los siguientes aspectos:

- **¿Qué información se debe producir?** Tanto información básica o de referencia que solo el sistema de información del Estado puede generar (número de estudiantes, docentes, instituciones, locales, programas educativos) como información más elaborada que apunte a dar cuenta de temas sustantivos (tales como: cobertura, atraso escolar, conclusión, logro académico, progreso hacia metas educativas, etc.).

Los temas específicos que respondan a lo anterior se han de traducir en un sistema de datos e indicadores cuyo nexo con la atención de las necesidades políticas de información debe ser explícito: el indicador x (o un conjunto de indicadores $\{a, b, c, \dots\}$) sirve para describir el fenómeno o monitorear la meta y , dado que da cuenta de los atributos α, β, γ . Esta identificación documentada de un sistema de datos e indicadores permite definir qué información es necesaria.

- **¿Cuáles son los instrumentos y/o dispositivos necesarios para producir la información?** Se debe identificar (primero) las fuentes secundarias que puedan compilarse y usarse, así como aquello (luego) que supone un esfuerzo de generación primaria de datos. Asimismo, se requiere revisar la consistencia entre las fuentes y los potenciales enlaces entre estas (por ejemplo, entre los directorios de instituciones educativas y la información sobre contratos docentes o la matrícula), prestando atención a temas centrales, como los períodos de referencia. Por ejemplo, cuando se combina información de matrícula por edades y de población, el desfase entre períodos de referencia puede producir verdaderos desastres.
- **¿Qué supone el desarrollo de instrumentos de generación primaria de datos?** Considerar los aspectos de registro, carga y recuperación de datos con los consiguientes procesos de crítica y consistencia, así como de manejo de omisiones e inconsistencias.

1.3 ¿QUÉ CONTAR PARA QUE LA INFORMACIÓN CUENTE?

El grado de éxito con el que se responda a esta pregunta depende de modo directo de la comprensión de la agenda de política educativa y, por lo mismo, será abordada de un modo más detallado en la segunda sección. Sin embargo, solo a manera de ilustración y considerando algunos temas recurrentes, podemos mencionar en forma preliminar los siguientes elementos generales:

- **Información sobre los atributos educativos de la población y/o los hogares.** Esta información es usualmente producida por las oficinas nacionales de estadística a partir de censos poblacionales o encuestas de hogares o personas, los que incluyen información sobre los niveles de escolaridad de la población (en algunos casos llamada «logro educativo» y expresada como años de estudio logrados o como último nivel o grado aprobado); también surge de estudios especializados sobre niveles de competencias de alfabetismo de la población y existe la posibilidad de vincular esta información a otros atributos de las personas y/o sus hogares, como los niveles de ingreso, condición de pobreza, área de residencia, pertenencia étnica, densidad del entorno letrado, estructura familiar, etc.¹⁹
- **Información sobre el comportamiento de las personas y las familias respecto de la educación.** Es decir, sobre temas como el gasto de las familias en educación, preferencias, trayectorias intergeneracionales, etc. Esta información es usualmente producida por las oficinas nacionales de estadística a partir de encuestas de hogares.
- **Información sobre el financiamiento (público y privado) de la educación.** Es usualmente producida por los ministerios de Economía o Hacienda mediante los registros administrativos sobre asignación y ejecución presupuestal de recursos públicos (incluidas donaciones al Estado); también por las oficinas nacionales de estadística, mediante encuestas de ingresos y gastos de las empresas y familias, para dar cuenta del gasto privado.
- **Información sobre la operación del sistema educativo.** Se trata de información generada tanto por los ministerios de Educación como por aquellos responsables de la educación superior. Incluye información usual como matrícula, logros académicos de los estudiantes, etc.

¹⁹ Este es el propósito del programa LAMP del UIS de la Unesco. Véase Unesco-UIS (2009b).

- **Información sobre las políticas educativas de un país y las normas, regulaciones y orientaciones vigentes.** Usualmente esta información es compilada (aunque sea de modo asistemático) por los ministerios de Educación como responsables a nivel global del monitoreo y de la orientación del sistema educativo nacional.

Esta enumeración de grandes áreas de información solo busca ser ilustrativa y hacer evidente dos temas: uno, los agentes involucrados en la producción de información educativa son múltiples y trascienden con creces los límites usuales de las unidades de estadística o de medición de logros académicos en los ministerios; y dos, la naturaleza de la información que puede ser necesaria para informar los procesos de debate, definición y ejecución de las políticas públicas también trasciende los límites usuales de lo que conocemos como «estadística educativa».

Asimismo, es importante destacar el rol de liderazgo que toca a los ministerios de Educación en la formulación y conducción de las políticas públicas en lo que atañe a la educación, lo que hace evidente que las unidades de información dentro de estos ministerios deben considerar como su área de interés al conjunto amplio de información que se requiere para estos fines.

Por su parte, la dimensión internacional de los sistemas de información educativa introduce elementos que tienen que ver tanto con la comparabilidad como con la existencia de objetivos globalmente reconocidos como claves en materia educativa.

La comparabilidad internacional de datos debe ser tenida en cuenta a efectos de facilitar, desde el diseño del sistema de información, la transferencia de la información nacional a esquemas de clasificación y definiciones internacionalmente utilizadas (principalmente la Clasificación Internacional Normalizada de Educación [CINE]), en el primer caso; o la definición de, por ejemplo, «repetidor», en el segundo. Esto se debe hacer así a efectos de asegurar que la información que el país reporta sea consistente con dichas clasificaciones y definiciones y para documentar cabalmente las transformaciones o adaptaciones, de modo que sean claramente identificables las razones por las que pueden presentarse diferencias entre las cifras usadas nacionalmente y aquellas contenidas en bases de datos internacionales.

1.3.1 DEFINIR Y DIFERENCIAR CONCEPTOS BÁSICOS

Algunos conceptos fundamentales deben ser claramente definidos y diferenciados como otro de los pasos iniciales requeridos para construir un sistema de información robusto.

Típicamente, un primer conjunto de elementos a este nivel involucra lo siguiente:

Institución, programa y local educativo

Cada uno de estos es un elemento distinto con diferentes propiedades susceptibles de ser registradas y medidas. Un manejo inapropiado de las diferencias puede llevar a problemas sustantivos en la operación regular de los sistemas de información.

Una **institución es una unidad de gestión**. Dependiendo de la normativa de cada país, una institución educativa puede, o no, ofrecer más de un programa educativo (que puede ser de distinto nivel: preescolar, primario, etc.) y operar en uno o más locales, que puede, o no, compartir con otras instituciones (educativas o no). Por lo mismo, es preciso distinguir cada uno de estos elementos.

Dicha institución puede ofrecer uno o más programas educativos. Los **programas educativos** son unidades educativas con un *currículum* y una estructura educacional propios que corresponden a un **nivel** (preescolar, primario, etc.) y se ofrecen de acuerdo a una **modalidad** (presencial, a distancia, para menores, para adultos, etc.).

Los **locales** son las instalaciones físicas usadas por una o más instituciones para ofrecer sus servicios o programas educativos. Asimismo, una única institución puede hacer uso de uno o más locales (próximos o en diferentes circunscripciones cuando opera anexos, satélites o sedes diferenciadas de duración limitada o no).

Cada uno de estos conceptos da cuenta de una entidad distinta. Si el sistema de información no diferencia entre ellos, puede llegar a tener serios problemas especialmente en la construcción de series temporales de datos.

Lamentablemente, en ocasiones estos problemas se originan en una combinación de ignorancia con su típica compañera, la soberbia, que lleva por ejemplo a sugerir la necesidad de crear códigos «inteligentes» para la identificación de los centros educativos, códigos que serían el identificador clave a usar en las bases de datos.

Estos códigos «inteligentes» suelen, por ejemplo, asociar la institución educativa a su actual jurisdicción geográfica, de modo que la institución recibe un código de la forma DDPPMMGGNN0000, que no es sino un identificador de departamento (DD) o Estado, dependiendo de la división geográfica del país, provincia (PP) y municipio (MM), seguido por un identificador de gestión (GG), uno de nivel educativo (NN) y un correlativo (0000). Pero, ¿qué pasa si la institución se muda a otro municipio o provincia?, ¿si cambia la demarcación territorial?, ¿si la institución ofrece más de un nivel educativo? En cualquiera de estos casos, la respuesta tenderá a ser que se requiere de un nuevo código para la misma institución, lo que no hace sino evidente que el sistema de codificación (o quien lo diseñó) no es tan «inteligente» como se pretendía.

Por lo anterior, un registro de instituciones educativas debe ser distinto de (aunque se debe vincular a) un registro de programas educativos y a uno de locales escolares (o usados para fines educativos).

Asimismo, cabe anotar también que las instituciones educativas como unidades de gestión pueden presentar diferentes afiliaciones institucionales. Por ejemplo, ser parte

de una asociación de escuelas independientes, del grupo que ofrece el programa de bachillerato internacional o de una red de escuelas con una gestión común (como Fe y Alegría, que opera en diversos países); pueden tener una filiación religiosa, contar con un promotor o propietario común (empresa, fundación, etc.) o ser parte de una franquicia; etc. Estas afiliaciones son importantes para entender, por ejemplo, la presencia de educación confesional, el crecimiento de la educación comunitaria, etc.

Niveles educativos

Los niveles educativos son una propiedad que corresponde a un programa educativo (no a una institución).

La CINE en su revisión 2011²⁰ considera nueve niveles educativos, a saber: (i) preescolar; (ii) primaria; (iii) baja secundaria; (iv) alta secundaria; (v) postsecundaria no terciaria; (vi) terciaria de ciclo corto; (vii) terciaria conducente a un primer grado académico; (viii) maestría; y (ix) doctorado o investigación avanzada.

Sin embargo, cada país define su propia estructura educativa en niveles según sus propias consideraciones. Así, por ejemplo, se tiene que un grupo grande de países ha establecido una Educación Básica (que usualmente reúne lo que la CINE conoce como primaria y baja secundaria, aunque en algunos casos, puede incluir parte del preescolar), asimismo hay países que no cuentan con una educación postsecundaria no terciaria. Del mismo modo, en muchos casos de América Latina y el Caribe, la educación de la primera infancia tiene un alcance mucho más amplio de lo que contemplaba la CINE 97 (ya que se inicia en edades más tempranas e incluye modalidades no previstas por ella)²¹.

Independientemente de lo anterior, el registro de programas educativos implica conocer el nivel educativo al que corresponde la oferta de los mismos. Por su parte, este se define con relación al contenido educativo previsto del programa, independientemente de la forma como este contenido se gestiona o brinda (modalidades).

Modalidades

Las modalidades son una propiedad de un programa educativo (no de una institución o de un nivel dado).

Si bien no existe una definición comúnmente aceptada de lo que es una modalidad, es preciso considerar que las propiedades de los programas educativos incluyen aspectos

²⁰ La primera versión de la CINE fue aprobada en 1975 por la Conferencia Internacional de Educación y adoptada por la Unesco en la Conferencia General de 1978. En 1997 la Conferencia General de la Unesco aprobó la revisión usada hasta 2013 (Unesco 1999a). La 34 Conferencia General de la Unesco (2007) inició un proceso de estudio y consulta que ha dado lugar a la revisión 2011 de la CINE (Unesco-UIS 2013a), que viene siendo usada desde 2014.

²¹ Uno de los aportes latinoamericanos a la revisión 2011 consiste justamente en haber ampliado la visión de la educación de la primera infancia.

relativos a: (i) el modo de prestación del servicio (presencial; a distancia; combinado; programas no-graduados; etc.); y (ii) la población objetivo (menores, jóvenes y adultos; personas con necesidades educativas especiales; educación bilingüe; etc.).

Así, una **institución educativa** podría por ejemplo ofrecer los siguientes **programas**: (i) educación preescolar escolarizada; (ii) primaria de menores; (iii) primaria de adultos presencial; (iv) primaria de adultos a distancia; (v) secundaria de menores; (vi) secundaria de adultos presencial. Un ejemplo más concreto sería el de una institución especializada en la **modalidad** de la educación de personas con impedimentos visuales que puede ofrecer un programa de educación primaria y otro de educación secundaria.

Las modalidades también varían en función de la organización de los tiempos de instrucción. Así un programa de un nivel dado puede perfectamente brindarse usando diversos esquemas de organización del tiempo (tiempo completo, tiempo parcial, ciclos acelerados, etc.).

Estudiantes, individuos, plazas; tiempo completo, tiempo parcial

Las preguntas que se formulan con relación al «número de estudiantes» o a diferentes atributos de estos pueden tener muy diferente naturaleza y propósitos, por lo que potencialmente son muy distintas.

Por ejemplo, preguntar por el número de personas que el sistema educativo atiende (cobertura) hace referencia a **individuos** en una población dada. Así, la respuesta pasa por contar el número de tales individuos que satisfacen la condición de ser atendidos por el sistema y que son parte del conjunto de población del que se quiere hablar.

Por su parte, la pregunta sobre el número de estudiantes que una institución educativa (o un grupo de instituciones educativas) puede atender (lo que a veces es llamado «capacidad de oferta»), en realidad, hace referencia al número de **plazas o vacantes** disponibles en dicha institución (o grupo de instituciones), independientemente de si están ocupadas o no. En este caso, se está midiendo la capacidad instalada para atender, pregunta cuya respuesta no se corresponde con un número de individuos.

De igual forma, si se está haciendo un estudio de necesidades de inversión y se quiere determinar **montos unitarios** (por alumno), se requiere trabajar con unidades equivalentes de tiempo completo. Los costos pueden ser distintos si los estudiantes son de tiempo parcial y, por lo mismo, los números de estudiantes deben transformarse a una unidad de costo equivalente. Esto tampoco es igual al número de individuos, ya que una persona individual puede contar como una fracción (si es estudiante a tiempo parcial). Esto es particularmente claro en la educación terciaria, ya que en muchos casos la carga lectiva de los estudiantes es medida en créditos y cada estudiante puede estar matriculado en un número de créditos diferente en cada período lectivo.

Docentes, individuos, plazas, contratos; tiempo completo, tiempo parcial

Con los atributos referidos a los docentes sucede algo similar que con los estudiantes.

Si lo que interesa es conocer el número de personas que se desempeñan como docentes, la respuesta pasa por un conteo de **individuos**. Si, por otra parte, lo que se requiere es conocer los costos de inversión en docentes, por ejemplo, por un número dado de estudiantes, usualmente se opera con **unidades de tiempo completo** (tanto de estudiantes como de docentes).

De igual forma, los puestos o plazas, así como los contratos de docentes no son necesariamente equivalentes al número de personas. Una **plaza**, puesto o ítem presupuestal es usualmente una unidad presupuestal con la que, dependiendo de las normas de cada país, se hace uno o más **contratos** a uno o más individuos.

En ambos casos (docentes y estudiantes), es necesario definir claramente el alcance de las preguntas para no cometer errores de sobre o subestimación, que se originan, por ejemplo, en potenciales dobles cuentas (si se quiere el número de individuos que son docentes y se responde con el número de contratos).

Lo anterior muestra que incluso las preguntas aparentemente más triviales que se puede hacer a un sistema de información (cuántas escuelas, cuántos alumnos, cuántos docentes) obligan a un trabajo técnico riguroso que permita entender lo que se está preguntando y, por lo mismo, a construir procesos de generación de información que hagan posible responder dichas preguntas de forma adecuada.

1.4 DIRECTORIOS Y TASAS DE RESPUESTA

Los directorios son un elemento central de los sistemas de información, en tanto: (i) permiten determinar la omisión en eventos censales; y (ii) son el marco de los eventos muestrales.

Un directorio actualizado permite conocer la omisión de los eventos censales, ya que hace posible contar con un punto de referencia claro acerca de las unidades sobre las que se debería estar reportando información. Conocer la omisión es de fundamental importancia tanto para evaluar la gestión de la información (si los niveles de omisión se encuentran dentro de límites manejables), como para estar al tanto de potenciales sesgos en el reporte de información (si la omisión se concentra en unidades con determinadas características –por ejemplo, instituciones educativas de gestión privada–) y, finalmente, también para poder estimar valores totales.

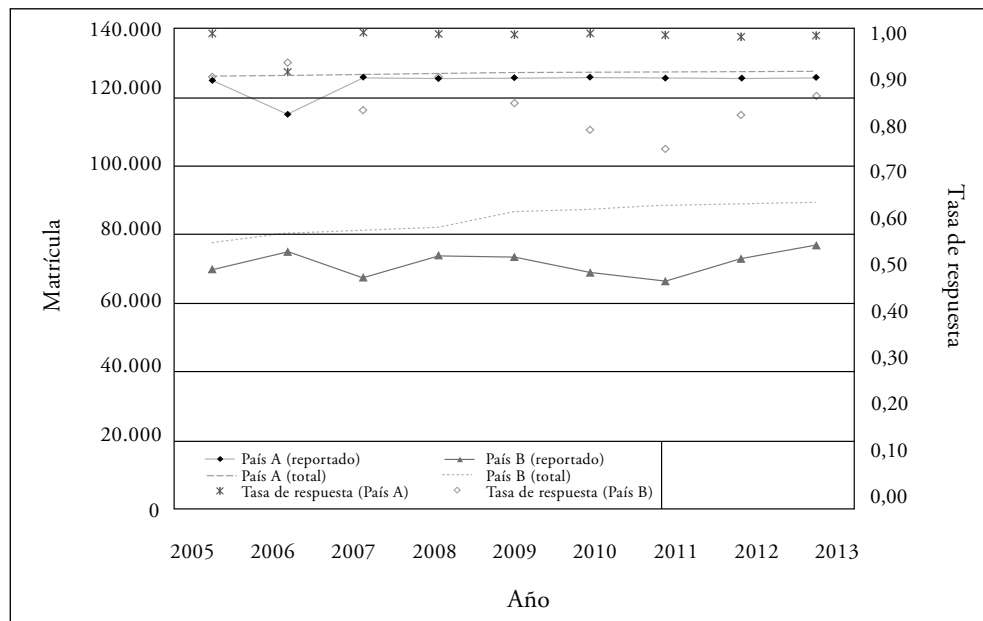
Los censos de población nos permiten ilustrar este punto. La información de tales censos usualmente se publica haciendo referencia a dos totales: (i) población total de un país y (ii) población nominalmente censada. La segunda corresponde a los datos efectivamente obtenidos en un evento censal, mientras que la primera permite corregir la omisión (es decir la población total es igual a la población nominalmente censada más los omisos). Esta corrección tiene la forma de una estimación. Dependiendo de los propósitos, se trabaja o bien con la información total, o bien con la población nominalmente censada.

En el caso de un censo escolar que cuenta la matrícula y que es realizado de modo anual, no hacer esta diferenciación puede tener efectos negativos de mucha importancia. El gráfico 2 ilustra estos problemas.

En el gráfico 2, en el caso del País A se aprecia una matrícula reportada constante a lo largo de los años, excepto en 2006, donde se observa una caída. Si la matrícula **reportada** fuese presentada como **total**, inmediatamente dispararía la pregunta acerca de qué sucedió en 2006 que la matrícula disminuyó. Sin embargo, en este ejemplo hemos podido contar con las tasas de respuesta de los censos de matrícula de cada año (estimadas usando los directorios correspondientes) y resulta que lo observado en 2006 obedece a una caída en las tasas de respuesta y no en la matrícula. En efecto, en el País A, las tasas de respuesta son siempre cercanas a 1, excepto en 2006, cuando cayeron a cerca de 0,9. Al contar con las tasas de respuesta, se puede estimar la matrícula total y ella muestra la estabilidad esperada. De esta manera, lo que parecía un problema con la matrícula en realidad lo es con el sistema de la información en un año dado (cuando un consultor externo forzó la introducción de cambios que supuestamente debían mejorar

los tiempos de acopio de datos), problema que, felizmente, se superó en el año siguiente (cuando no se renovó el contrato de dicho consultor).

Gráfico 2. Matrícula total y reportada en educación primaria, países A y B, 2005-2013



Nota: información hipotética creada para efectos de ilustración.

En el caso del País B la matrícula reportada muestra un carácter errático, lo que debería llamar la atención de todos los interesados. Al contar con las tasas de respuesta, se observa claramente que la matrícula reportada no está representando adecuadamente la tendencia real al ascenso de tal matrícula, sino los patrones erráticos de las tasas de respuesta. En efecto, estas varían de modo asistemático y muestran una importante caída entre los años 2009 y 2011 (cuando el consultor, que estuvo en el País A previamente, logró una importante influencia en el País B).

Estos ejemplos permiten ilustrar la importancia de contar con buenos directorios y distinguir la información reportada de la total, asumiendo con toda seriedad la tarea de estimar las omisiones. En la sección tercera se presentarán algunas ideas acerca de los procedimientos para lidiar con la omisión y hacer estimaciones.

Ahora bien, es posible construir directorios para cada una de las unidades de observación: (i) instituciones educativas; (ii) programas educativos; (iii) docentes; y (iv) estudiantes. Sin embargo, la elaboración de directorios no es necesariamente una tarea fácil y su mantenimiento también presenta desafíos. Así, las preguntas centrales son:

¿cuáles son los directorios que se necesitan? y ¿cómo es posible crearlos y darles un adecuado mantenimiento?

Un directorio de instituciones educativas es un elemento básico que, en un escenario esperable, puede ser construido a partir de las acciones administrativas que regulan la operación de dichas instituciones: autorizaciones de funcionamiento (resoluciones o mandatos de creación o apertura) y resoluciones de cierre temporal o definitivo. Es de esperar que la autoridad educativa cuente con mecanismos claros para estas acciones administrativas e, idealmente, que lleve un registro de las mismas. Ese **registro administrativo** es la columna vertebral del directorio.

Sin embargo, este escenario puede no observarse por dos circunstancias: (i) las autoridades educativas operan en un esquema descentralizado y, por lo tanto, no existe una sola fuente de información, sino potencialmente varias, que no siempre operan con el mismo nivel de eficacia y de modos consistentes entre sí; y/o (ii) la gestión administrativa nunca encontró una razón para hacer un mantenimiento ordenado de estas acciones en un directorio o en una base de datos y, por lo tanto, existe un universo disperso y potencialmente caótico que contiene el registro de los actos administrativos con una forma de difícil recuperación para un sistema de información.

En casos como estos es posible que las unidades de información se vean precisadas a tomar la iniciativa en la construcción del directorio; sin embargo, es necesario que esta no sea una iniciativa aislada a cargo exclusivamente de las unidades de información, sino que sea asimilada a la gestión administrativa para que se sostenga en el tiempo de la manera más eficiente posible.

Ahora bien, ¿cómo puede una unidad de información crear un directorio de instituciones educativas? La solución pragmática suele estar dada con el uso de una combinación de dos elementos como punto de partida: (i) el levantamiento de información básica sobre las instituciones educativas en un evento censal; y (ii) la verificación de la información (en busca de potenciales omisiones) en un trabajo conjunto con las unidades encargadas de la gestión local (por ejemplo, autoridades educativas a nivel municipal, distrital o provincial, dependiendo de la organización propia del país). Así, un evento censal da un punto de partida que luego se suplementa mediante la consulta a los órganos de gestión. Este es un trabajo arduo y que introduce las diferencias en la efectividad de los distintos órganos de la gestión como variable que explica el grado de éxito de la empresa.

Asumiendo que este esfuerzo dio un resultado lo suficientemente bueno, la pregunta que sigue es: ¿cómo mantener actualizado dicho directorio? Si la construcción del directorio inicial resulta una tarea compleja y con probabilidades de éxito que pueden ser muy variables, el mantenimiento es aún más complejo, a menos que se asegure su eslabonamiento con el proceso administrativo de apertura y cierre de instituciones. Esto supone que el trabajo cooperativo con las áreas de gestión debe mantenerse, lo cual solo es sostenible en el largo plazo si estas también se ven beneficiadas en su trabajo por dicha

tarea, lo que nos lleva de regreso a la necesidad de enraizar los procesos en la propia gestión. Asimismo, este enraizamiento puede requerir acciones normativas específicas que lo hagan posible (incluyendo potenciales esquemas de sanción a las unidades que no reporten información oportunamente –pero como cualquier esquema sancionador, este solo tiene sentido si es que la administración cuenta con la capacidad para detectar el incumplimiento y ejecutar la sanción–) y acciones informáticas de soporte que lo viabilicen.

En cualquier caso, la actualización de los directorios no consiste en reemplazar la versión actual por la precedente, sino en mantener un registro histórico. Es decir, se debe contar con un directorio padrón para cada ejercicio o año lectivo y asegurar que los directorios de los distintos períodos sean compatibles a efectos de permitir la construcción de series históricas. Alternativamente, se puede mantener un único directorio con información registrada en cada año.

Estas consideraciones relativas a los directorios de instituciones se aplican de igual forma a los programas, dado que los actos administrativos que regulan la apertura o cierre de instituciones y de programas suelen conjugarse. Una institución está asociada a un número de programas y cada programa a una institución.

Los directorios de docentes y estudiantes tienen una naturaleza diferente por referirse a registros de personas. Mientras que el registro de instituciones que operan en un ámbito de interés público como la educación resulta relativamente más simple, el registro de información individual supone otras normas sobre la reserva y la protección de la intimidad que deben ser tomadas en cuenta y que, por lo mismo, plantean otros desafíos éticos y técnicos vinculados a la protección de la información²².

Dada esta situación, la primera pregunta que es siempre de aplicación aquí se hace aún más perentoria: ¿es realmente necesario componer un directorio de personas?, pregunta que se puede formular también de otras formas: ¿es realmente necesario contar con un registro de información individualizado que permita la identificación de las personas? ¿de verdad alguien va a usar esa información para algún propósito significativo, o el tema solo está en la agenda porque a alguien se le ocurrió que es técnicamente posible e «interesante»?

Asumiendo que la respuesta a esa primera pregunta es positiva, existe un sinnúmero de consideraciones que se refieren a la construcción y mantenimiento de este tipo de directorios. De estas, debemos enumerar las siguientes: (i) ¿cómo se vinculan estos directorios a los registros nacionales de identidad?; (ii) ¿es posible construir estos directorios a partir de registros administrativos existentes, como las actas de matrícula para los estudiantes o los contratos de los docentes?; (iii) ¿es esto posible o deseable solo con relación a las instituciones y programas educativos gestionados por el sector estatal?

²² Sobre este tema, véase el capítulo XII del *Manual de organización estadística* de la UNSD (2004).

1.5 SISTEMAS DE CAPTURA DE DATOS

¿Cómo acopiar la información? Para responder a esta pregunta es preciso considerar lo siguiente: (i) ¿qué información se necesita?; (ii) ¿con qué periodicidad?; (iii) ¿hay alguien que ya la esté produciendo?

Este acápite está destinado a abordar temas relativos a la producción **primaria** de información. Es decir, no considera aspectos vinculados al manejo de fuentes secundarias (información generada por «otros» mediante sus particulares procedimientos y mecanismos y atendiendo a sus propias necesidades); aunque este segundo grupo de información es muy importante no solo para complementar la información que un sistema propio produce, sino también para el análisis de la validez de los propios datos.

1.5.1 ESTADÍSTICAS Y REGISTROS

La generación primaria de datos estadísticos se basa en dos tipos de procedimientos principales: (i) eventos especiales de levantamiento censal o muestral de información para propósitos estadísticos; (ii) captura de información estadística a partir de registros creados para propósitos no estadísticos sino de gestión.

Un ejemplo puede ilustrar la diferencia entre estos dos procedimientos. El número total de personas matriculadas en un programa educativo dado puede ser obtenido: (i) mediante un cuestionario administrado a la persona responsable de dicho programa (un cuestionario diseñado para fines estadísticos y administrado según las normas propias de los eventos estadísticos); o (ii) mediante la captura de la información de las actas de matrícula de dicho programa. Es cierto que las actas de matrícula existen para propósitos no estadísticos, pero la captura (total o parcial) de su contenido puede permitir generar información estadística. Adicionalmente, es posible pensar en un evento estadístico que requiere que los respondientes capturen (al menos en teoría) la información de los registros administrativos existentes.

Cada una de estas formas de operar tiene sus propias particularidades y se regula por su propio conjunto de reglas. Los eventos estadísticos se regulan por la normativa de los sistemas estadísticos nacionales, mientras que la captura de información de registros administrativos se regula por la normativa propia de cada uno de estos.

Los instrumentos de generación de información son diferentes en cada caso. Mientras que en el caso de los eventos estadísticos se utilizan instrumentos (formularios, cuestionarios, pruebas) especialmente diseñados para lograr los propósitos de los sistemas de

información, en el caso de los registros administrativos es necesario crear un dispositivo de enlace que permita la captura de la información ya existente.

En un escenario típico, ambos procedimientos (estadísticas y registros administrativos), así como su combinación, se encuentran presentes.

En los casos en los que los eventos estadísticos requieren que los respondientes usen los registros administrativos como fuente para sus respuestas, es particularmente importante que el diseño de los instrumentos de recolección de información esté alineado con los instrumentos administrativos existentes en términos conceptuales, de nomenclatura y de períodos de referencia. Por ejemplo, la captura de la matrícula a partir de las actas de matrícula requiere tener en cuenta el momento en el que dichas actas se «cierran» o la manera como en estas se registran las edades de los estudiantes (si es en años cumplidos, teniendo en cuenta cuál es el período de referencia para esto).

La aplicación de formularios puede, por su parte, darse de modo autoadministrado o mediante una administración ejecutada por la unidad organizacional a cargo de las tareas de información. El primer caso es el correspondiente a formularios que son enviados para que las personas a cargo los diligencien y los retornen de acuerdo a determinadas instrucciones. El segundo caso ocurre cuando personal trabajando para la unidad de información educativa visita, por ejemplo, las instituciones educativas y administra los instrumentos.

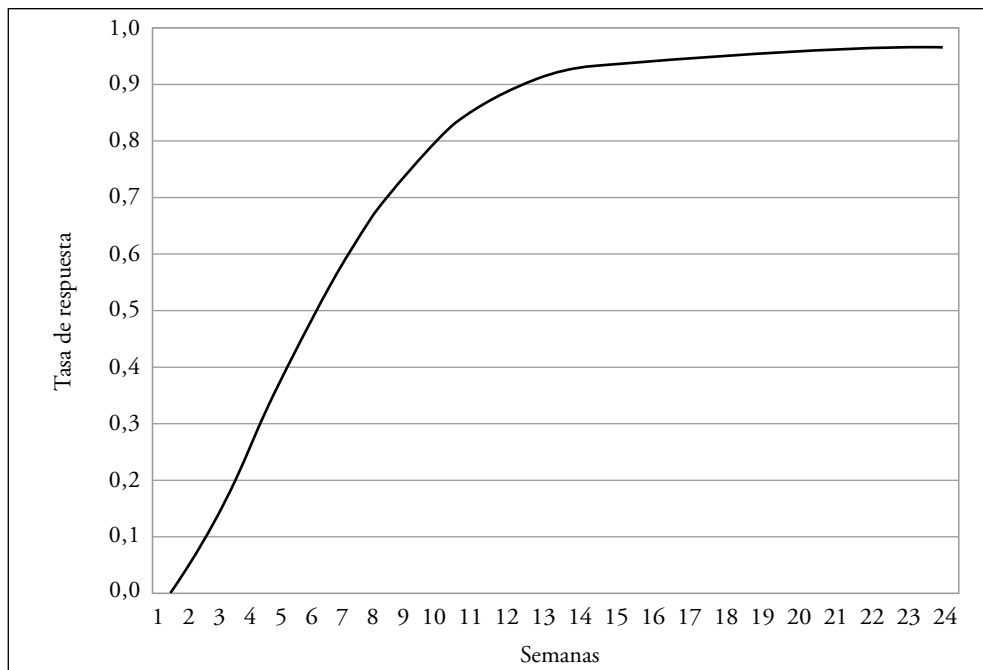
En ambos casos pueden surgir problemas. El primero plantea principalmente aquellos de control de la calidad de las respuestas y de potencial no respuesta; mientras que el segundo tiene sobre todo dificultades operacionales, de costos y de potenciales rechazos.

1.5.2 FECHAS DE CIERRE DEL LEVANTAMIENTO (CONVIVIR CON LA OMISIÓN PERO CONOCERLA)

Bajo cualquier esquema de relevamiento de información (administración central o autoadministración de los instrumentos) lo más probable es que se cuente con un número de casos no ubicados y rechazos o con omisiones. Una tasa de respuesta de 1 (o 100%) es pocas veces lograda y, la verdad, esto no es un problema mayor.

Usualmente, es necesario encontrar un balance entre la tasa de respuesta a obtener y el tiempo en el que se requiere la información. Teniendo conocimientos previos sobre la operación de los propios sistemas de información (si se registra la fecha de reporte) es posible estimar cuántas unidades de tiempo adicionales son requeridas en cada momento dado para un incremento determinado en la tasa de respuesta. Usualmente, esta relación tiempo adicional/incremento en las tasas de respuesta tiene una forma como la del gráfico 3.

Gráfico 3. Relación entre las unidades de tiempo requeridas en la generación de información y las tasas de respuesta obtenidas, País A, Censo Escolar 2008



Nota: información hipotética creada para efectos de ilustración. Los conocimientos previos de gestión del sistema de información permiten medir los tiempos reales para cada evento de levantamiento de información. Con los datos correspondientes a diferentes momentos y eventos es posible construir curvas típicas para cada uno de estos.

Si bien el gráfico 3 muestra información hipotética, esta es suficiente para ilustrar el punto que aquí se quiere destacar: en las primeras semanas de la operación de acopio, cada semana adicional representa un incremento importante en las tasas de respuesta hasta que se llega a un punto en el que cada semana adicional deja de traducirse en una ganancia significativa en tasas de respuesta y se convierte, de modo creciente, en un problema adicional respecto de la oportunidad de la información.

Así, la estimación del momento de «corte» o cierre del proceso debe hacerse combinando criterios diversos que definen qué es lo aceptable en términos tanto de los tiempos requeridos (oportunidad de los datos) como de cobertura del evento de generación de información (tasa de respuesta). Usando como referencia el gráfico 3, ¿vale la pena esperar hasta lograr una cobertura cercana al 95% en la semana 17, o sería suficiente con una cobertura del 90% en la semana 11? No hay una respuesta única a una pregunta como esta, ya que el óptimo depende del contexto: no es lo mismo una dilación de seis semanas que una de seis meses, tampoco es lo mismo una dilación en un sistema en formación que en uno que se supone que viene operando de modo estable, etc.

La observación empírica de este fenómeno debe ayudar a tomar la decisión de en qué momento es importante «cerrar» el proceso. El «cierre» permite contar con información que, de ser necesario, puede luego corregirse con nuevos datos, pero que en principio debe ser suficiente para brindar estimados robustos de, por ejemplo, la matrícula total.

El tema principal consiste en diseñar una estrategia para manejar la omisión. Si la omisión se presenta de manera más o menos homogénea a través de toda la población observada, entonces es relativamente más fácil de manejar. El problema potencialmente más importante (y probable) es que la omisión no sea aleatoria y, por lo mismo, que se concentre en determinados segmentos de la población (por ejemplo, instituciones educativas en zonas rurales, instituciones educativas de gestión privada o de una determinada jurisdicción político-administrativa)²³. En estos casos, el «cierre» del proceso de generación de información requiere de esfuerzos adicionales para determinar y, en lo posible, subsanar las omisiones hasta un nivel que sea comparable al del resto del universo observado o, en su defecto, requiere procedimientos diferenciados de estimación de la omisión.

Omisiones con sesgos menores y conocidos son perfectamente manejables por un sistema de información, ya que el conocimiento de la omisión y sus características permite hacer estimaciones²⁴.

1.5.3 PERÍODOS DE REFERENCIA Y CÁLCULO DE EDADES

Un tema clave en la generación de información es el período de referencia de la misma. El período de referencia es un atributo básico de la información, como lo es la población a la que se refiere. Un dato sin fecha (o con una fecha ambigua) es uno que puede llevar a errores significativos que se pueden traducir en decisiones impropias y/o en la pérdida de credibilidad del sistema de información, ya que muestra escasa rigurosidad profesional.

Este tema se manifiesta en varios aspectos de los sistemas educativos, entre los que cabe destacar: (i) el ciclo lectivo (cuando se dice matrícula del año x , ¿a qué período se hace referencia?: ¿al ciclo lectivo iniciado en dicho año calendario?, ¿al terminado en ese año calendario?, ¿al inicio del ciclo?, ¿al final del ciclo?, ¿hay un único ciclo en todo el país?, etc.); (ii) el cálculo de las edades de las personas (se debe considerar edades en años cumplidos a una fecha dada); (iii) la correspondencia entre la información financiera (de un año fiscal dado) con la del ciclo lectivo.

El tema de las edades es uno particularmente delicado que hemos de ilustrar con el caso de las tasas de matrícula en Chile.

²³ Nótese que aquí se hace referencia a instituciones y no a programas educativos. Esto es así debido a que las instituciones refieren a los agentes que reportan, o no, la información y, en este caso, esa es la preocupación relevante y no aquella vinculada al contenido de las experiencias educativas.

²⁴ El tema de las estimaciones es abordado en la tercera sección.

Hasta 2005, las tasas netas de matrícula de la educación primaria²⁵ en Chile resultaban sorprendentemente bajas para un país que, según lo mostraba otra evidencia, tenía prácticamente resuelto el problema del acceso a dicho nivel y mostraba un grado de conclusión de estudios próximo a la universalidad. La situación resultaba particularmente llamativa al compararse los datos de este país con los observados en el resto de la región.

Esta situación motivó un trabajo conjunto entre el Departamento de Estudios del Ministerio de Educación de Chile y la Orealc de la Unesco destinado a identificar, abordar y resolver el problema.

En un sistema educativo con cobertura oportuna perfecta, es decir, donde todas las personas ingresan al sistema oportunamente y cursan cada año de estudios con la edad oficialmente establecida, lo que cabe esperar es tasas netas de ingreso²⁶ al primer grado (es decir, la proporción de personas con la edad de ingreso que efectivamente acceden al primer grado) así como tasas de matrícula oportuna en cada grado (la proporción de personas que teniendo la edad esperada para cursar el grado *g* se encuentra efectivamente en dicha condición) iguales a 1. Si en este caso se observase un sesgo en el registro de las edades (por ejemplo, que los alumnos aparezcan con una edad un año superior a la real), esto llevará a que de modo aparente se cuente con muy pocos estudiantes de la edad inicial en la que deben cursar el nivel y con las edades que corresponden a cada grado. Si el nivel tiene seis grados y la edad de inicio es seis años, esto significará que el numerador de la tasa neta de matrícula del nivel cuenta básicamente con población de siete a once años, es decir, un sexto menos de lo que debería ser. Por esta razón, en esta situación hipotética «perfecta», en vez de observarse una tasa del 100% se tendrá una tasa de 83,3% por el solo efecto del tratamiento sesgado de las edades.

Siendo el caso chileno uno en el que es esperable observar tasas de matrícula próximas al 100%, el hecho de que las tasas (para la educación primaria computada de acuerdo a la CINE con seis grados) hayan sido aproximadamente un sexto inferiores a lo esperable sugirió que el manejo de los períodos de referencia para el cálculo de las edades podría explicar el sesgo encontrado.

En efecto, lo que se pudo observar es que el período de referencia usado para fines estadísticos (edades en años cumplidos al 31 de diciembre) no se correspondía con el estipulado en el cuerpo normativo como fecha para definir la edad normativa o esperada para cada grado (años cumplidos al 30 de abril hasta el año 2005 y al 30 de junio a partir

²⁵ En Chile no existe un nivel conocido como educación primaria. Desde la década de 1960, la educación básica está compuesta por los grados 1 a 8 y la educación media por los grados 9 a 12. Cuando se habla de primaria se está haciendo referencia a la adaptación del sistema chileno a la CINE, según la cual se considera primaria a los grados 1 a 6, como baja secundaria a los grados 7 y 8 y como alta secundaria a los grados 9 a 12.

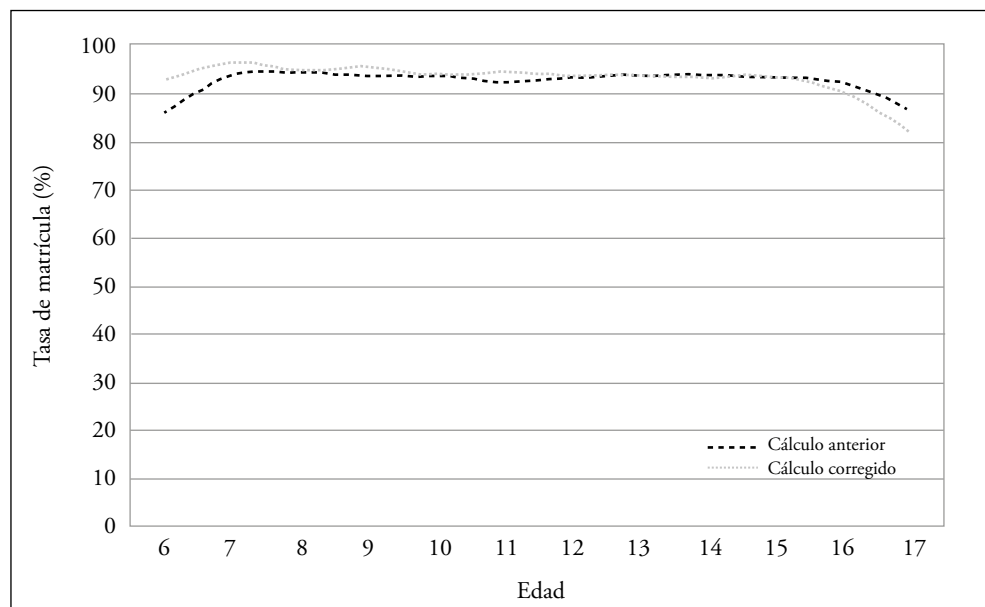
²⁶ Es decir, la proporción de personas que ingresan al primer grado con la edad esperada de acuerdo a la normativa. En el caso de Chile, esta tasa mostraba un absurdo valor de 38% en 1998 (PRIE 2002), lo que, evidentemente, llevaba a protestas o muestras de total desacuerdo por parte de la unidad del ministerio encargada de la educación de la primera infancia.

de 2006), lo que consiguientemente llevaba a un sesgo en el cálculo de las edades. En efecto, si por ejemplo los estudiantes debían tener seis años (al 30 de abril) para cursar el primer grado y, sin embargo, la edad se computaba al final del ciclo escolar (diciembre), era esperable que los estudiantes de acuerdo al reporte estadístico apareciesen con edades superiores a las definidas, ya que, en efecto, al término del año escolar una gran mayoría de los estudiantes eran un año mayores (en años cumplidos) que al inicio.

Desde hace algunos años se ha implantado en Chile un sistema de registro administrativo de los estudiantes (Registro de Estudiantes de Chile, RECH) que cuenta con información sobre la fecha de nacimiento de cada individuo. La base de datos del RECH permite, por tanto, recalculer las edades con cualquier período de referencia que se desee. A estos efectos, el Departamento de Estudios del Ministerio de Educación de Chile obtuvo reportes de matrícula por edades para la educación básica usando tres fechas de referencia: (i) 31 de diciembre, (ii) 30 de abril y (iii) 30 de junio. Dicha información permitió identificar y corregir los sesgos, tal y como se presenta en la publicación de indicadores del año lectivo 2006 (Chile-Ministerio de Educación 2007).

El gráfico 4, construido sobre la base de información de esa publicación, ilustra claramente las diferencias observadas dados los diferentes períodos de referencia usados para el cálculo de las tasas de matrícula por edades simples.

Gráfico 4. Tasas de matrícula por edades simples, 6 a 17 años, Chile, 2006 (cálculo anterior y cálculo corregido ajustando el período de referencia usado para el cálculo de las edades)



Fuente: Chile-Ministerio de Educación (2007: 28).

Como era de esperar, las diferencias se concentran en los extremos (un mayor número de personas de seis años y un menor número de personas de diecisiete en el nuevo cálculo) debido al «desplazamiento» de las edades que resultaba del período de referencia previamente usado.

Gracias al mejor ajuste metodológico resultante de controlar adecuadamente los períodos de referencia para el cálculo de las edades, las tasas netas de matrícula de Educación Básica y Media de Chile en 2006 muestran valores que se corresponden mejor con la realidad educativa del país (94 y 81%, respectivamente) en vez de los que resultarían de usar el procedimiento anterior (88 y 71%, respectivamente)²⁷.

Finalmente, la imprecisión derivada de un inadecuado control de los períodos de referencia no es solo un problema metodológico serio, sino que ello tiene potenciales impactos negativos en la política. Por ejemplo, con ocasión del Seminario Las Transformaciones Demográficas en Chile y sus Consecuencias para las Políticas Públicas, organizado por el Instituto Nacional de Estadística de Chile con la colaboración del Centro Latinoamericano y Caribeño de Demografía de la Cepal (Celade) el 28 de septiembre de 2005, es decir, antes de hacerse la corrección aquí descrita, la presentación del subsecretario de Educación de entonces señalaba que, en el caso de la educación básica, Chile debía atender a «una población que disminuye en número pero que aumenta en matrícula» (Montt 2005: 15), ya que se asumía la necesidad de elevar las tasas de matrícula entonces observadas. Más allá de que el subsecretario haya hecho una presentación que contenía una afirmación tan errada como esta, el problema mayor es que si las estadísticas son tomadas en serio sería posible decidir, por ejemplo, sobre asignación de recursos de maneras que resultarían absurdas. Por otra parte, si las estadísticas no son tomadas en serio, entonces, tenemos un problema mayor de legitimidad del sistema de información.

Pero no solo se trata de controlar los períodos de referencia que competen directamente a los procesos internos. También es preciso considerar qué sucede con los períodos de referencia de información tomada de otras fuentes, como es el caso de aquella sobre población por edades, que usualmente se origina en proyecciones de población al 30 de junio de cada año.

En países donde la fecha de ingreso a un determinado nivel no se fija usando valores enteros en la edad el problema es el mismo: si se dice, por ejemplo, seis años y medio al 31 de diciembre, eso es equivalente a siete años al 30 de junio del año siguiente. Así, si

²⁷ Cuando un procedimiento similar se aplique el cálculo de las llamadas **tasas netas de ingreso** (porcentaje de la población en edad de ingresar al primer grado que efectivamente se encuentra en dicha condición), el cambio será aun más marcado por tratarse de una edad singular. De hecho las cifras no se encuentran disponibles ni a nivel nacional ni a nivel internacional, pero es posible prever que con los períodos de referencia adecuados estas tasas resulten superiores a 80% o incluso a 90% (lo que es una mejor representación de este aspecto de la realidad educativa chilena), en vez del valor por debajo de 40% que se solía publicar. Nótese que (al momento de revisar esta nota, marzo de 2015) en el centro de datos del sitio web del UIS solo existen tasas netas de matrícula (pero no de ingreso) para Chile entre 2007 y 2012.

la información sobre matrícula por edades en años cumplidos se produce con referencia al 30 de junio, no existe ninguna dificultad mayor para los cálculos.

1.5.4 EL PROBLEMA DE LA «OPORTUNIDAD» DE LOS DATOS Y CÓMO ABORDARLO

Un tema central a enfrentar en el diseño y operación de los sistemas de generación de información tiene que ver con la oportunidad de la misma. Como se mencionó a propósito del «cierre» en los ciclos de generación de información, es posible identificar un *trade-off* entre las tasas de respuesta y los tiempos requeridos. Así, tomar una decisión acerca del momento de cierre tiene que ver con las ganancias marginales esperadas en términos de las tasas de respuesta por cada unidad de tiempo adicional y, lo que no se mencionó en ese momento, con el costo que tiene cada unidad de tiempo adicional a invertir.

En efecto, la inversión de tiempo tiene un costo en términos de su impacto en la oportunidad de los datos y este costo no es siempre el mismo, se eleva cuando el plazo se alarga, hasta llegar a un punto en el que puede comprometer la credibilidad del sistema y, por tanto, ser extremadamente oneroso. Un censo escolar anual que no logra cerrarse en, por ejemplo, dos años termina siendo evidencia de serios problemas de diseño y operación y los usuarios se ven enfrentados a una situación que quebranta su confianza; y un sistema de información que no es confiable termina colapsando.

Entonces, ¿cuándo debe cerrarse un ciclo de producción de datos? La respuesta a esta pregunta requiere enfrentar con precisión una pregunta anterior que toca al diseño de los instrumentos y procesos de generación de información: ¿qué información es requerida, por quién y en qué momento?

No toda la información brinda el mismo tipo de evidencia para todos los fines. Por ejemplo, si se está planificando la entrega de materiales educativos (digamos textos escolares) a cada estudiante a inicios del año escolar, es lógico asumir que se necesita información sobre el número de estudiantes al menos en dos momentos: (i) al definir la impresión de los materiales (es decir, cuando se inician los procesos de contratación de los servicios requeridos) y (ii) al distribuirlos.

La primera necesidad debe ser satisfecha con mucha antelación y es imposible que se haga con un número exacto de estudiantes, ya que dicha cantidad (es decir, la realidad que representa: la matrícula al inicio del ciclo lectivo siguiente) aún no existe, por lo que es necesario **proyectar** el volumen y, al menos, una primera aproximación de cómo se distribuye ese volumen por áreas geográficas. Por su parte, la segunda necesidad requiere información mucho más precisa, ya que, con seguridad, involucra contar con cantidades de estudiantes que han de recibir los materiales por cada institución y programa educativo y por cada grado. De igual forma, si se requiere distribuir estos textos, digamos, en la primera semana de las actividades lectivas, lo óptimo sería contar con la información

de matrícula al inicio de dichas actividades; sin embargo, los procesos de acopio y procesamiento de información suelen tomar tiempos mayores a una semana; entonces, ¿cómo proceder? En este caso, resulta claro que un sistema de estadística educativa solo podría proporcionar las mejores proyecciones posibles, pero no números exactos de estudiantes. De hecho, ni siquiera en el hipotético caso de que se contase con el registro administrativo de la matrícula en línea y en tiempo real sería posible contar con la información «exacta» por, al menos, dos razones: (i) la planificación de la entrega se hace con anterioridad, así que no se puede tener la cifra «final» con dicha antelación, pues aún no existe; (ii) los procesos de matrícula no siempre se cierran antes del inicio de las actividades lectivas, porque hay estudiantes que por distintas razones solo se registran una vez iniciado el ciclo lectivo o hay traslados.

El ejemplo anterior ilustra varios problemas a tener en cuenta: (i) la necesidad de definir con precisión lo que se requiere, es decir, enfrentar el hecho de que en diversas ocasiones los usuarios pueden demandar cosas que son lógicamente imposibles (como tener la cifra final de matrícula antes de que el proceso de matrícula haya concluido), lo que deriva en trabajar cooperativamente para aclarar el compromiso necesario entre lo deseable y lo posible; (ii) la necesidad de definir cuál es el mejor dispositivo de generación de información para atender el requerimiento, es decir, ¿pueden ser suficientes las proyecciones?, ¿qué tipo de proyecciones, con qué margen de error, con qué niveles de desagregación?; o, por el contrario, ¿se necesita realmente la información de los registros administrativos de matrícula?

La falta de entendimiento mutuo sobre estos elementos subyace a muchos desencuentros entre los usuarios y los generadores de información. Al parecer, los usuarios «siempre» desean el dato «actualizado» y con el mayor nivel de detalle posible y eso no siempre es factible ni, necesariamente, deseable.

Por ejemplo, decisiones de política nacional sobre la expansión de la cobertura requieren menos de un dato «exacto» y «actualizado» que de una buena serie histórica, una proyección de las tendencias y un buen análisis de ambas.

Supongamos que en el País B en el año 2013 se cuenta con información robusta sobre la cobertura educativa por edades simples para un período de veinte años. El análisis de dicha serie (cómo ha evolucionado; a qué ritmo; con qué diferencias entre, por ejemplo, áreas urbanas y rurales) es mucho más útil para la política que saber si se cuenta con una matrícula para 2015 (si es que este número fuera posible de obtenerse en tiempo real) de 1.257.127 en educación primaria, que la matrícula en 2014 fue de 1.256.145, o que las tasas netas de matrícula de 2013 y 2014 fueron de 95,8% y 95,7%, respectivamente.

De hecho, el caso de las tasas de matrícula es particularmente ilustrativo, pues estas se componen con la información de matrícula y con **proyecciones** de población. El hecho de que un elemento de su cálculo sean proyecciones hace que el resultado sea siempre un estimado, por muy «exacta» que sea la información sobre matrícula.

Un segundo ejemplo puede reforzar aún más este punto. En 2004 nos encontrábamos tratando de reconstruir un sistema de estadística educativa en un país latinoamericano. La situación en la que tal sistema estaba era dramática y se necesitaba un esfuerzo grande y sostenido para revertirla. En ese contexto, en una reunión con usuarios se suscitó una discusión muy compleja acerca del siguiente tema: un donante había ofrecido una cantidad significativa de recursos financieros para un programa de textos escolares, pero requería que el ministerio hiciera, en un período muy breve de tiempo, una propuesta documentada que incluyera números «exactos» de matrícula para tres grados de educación primaria en los siguientes tres años escolares. La discusión se centraba en: «el país no puede esperar a que el sistema de estadística opere como ustedes quieren, necesitamos esos datos ya, si no perderemos el financiamiento».

La buena intención de las personas junto con una extremada miopía primaba en este debate, que al final se resolvió de la única manera que aconsejaba la desesperación: alguien hizo las mejores proyecciones que pudo (con procedimientos nunca claros y con información base de muy dudosa calidad), con lo que el ministerio quedó contento, al igual que el donante, ya que a la mayor parte de los involucrados solo le importaba hacer un uso ritual de los números, y si estos carecían de validez daba más o menos lo mismo. Por supuesto, el resultado adicional fue que quienes estaban desesperados por contar con «buenas estadísticas» nunca se interesaron en invertir en tenerlas y, años después, el sistema de información seguía sin poder recomponerse y, con seguridad, se replicaban discusiones como la aquí reseñada.

La estadística es una disciplina cuyo propósito principal no consiste en contar con información «exacta», sino justamente en saber que esta no lo es, determinando sus márgenes de error para poder manejarlos. Mejor que una relación fetichista con la «exactitud» de los datos es saber cuál es el margen de confiabilidad de lo que hacemos²⁸.

Por lo mismo, un factor de éxito de los sistemas de información está dado por su capacidad para satisfacer a los usuarios partiendo de una mejor calificación de sus demandas, de modo que estas sean mejor atendidas con los mecanismos que mejor correspondan.

Una crítica usual a los datos internacionales de educación es que no son oportunos. Por ejemplo, en marzo de 2009, el centro de datos del sitio web del UIS contaba con información para el año escolar terminado en 2007, lo que suele generar ciertas

²⁸ A inicios de la década de 1990, un buen colega me comentaba que él siempre publicaba sus estimaciones con al menos tres decimales, ya que esto daba una impresión de exactitud que hacía que los usuarios «no molestaran» preguntando cómo se hicieron las estimaciones. Siendo este un razonamiento muy «sensato», ya que permitía que las discusiones se enfocaran en los temas en los que debían enfocarse, es preciso señalar que, en el largo plazo, esto refuerza la relación fetichista con los «datos» y, por lo tanto, no contribuye a un mejor uso de la información.

demandas o críticas²⁹. Cabe anotar, sin embargo, que dichos datos no incluían la información de los países de la (OECD), que publicaron los suyos seis meses después **sin que nadie se quejara por ello**. La falta de queja obedece a que los usuarios entienden que el tipo de información que se publica a nivel internacional no puede ocurrir en un momento más próximo en el tiempo, ya que supone, primero, que cada país termine su propio proceso de generación de información y, segundo, que se haga la transformación a términos internacionalmente comparables, lo que implica procesos propios de control que contribuyen a asegurar la robustez de la información. Por otra parte, los indicadores que se calculan a nivel internacional no refieren a fenómenos que cambien de modo muy dramático de un año a otro y, de hecho, si estos datos sugiriesen cambios fuertes, lo más probable es que hubiera problemas con tales datos más que fenómenos propios de la realidad educativa.

La necesidad de hacer proyecciones es inherente a un sistema de información, ya que permite anticipar eventos y, por lo mismo, contribuir de mejor manera a los procesos de toma de decisiones. Por lo tanto, cómo hacer proyecciones es un tema de crucial importancia que abordaremos en la tercera sección.

²⁹ Por cierto, estos tiempos se han acortado de modo importante en los últimos años y a inicios de 2014 varios países ya contaban con información del año escolar terminado en 2013.

1.6 LA BASE TECNOLÓGICA

La plataforma tecnológica a usarse para la captura, el almacenamiento y la recuperación de la información es un tema que ocupa parte importante de las preocupaciones de los gestores de los sistemas de información. Siendo una preocupación completamente válida, sin embargo, es importante que no ocupe un lugar que no corresponde; es decir, se debe asegurar que la discusión sobre las opciones tecnológicas no vele los temas sustantivos que definen la operación y los propósitos del sistema de información.

Las herramientas tecnológicas son elementos de **apoyo** que han de facilitar el trabajo en diversos aspectos, pero no debe permitirse que el diseño de los sistemas resulte constreñido por la plataforma tecnológica. En el momento en el cual las personas a cargo del soporte tecnológico terminan definiendo qué es lo que se puede y qué es lo que no se puede hacer, el sistema se encuentra trastocado.

La discusión sustantiva es qué es lo que se necesita y la plataforma tecnológica debe facilitar la manera de responder a ello; no al revés. En ese sentido, la discusión sobre motores de bases de datos, o sobre programas estadísticos, es secundaria y más parece obedecer a las preferencias y gustos de algunos consultores o funcionarios que a las necesidades públicas.

Casi cualquier motor de base de datos o programa estadístico **bien utilizado** puede ser lo suficientemente bueno si es que las necesidades de soporte tecnológico del sistema están claramente definidas y se sabe operar las herramientas a usar³⁰.

Así, es importante invertir los esfuerzos necesarios para desarrollar los aspectos centrales del diseño y considerar que la elección de una plataforma tecnológica es una decisión institucional que implica más de un actor. Por ejemplo, pueden existir políticas institucionales destinadas a favorecer el uso de *software* libre o de arquitectura abierta, o convenios comerciales con determinados fabricantes de *software* que son la base para

³⁰ Conocí una situación en la que se usaba un *software* estadístico muy potente (y con una licencia anual muy cara también) y, sin embargo, las personas que lo operaban tenían dificultades conceptuales básicas que implicaban, por ejemplo, no saber definir lo que es un caso válido o un caso perdido (*missing*) o qué es un cuartil. Evidentemente, este no es un problema que se resuelva con el *software* y, en un caso como tal, no sirve de nada contar con uno muy potente. Algo análogo sucede con el Programme for International Student Assessment (PISA) de la OECD, donde se aplican técnicas (estimar errores estándar y computar ponderadores replicados) en casos en los que no corresponde (donde PISA se aplica de modo censal). El informe técnico (OECD 2012b: 140) comete el despropósito de afirmar que, al haber omisiones, se ha usado el dato censal como si fuese una muestra, desconociendo que la omisión censal no es aleatoria y, por lo tanto, si se tratase de una muestra esta no sería probabilística, por lo que no hay forma de estimar dichos errores o computar tales ponderadores.

definir los productos a utilizar. Esto se aplica también a los sistemas de captura de datos, incluidos los esquemas de reconocimiento óptico. En todo caso, es importante considerar el efecto de las características que determinadas opciones imponen sobre el proceso. Por ejemplo, el uso de formatos de reconocimiento óptico mediante marcado de burbujas en un levantamiento de información realizado con entrevistas puede llevar a que la duración total de estas tome un tiempo mayor (marcar cada respuesta mediante un cuidadoso llenado de burbujas supone tiempo) y ello puede afectar la calidad de las respuestas.

En cualquier caso, hay tres elementos básicos en la decisión: (i) asegurar que la opción que se tome sea **sostenible**, es decir, que se trate de un producto o conjunto de productos que tengan respaldo y de los cuales el equipo técnico se pueda apropiar y manejar solventemente; (ii) asegurar que la opción sea **coherente** con otras políticas institucionales existentes; y (iii) asegurar que las personas puedan operar la solución tecnológica elegida, es decir, garantizar que cuenten con la capacitación requerida para maximizar la inversión hecha en *software*, de modo que este sea un **facilitador** del trabajo a desarrollarse.

Ya que hay distintos productos que pueden ser muy útiles para la tarea, la elección no debe ser una que corresponda al gusto o preferencias de alguien, sino una que se ajuste a las características institucionales en las que se ha de operar. Asimismo, debe mencionarse que, en ningún caso, una opción válida consiste en elegir un paquete tecnológico que lleve a la permanente dependencia de quien lo provee, ajusta, modifica y/o actualiza. Si bien hay personas que buscan acuerdos que de esta forma garantizan una longeva relación de dependencia de la consultoría o, la **inmortalidad del consultor o contratista**, estos acuerdos son completamente perjudiciales para el interés público y deben ser evitados.

1.7 ¿REGISTROS ADMINISTRATIVOS O INFORMACIÓN ESTADÍSTICA?

En cierta medida, la discusión anterior ya anticipó parte del contenido del presente acápite. En efecto, en este momento debe resultar claro que los sistemas de información se pueden alimentar tanto de información generada mediante eventos estadísticos, como mediante información capturada de sistemas de registro administrativo creados para fines diferentes a la generación de información.

Por lo mismo, el primer elemento que debería quedar claro es que la diferencia entre la estadística y el registro administrativo no está dada por la institución a cargo de generar la información, sino por el procedimiento y la base institucional que lo sustenta. Desde ese punto de vista, la opinión escuchada demasiadas veces acerca de que las oficinas nacionales de estadística producen datos estadísticos a través de censos y encuestas de hogares mientras que los ministerios de línea (incluyendo a los de Educación) generan información de registros administrativos es, sin lugar a dudas, un sinsentido. Cualquier institución puede hacer ambas cosas, ya que la afiliación institucional no define el carácter de la información.

De hecho, los ministerios de Educación suelen generar información de ambas formas. Por ejemplo, los operativos de medición de logros académicos de los estudiantes son, mayormente, eventos estadísticos desarrollados con propósitos y métodos estadísticos. La excepción serían los casos donde dichas pruebas cumplen una función en los procesos de certificación individual de competencias, pero estos son realmente excepcionales, por lo menos en América Latina. Del mismo modo, los censos escolares son eventos estadísticos, incluso cuando pretenden (pero no aseguran) que parte de su información sea tomada de los registros administrativos, ya que, a fin de cuentas, la información es reportada por un informante que puede o no haber hecho uso de dichos registros.

Sin embargo, ¿por qué es importante distinguir entre la información generada mediante eventos estadísticos y la que resulta de registros administrativos? Desde el punto de vista del uso de la información, esta diferencia puede resultar trivial, sin embargo, no lo es desde el punto de vista del diseño y la operación de los procesos de generación y diseminación de la información.

Los registros administrativos y la información estadística obedecen a propósitos diferentes y, por lo mismo, es necesario tomar en cuenta estos propósitos a la hora de hacer uso de la información. Asimismo, esos propósitos definen estructuras institucionales propias para la generación y diseminación de la información que deben ser tenidas en cuenta, así como pautas legales para el manejo de la información.

Por ejemplo, en un país en el que se observa dificultades para la generación de información relativa a los programas educativos gestionados por instituciones privadas, podría ser posible recurrir a **fuentes alternas de información**.

Una de estas podría ser la administración tributaria. Considerando que la prestación de servicios educativos suele ser una actividad exenta de algunas obligaciones tributarias sustanciales (pago de tributos), es posible que las instituciones que hacen uso de estas exenciones sí deban satisfacer algunas obligaciones formales (como reportar información) justamente para mantener la prerrogativa tributaria que se les concede.

Ahora bien, ¿es dicha información accesible, dadas las normas legales vigentes sobre reserva tributaria? De no serlo, ¿qué pasaría si la administración tributaria llegase a compartir dicha información? Claramente se tendría no solo una situación de violación de una norma legal y del derecho de las personas a la confidencialidad por dicha norma protegido, sino también se mermaría la credibilidad de la operación de la administración tributaria, lo que acarrea un perjuicio importante para la colectividad. Así, la confidencialidad en este caso es un mecanismo básico para proteger no solo a los individuos frente a un potencial uso arbitrario de la información sino también para asegurar que la colectividad cuente con una administración tributaria efectiva. La reserva tributaria existe para servir ciertos propósitos que son de interés público, por eso su violación acarrea problemas que probablemente son más graves que los beneficios supuestos que dicha violación traería.

En el caso de que la normativa legal sí permita a la administración tributaria compartir dicha información, se requiere conocer cuáles son los términos en los que esto es posible: ¿mediante registros individuales para información que se entiende no compromete los principios que la reserva protege?, ¿mediante reportes agregados que la administración tributaria podría o debería producir a solicitud de la autoridad educativa?

Por su parte, los **datos estadísticos** suelen verse protegidos también por normas de reserva o secreto estadístico. Estas normas tienen como propósito usual salvaguardar la credibilidad del sistema (lo que es de interés público) y ofrecer a los informantes garantías que lleven a un reporte veraz de información (lo que también es de interés público). Imaginemos un caso hipotético en el que una oficina nacional de estadística (ONE) genera información sobre ingresos y gastos de las familias y la comparte, individualizando a las personas del segmento de mayores ingresos, con la administración tributaria, la cual, entonces, compara estos datos con sus propios registros y observa, en algunos casos, inconsistencias que desatan procesos de auditoría de esos contribuyentes. Es claro que si este fuera el caso los informantes se verán impulsados a no declarar verazmente (al menos la información estadística) en el futuro y, por lo mismo, la capacidad de operación del sistema estadístico se vería mermada.

De igual forma, compartir **información protegida** con otros agentes públicos o privados no solo se encuentra prohibido por las normas, sino que dicha prohibición obedece a la necesidad de proteger el interés público. Por ejemplo, empresas privadas pueden ver

impropiamente beneficiada su participación en el mercado (en detrimento no solo de sus competidores, sino de los propios consumidores) mediante el acceso a información detallada sobre potenciales clientes (por ejemplo, instituciones educativas de gestión privada en las que podrían vender sus productos o servicios); o un partido político puede obtener alguna ventaja frente a sus adversarios al recibir información protegida sobre la población, afectando el derecho de las personas a un proceso electoral justo.

Las normas sobre reserva de la información obedecen (o, por lo menos, deberían obedecer) a un propósito consistente en salvaguardar el interés público, incluyendo la protección de la privacidad de los individuos. No es otro el argumento. La mejor forma de tratar con las normas de reserva de la información es, primero, ser consciente de ellas y de que estas son de naturaleza y alcance diferente, dependiendo de la propia naturaleza de la información (que afectan la privacidad, etc.); y, segundo, operar frente a la normatividad legal de la forma en que debe hacerse: **nadie está impedido de hacer lo que la ley no prohíbe y nadie está obligado a hacer lo que esta no manda.**

Este último principio tiene que ver también con la tendencia tecnocrática a poner **mayores restricciones al acceso** a la información que las establecidas legalmente y que aquellas que la protección del interés público sugiere. Por ejemplo, no es deseable restringir el acceso a la información por temor a su «mal uso».

La información generada por las instituciones públicas en su accionar es, por definición, de carácter público, a menos que una norma específica indique lo contrario (por cierto, a efectos de salvaguardar el interés público y no por otra causa). Restringir el acceso a la información para evitar un «mal uso» de la misma es equivalente a decir que las personas no deberían votar ya que pueden «elegir mal» o que el parlamento no debería, por ejemplo, legislar en materia tributaria, sino que lo habría de hacer un grupo de expertos (lo que violenta principios democráticos básicos).

Los derechos fundamentales de las personas no pueden coactarse con un argumento tecnocrático, las personas tienen derecho incluso a equivocarse y justamente los mecanismos democráticos son los que permiten identificar los yerros, enmendar y mejorar. Un régimen tecnocrático carecería de los mecanismos de control para mejorar y terminaría ensimismado en sus propios errores. A fin de cuentas, quién dice que los «expertos» no pueden cometer errores garrafales o establecer tiranías (Easterly 2013). De hecho, una parte de los problemas o de la falta de progreso educativo obedece justamente a la aplicación de juicios «expertos» carentes de sentido³¹.

³¹ Un ejemplo de esto es la «recomendación», que se viene repitiendo desde hace aproximadamente cinco décadas, que sugiere que los países deben invertir «al menos» el 6% de su producto bruto en educación. Más adelante se explica en detalle el absurdo detrás de esta afirmación repetida por innumerables «expertos», baste aquí con señalar que nadie que afirme esto puede argumentarlo. El argumento basado en afirmar que «la Unesco lo recomienda» —ya que una afirmación como esta se encuentra en Unesco (1996: 187) y en declaraciones de conferencias internacionales organizadas por la Unesco— es un típico caso de argumento *ad verecundiam*, o de recurso a la autoridad, lo que es una forma de falacia.

1.8 LA ESTADÍSTICA EDUCATIVA COMO PARTE DE UN SISTEMA ESTADÍSTICO NACIONAL

Los sistemas de información del sector público componen, o deberían componer, un cuerpo coherente que brinde un servicio a la ciudadanía en general y a la gestión pública en particular. Sin embargo, esto no siempre sucede, aunque existen diversos esfuerzos de integración y convergencia realizados con diferentes grados de éxito.

Uno de estos esfuerzos está vinculado a la existencia de un SEN, que usualmente está formalmente establecido mediante un cuerpo legislativo y cuenta con un instituto u ONE como institución que cumple una labor que bien puede ser normativa, de liderazgo técnico o, al menos, de coordinación entre ella misma y las unidades sectoriales de estadística (las oficinas de estadística de los ministerios de línea u otras agencias públicas).

Los niveles de coordinación y coherencia del trabajo entre las instituciones son variados como variadas son las dinámicas institucionales de los países, lo que hace que los SEN sean más o menos sólidos, o incluso realmente inexistentes.

Dichas coordinaciones usualmente se han estructurado alrededor de cuatro grandes campos: (i) estadísticas económicas; (ii) estadísticas demográficas; (iii) estadísticas de energía y medio ambiente; y (iv) estadísticas sociales.

En el primer caso, el de las estadísticas económicas, las coordinaciones entre la ONE, el Banco Central, el Ministerio de Economía o Hacienda y otras instituciones juega un papel central; y el trabajo se estructura alrededor de una agenda sobre la cual existe abundante documentación y estándares internacionales establecidos. Los sistemas de cuentas nacionales, precios al consumidor, finanzas públicas, balanza de pagos, etc. no solo tienen una documentación técnica abundante, sino que su operación está en el centro del interés de organismos nacionales e internacionales (especialmente de carácter financiero) que, por lo mismo, mantienen un seguimiento y se interesan en su fortalecimiento.

En el caso de las estadísticas demográficas y los censos poblacionales y de vivienda, así como en el de los estudios demográficos (incluyendo proyecciones, dinámica de la transición demográfica, análisis de fenómenos migratorios, etc.), ocupan a los sistemas de información, incluyendo a la ONE, y, donde existen, a consejos nacionales de población o agencias equivalentes.

Algo similar ocurre en el caso de las estadísticas vinculadas a la energía y el medio ambiente.

Finalmente, las estadísticas sociales son un campo muy amplio que aborda temas que van desde la medición de la pobreza hasta la de las variables educativas, pasando por el empleo, la salud y otros.

Ahora bien, cuando se observa los sistemas de información de América Latina y el Caribe es claro que el mundo de las estadísticas económicas suele aparecer como el más robusto y estable, con series de datos mejor constituidas, con normas y procedimientos normalizados. Una de las razones para esto es el interés de diversas agencias con capacidad tanto para movilizar recursos y decisiones, así como para producir mayor documentación tendiente a su estandarización (como el DQAF³² y otros instrumentos existentes).

Las estadísticas sociales, por su parte, en muchas ocasiones no operan como un cuerpo coherente, ya que las labores en información de la ONE, por un lado, y las conducidas por los ministerios de línea, por otro, no siempre se encuentran propiamente coordinadas.

Un ejemplo claro de malas prácticas en este terreno está dado por la usual «competencia» entre el uso de las «tasas de matrícula» producidas por los ministerios (con información de matrícula de los censos escolares) y las «tasas de asistencia» generadas por las ONE mediante censos o encuestas de hogares. Exploremos este tema con un poco de detalle.

Las tasas de matrícula (que veremos con mayor precisión en la siguiente sección) usualmente buscan responder a preguntas vinculadas al **acceso** a los sistemas educativos (¿cuántas personas acceden?) o, dicho de otra forma, a la **cobertura** de estos servicios (¿cuántas personas son atendidas por el sistema educativo?) y expresan esto como una proporción de una población dada³³. Por su parte, las tasas de asistencia buscan dar cuenta de la población que **asiste** a un programa educativo.

Aquí hay dos conceptos diferentes: el de acceso o cobertura y el de asistencia. Mientras que el primero constata el hecho de que las personas cuentan con el servicio educativo, el segundo **podría** sugerir una medida del «uso» de este. El primero verifica el hecho clave del acceso, mientras que el segundo da cuenta de algo «más».

Así, una primera diferencia es conceptual, en el sentido de que no son tasas que midan lo mismo. Sin embargo, hay otros problemas a considerar y, el primero, es que esta diferencia conceptual parece que no es bien entendida, dado que con frecuencia ambos conceptos se usan como medidas alternativas. Por definición, la «asistencia» debe

³² El Fondo Monetario Internacional (FMI) ha preparado versiones específicas del DQAF para: (i) cuentas nacionales; (ii) índices de precios al consumidor; (iii) índices de precios al productor; (iv) estadísticas de las finanzas públicas; (v) estadísticas monetarias; (vi) estadísticas de balanzas de pagos; y (vii) estadísticas de deuda externa. Todas estas versiones se encuentran disponibles en su Data Quality Reference Site (<<http://dsbb.imf.org/Pages/DQRS/DQAF.aspx>>).

³³ En realidad, como se discutirá en la siguiente sección, esto es lo que hacen las llamadas **tasas netas de matrícula** y las llamadas **tasas de escolarización por edad específica**, pero no es el caso de las llamadas **tasas brutas de matrícula**, que miden el volumen de esta y que son ratios y no tasas. Aclarar estos términos nos ha llevado a la necesidad de sugerir abandonar algunos nombres en favor de otros que expresen mejor lo que se mide y también a abandonar algunas medidas en favor de otras mejor definidas. Estos temas se abordan en detalle en la segunda sección.

ser menor o igual al acceso. Sin embargo, aquí surge un segundo problema, ¿qué quiere decir «asistir»?

Tanto el concepto de acceso como el de asistencia suponen una referencia en el tiempo. En el primer caso, es acceso en un ciclo lectivo dado (por eso las tasas de matrícula son anuales); mientras que en el segundo esto no es tan claro y depende más bien de cómo se operacionaliza el concepto en el momento de capturar la información, es decir, al determinar cuál es el período de referencia de la información y cómo este se vincula al propósito.

En una situación normal, las personas que acceden a un programa educativo asisten a él con cierta frecuencia. Por ejemplo, sería innegablemente importante conocer, en un ciclo escolar compuesto de 1.000 horas, 200 días o 40 semanas, cuál es la proporción de dicho tiempo en la que las personas **asisten**. Sin embargo, no es esto lo que suelen medir las tasas de asistencia y aquí se presenta un problema muy importante: **la forma** como se formula la pregunta sobre la asistencia y **el momento** en el que se hace **definen lo que se está midiendo** y esto puede ser, y de hecho es, extremadamente variable.

Aquí tenemos varios escenarios posibles. Por ejemplo la formulación de la pregunta al respecto puede ser una de las siguientes:

- **¿Se encuentra usted asistiendo de modo regular a un programa educativo durante el presente período lectivo?** Esta pregunta mediría algo muy parecido a la matrícula, sin embargo, el elemento «modo regular» introduce una ambigüedad difícilmente manejable. ¿Qué entiende el entrevistado por esto?
- **¿Ha asistido usted a algún programa educativo durante la última semana/mes/año?** Dependiendo del período de referencia y el momento de administración de la pregunta (usualmente durante una encuesta de hogares que se administra durante un período prolongado de días o, incluso, de modo continuo), ella puede derivar en muchas direcciones no manejables.

A continuación se presenta algunos escenarios posibles de respuestas a estas preguntas indicando entre paréntesis elementos que pueden ser parte de la respuesta pero no mencionarse porque la pregunta no suele pedir una explicación:

- (i) No asistí la semana pasada /el mes pasado (pues estábamos de vacaciones).
- (ii) No asistí la semana pasada (pues estaba enfermo, pero la próxima semana me reincorporo).
- (iii) Sí, este año he ido (dos días al inicio y luego decidí no continuar y no me matriculé).

Como queda claro, aquí hay muchos cabos sueltos que conspiran contra el valor de la información.

Esto, por otra parte, sugiere un problema serio de diseño en las preguntas sobre asistencia en las encuestas de hogares, ya que, al parecer, el propósito de la pregunta no siempre es preciso y, en ocasiones, existe una cierta cerrazón a revisar lo que se tiene con el manido argumento (que en realidad es una coartada para no mejorar) de «mantener las series».

Por otro lado, la parte del año en la que se administra la pregunta (especialmente cuando en las encuestas el momento es variable) complica las cosas no solo con relación a los períodos vacacionales, sino y fundamentalmente con relación al cálculo de las edades. Este es un tema fundamental en educación, especialmente cuando se trata de conocer el acceso o la condición de matrícula de las personas. Como se anotó anteriormente con la ilustración sobre el caso de Chile, esto puede producir importantes sesgos a la hora de calcular tasas de matrícula o de asistencia, si estas últimas son usadas con el mismo propósito que las primeras. Esto, por otra parte, puede resultar manejable si es que la encuesta indaga sobre las fechas de nacimiento, pero si solo requiere información sobre edades en años cumplidos o año de nacimiento, entonces es una potencial fuente de «suciedad» en los datos difícilmente manejable.

Asimismo, las tasas de matrícula presentan sus propios problemas como, por ejemplo, arrojar valores teóricamente imposibles como aquellos que son superiores al 100%³⁴. Estos valores se explican por tres factores que pueden actuar de modo independiente o conjunto:

- (i) la matrícula está sobrerreportada;
- (ii) las proyecciones de población usadas para el cálculo de las tasas están subestimadas; y
- (iii) los períodos de referencia de la información de matrícula por edades y de la proyección de la población por edades no se corresponden y existe una variación estacional en los nacimientos que deriva en diferencias suficientes para producir estos resultados.

Dados estos problemas, sumados a las ventajas derivadas de usar información de encuestas de hogares (la posibilidad de relacionar la información educativa de los individuos con otros atributos del mismo o de su hogar, como niveles de ingreso, pertenencia étnica, etc.; relaciones que no son posibles con los datos de matrícula de los ministerios),

³⁴ En el caso de las **tasas netas** y las **tasas específicas por edades simples**, esto es teóricamente absurdo; en el caso de las llamadas **tasas brutas**, esto es perfectamente posible, pero ahí el absurdo es otro, que será discutido más adelante.

las tasas de asistencia han aparecido como una buena alternativa a las tasas de matrícula con claras deficiencias.

Si bien esto parece una reacción sensata, en el mediano y largo plazo produce tres complicaciones:

- (i) no se busca resolver los problemas de origen con los datos de matrícula;
- (ii) se agudiza la falta de credibilidad del sistema de información (o al menos de uno de sus componentes); y
- (iii) se tiende a ignorar las ya mencionadas limitaciones de las tasas de asistencia, con lo cual se siguen perpetrando errores que a su vez derivan en la debilitación del sistema de información.

Como ya se señaló, las ONE a través de los censos de población y, sobretudo, de las encuestas de hogares tienen la posibilidad de generar información muy rica sobre temas educativos (como las ya mencionadas asociaciones con atributos del individuo y del hogar, el gasto de los hogares en educación o información educativa de personas actualmente no matriculadas en ningún programa educativo), información imposible de ser generada desde los ministerios de Educación.

Por su parte, la riqueza de la información generada por los ministerios tampoco puede ser desdeñada, y también debe construirse sobre el propio conocimiento técnico que viene del campo de la educación. Por ejemplo, este último lleva a cuestionar o redefinir de modo sustantivo el valor de información como la expresada mediante las llamadas **tasas de alfabetismo**³⁵ o a plantear preguntas claves sobre el período de referencia para el cálculo de las edades usualmente no previstas por las ONE.

Así, existen claras razones prácticas que se añaden a otros principios de la acción pública³⁶ que inducen a pensar que es conveniente, necesario y obligatorio operar como un sistema nacional de información que sume lo que pueden brindar diferentes agentes; es decir, coordinar tareas y cooperar en la definición de los elementos involucrados en la producción de información y en el análisis de la misma.

Usualmente, las ONE convocan a consejos de coordinación con las oficinas sectoriales, del mismo modo que la División de Estadística de las Naciones Unidas (UNSD) lo hace con los demás organismos estadísticos a escala internacional en el Comité de Coordinación de las Actividades Estadísticas de las Naciones Unidas. El desafío es convertir estos espacios en oportunidades de cooperación y desarrollo y no solo en rituales burocráticos. La inoperancia, al igual que la «competencia» absurda y la no coordinación crean confusión y minan todo el sistema en un área muy sensible: la credi-

³⁵ Véase Unesco-UIS (2009b) y, más adelante, el acápite 2.4.8 «Medición de aprendizajes con consideraciones sobre el caso del alfabetismo adulto».

³⁶ Sobre esto, véase UNSD (2005, 2013).

bilidad. La credibilidad es un atributo fundamental de la información: si esta no es verosímil, no es nada.

La coordinación y el trabajo conjunto pasa por desarrollar esfuerzos para contar tanto con marcos conceptuales comunes como con definiciones comunes y por estar abiertos a conversar y modificar lo propio. Durante los primeros años de este siglo, la Comisión Económica para América Latina y el Caribe de las Naciones Unidas (Cepal) y la Orealc-Unesco Santiago iniciaron un proceso de este tipo que ha dado frutos en el trabajo de ambas organizaciones³⁷.

³⁷ Esfuerzos como el del Consorcio Paris21 orientados a que los países construyan estrategias nacionales de desarrollo estadístico suponen también este tipo de convergencia en el plano nacional.

1.9 LA DISEMINACIÓN DE LA INFORMACIÓN

Este tema suele ser objeto de tratamiento en la discusión sobre los sistemas de información; sin embargo, usualmente se trata **después** de diseñar los sistemas de producción de información, bajo la forma típica: «Bueno, ya tenemos los datos, ahora, ¿qué hacemos con ellos?».

Este enfoque tiene un problema de partida: las opciones de diseminación deben ser discutidas **desde el inicio**, pues no se trata de una actividad «adicional», sino de una que toca directamente a los propósitos de la información. Esta se produce para informar a alguien y el proceso de hacerlo tiene una forma que es directamente dependiente de las características de ese alguien y de los usos potenciales que ha de darle a la información.

Si la información no se disemina, no informa a nadie y, por lo tanto, no existe como información. El discurso recurrente que afirma que la «información es poder» y que, por esa razón, algunos buscan reservársela para sí mismos y no la «sueltan» es un sinsentido que solo deriva en la autodestrucción de dicha práctica. Si la información no se disemina, no es nada y, por lo tanto, el «poder» de quien detenta nada es insignificante y chocante.

En realidad, la circunstancia que suele estar más asociada con retener o no publicar información es el temor a lo que esta puede mostrar tanto en términos sustantivos como técnicos. Es decir, se oculta información porque no es «conveniente» o porque no se quiere exhibir las debilidades que esta tiene. En ambos casos, el resultado es el mismo, el sistema de información ve afectada su credibilidad y se violentan los derechos básicos de las personas a acceder a datos que les pertenecen.

Un ejemplo puede ilustrar mejor este tema. En 1998, la Orealc-Unesco Santiago culminaba la preparación del primer informe de resultados de su primer estudio regional sobre logros académicos de estudiantes de tercer y cuarto grados de educación primaria (Unesco Santiago 1998) y circulaba los borradores del mismo entre los países participantes. Luego de leer el borrador del informe, el gobierno del Perú notificó a la Unesco su voluntad de que los resultados de dicho país no fueran incluidos en la publicación³⁸.

³⁸ La comunicación entre el Ministerio de Educación del Perú y la Unesco Santiago que derivó en la solicitud de no publicar los resultados del Perú se basó en la identificación de algunos problemas técnicos que sugerían posponer la publicación hasta su subsanación; dado que la Unesco Santiago no estaba de acuerdo con la necesidad de posponer para subsanar (debido a la presión por publicar que se originaba en las organizaciones que financiaban el estudio), el Perú pidió no ser incluido por no estar de acuerdo con los procedimientos seguidos. Hoy es sabido que este estudio tuvo muchas debilidades técnicas que han sido progresivamente superadas en ediciones posteriores; sin embargo, dichas debilidades tornaron imposible, por ejemplo, que el segundo estudio regional pudiese identificar algún tipo de tendencias o cambios respecto del primero, ya que al mejorar el diseño y los procedimientos seguidos no hubo forma de hacerlo comparable al primero.

En febrero de 2001, meses después del estrepitoso colapso de un gobierno encabezado por quien fuera el presidente del país desde 1990 (hoy condenado por la justicia bajo cargos de corrupción y violación de los derechos humanos), el entonces establecido gobierno de transición impulsó nuevos vientos de transparencia y, en ese contexto, el Ministerio de Educación del Perú publicó dichos resultados (Perú-Ministerio de Educación 2001) y señaló por qué la publicación inicial de la Unesco no incluyó los datos del Perú:

[...] una de las razones... [cabe preguntarse cuáles son las otras] fue que algunas de las autoridades políticas, en su momento, consideraron que los resultados podrían ser utilizados para argumentar que la gestión educativa del último gobierno había sido deficiente. El temor provenía de la posición rezagada de estudiantes peruanos respecto de sus similares de muchos países de la región que participaron en el estudio [...]. Estas autoridades no entendieron, o temieron que el público no entendiera, que los resultados de los estudiantes no dependen exclusivamente de la gestión de un gobierno en particular (Perú-Ministerio de Educación 2001: 1).

Este texto permitiría hacer muchos comentarios, empezando por su timidez para denunciar lo que se presenta como una violación del derecho de las personas a acceder a información que les pertenece y siguiendo con la racionalización de los posibles temores acerca de que alguien no «entienda» (parece que solo los expertos entienden) dicha información. Sin embargo, el propósito de traer este ejemplo se centra en lo siguiente:

- Si se trató de una decisión política subalterna (lo que no parece ser el caso), el hecho de no publicar resultados que son de propiedad pública no es una decisión que pueda legítimamente tomar un gobierno y no existe ningún argumento político posible para justificarlo. Esta decisión consistiría en una violación de principios básicos que deben regir los sistemas de información, como la independencia y la transparencia.
- El supuesto ocultamiento de información produjo un resultado de una intensidad que probablemente haya sido mayor que la que hubiese tenido su publicación, aunque, evidentemente, esto es algo que no se puede comprobar. El argumento supuesto que sustentó esta decisión se convirtió en «evidencia» del carácter antidemocrático del gobierno y de su deseo de ocultar información seguramente muy negativa (el sentido común sugiere que nadie va a esconder «buenas noticias»). Esto derivó muy rápidamente en un debate muy cargado en el terreno emocional y

Así, la decisión peruana de 1998 no pareciera haber sido motivada por un afán de manipulación política de los resultados, lo que no hubiese sido coherente con el hecho de que **simultáneamente** el ministerio de Educación del Perú solicitó participar en el estudio de la OECD (PISA).

en la aparición de un lugar común reforzado por las emociones del momento: «Los datos se ocultan porque el Perú quedó último en la región» (cuando el estudio cubrió solo doce de más de cuarenta países de ella), afirmación que se fue transformando en: «La educación peruana es la peor de la región», de tal suerte que nueve años después un ministro de Educación llegó a afirmar en una entrevista, refiriéndose a la participación del Perú en un estudio completamente distinto (Programme for International Student Assessment [PISA], de la OECD), lo siguiente:

Ministro: En noviembre del año pasado [2006] ya se tomó la nueva evaluación PISA.³⁹

Periodista: ¿Y estamos peor?

Ministro: Probablemente. Pedí un adelanto de resultados y me dijeron que estos no eran muy alentadores⁴⁰. En la anterior evaluación quedamos penúltimos, solo por encima de Haití, entre un conjunto de 40 países.⁴¹

Es decir, nueve años después, una discusión basada en el supuesto ocultamiento de información lleva al total sin sentido: Haití nunca ha participado de PISA, pero si el Perú es «el peor de la región» el prejuicio sugiere que debe estar más o menos como Haití, y la evidencia debe ser PISA, ¿o hay otra?

De nuevo tenemos un caso en el que el mal manejo del sistema de información (la interpretación sesgada de una decisión que, pasada la euforia de esos años, hoy podría parecer razonable) contribuye a que una autoridad nacional haga afirmaciones respecto de las cuales la vergüenza es un sentimiento recomendable. Pero, más importante que esto, el mal manejo lleva a desinformar, lo que es aun peor que no generar información y puede terminar conduciendo a decisiones dañinas con respecto al ejercicio del derecho a la educación por parte de las personas. La imagen de la educación peruana como «la peor» es una que impacta en la probabilidad de que, por ejemplo, se pueda reclutar a los más talentosos egresados de la educación secundaria para que estudien para ser docentes y, además, hace que se pierda de vista que los niveles de aprendizaje de los estudiantes han mejorado desde, por lo menos, el año 2001.

Entonces, volvamos a la pregunta de partida: ¿cómo diseminar la información? La respuesta a esta pregunta es, como siempre: depende. Depende íntegramente de las características de los usuarios y de sus necesidades, de sus usos potenciales y de las propias propiedades de la información. No es posible pensar en una solución única para todos

³⁹ Nótese que el Perú no participó de PISA en 2006, así que hay un equívoco desde la propia afirmación inicial del ministro, quien probablemente quería referirse al segundo estudio latinoamericano aplicado entonces.

⁴⁰ Los datos eran procesados por la Unesco Santiago y para esa fecha no había «adelantos», así que no hay forma de determinar a qué se estaría refiriendo el ministro.

⁴¹ Entrevista en el diario *El Comercio* (15 de julio de 2007).

y como los recursos no son ilimitados hay que jerarquizar las necesidades y enfrentarlas de acuerdo a las posibilidades.

Entonces, ¿quiénes son los usuarios y que necesitan?

- Empecemos por la parte fácil. Hay usuarios que necesitan la información como **material de consulta o referencia** y/o como insumo para sus propias labores de **estudio o investigación**: universidades, centros de investigación, expertos, etc.

Sin embargo, estos desarrollan sus actividades para sus propios fines privados y no es función del Estado subsidiar dichas actividades, al menos desde los sistemas de información (se puede ofrecer subsidios a través de otras vías, como, por ejemplo, becas de investigación, ya que la producción de conocimiento también es un tema de interés público). Entonces, ¿cómo brindar el máximo servicio a estos usuarios minimizando los costos de dicho servicio?

Tradicionalmente, este ha sido el público al que explícita o implícitamente se han abocado los esfuerzos de diseminación. Pensando bien las cosas, los tradicionales anuarios estadísticos, como los directorios telefónicos, siempre han sido materiales de referencia o consulta y no soportes de procesos de toma de decisión.

El desarrollo de las tecnologías de la comunicación e información, sin embargo, torna obsoleta la producción de materiales de referencia en formato impreso.

Los sistemas de consulta en línea (que permiten consultas personalizadas) y la diseminación de bases de datos son justamente lo que mejor pueden (o deberían poder) aprovechar quienes desarrollan labores de estudio e investigación. Los costos de instalación de estos sistemas pueden ser significativos, pero su mantenimiento no lo es, y constituye una tremenda oportunidad en términos de garantizar el acceso público a información que pertenece al dominio público.

Como principio, mientras mayor el volumen de información disponible en línea, mejor. Sin embargo, este principio debe conciliarse con la posible existencia de información protegida por normas de reserva y, por lo mismo, el acceso a ciertos datos (cuando esto es posible según la normativa vigente) debe controlarse mediante compromisos de confidencialidad que sirvan como acuerdos contractuales entre el usuario que accede a la información protegida y la institución a cargo de su custodia. Asimismo, es de capital importancia para el desarrollo riguroso de las labores de explotación de la información que esta se encuentre adecuadamente documentada. Ningún investigador que se precie de serio puede usar datos sin conocer al menos los elementos centrales de los metadatos correspondientes⁴². La investigación

⁴² A nivel internacional un conjunto importante de agencias (Bank for International Settlements, Eurostat, OECD, Banco Central Europeo, FMI, Naciones Unidas y Banco Mundial) vienen trabajando de modo conjunto en el desarrollo de estándares para el intercambio de datos y metadatos. Al respecto, véase el sitio web de esta iniciativa, llamada Statistical Data and Meta-Data Exchange (<<http://sdmx.org>>).

sería requiere que el investigador **se ensucie las manos**, lo demás es una práctica de la que se hace bien en sospechar.

Todo esto, por su parte, tiene un impacto positivo adicional sobre los sistemas de información, ya que la construcción de sistemas de consulta en línea fuerza a que las bases de datos se normalicen o estandaricen así como impone ciertos requisitos mínimos de consistencia de la información.

Estos usuarios son muy importantes desde el punto de vista que considera positivo sacar el máximo provecho a la información existente (la que es de interés público) y también para descubrir potenciales problemas con ella. De hecho, esta última posibilidad es la que lleva a algunos a restringir el acceso a la información, ya que existe el temor de que una vez que los usuarios encuentren problemas se afecte la credibilidad del sistema. Ciertamente esto puede ser visto de tal manera; sin embargo, como en todos los casos, el sentido de una acción depende del contexto: en una situación de celos mutuos esas son reacciones esperables, pero en el marco de construir relaciones de confianza justamente la apertura y la transparencia son, en adición a su valor intrínseco como principios democráticos, tremendamente funcionales. Promover una relación de confianza hace que las críticas originadas en la identificación de posibles falencias de la información tengan siempre un carácter constructivo, ya que todos los actores involucrados entenderán que existe un interés común en mejorar la información. Por supuesto, siempre existe la posibilidad de que algunos puedan actuar de mala fe, pero ese es un riesgo que no se controla limitando el derecho de **todos** a la información, por lo que no es un argumento ni legítimo ni operacionalmente válido. La identificación de problemas es el primer paso para mejorar y, en ese terreno, los usuarios especializados pueden hacer contribuciones sustantivas.

- Un segundo tipo de usuarios está dado por las **altas autoridades nacionales** en educación (ministros, viceministros, etc.). En relación a la información sobre el sector, es importante considerar cuál es el nivel de responsabilidad de estos, así como su naturaleza.

Las máximas autoridades tienen una responsabilidad política fundamental (y no técnica), por lo que no tienen por qué ser necesariamente expertos. Así que es posible que un nuevo ministro suela requerir un primer conjunto de información básica que dé cuenta de los atributos centrales del sistema educativo de un país, una visión panorámica y concisa a partir de información clave.

En la Oficina de Planificación del Ministerio de Educación del Perú solíamos contar con un pequeño documento (no más de quince a veinte páginas) con ese carácter y al que informalmente llamábamos el **desasnador**.

Debo mencionar, sin embargo, que la utilidad del **desasnador** solía ir más allá de introducir a nuevas autoridades a la realidad del sector. Llamaría sobremanera la

atención el número de **expertos** que encontraba en dicho compendio información que le resultaba novedosa y muy reveladora.

Esta información básica comprendía aspectos de: (i) matrícula, instituciones y programas por niveles (para dimensionar el sistema); (ii) distribución por gestión y región; (iii) gasto público y privado en educación (por niveles y comparado con otros países); (iv) número de docentes por nivel y gestión; (v) perfil educativo de la población adulta; y (vi) indicadores básicos de cobertura, cobertura en edad, conclusión de estudios, repetición, etc.

Junto a esto, las autoridades suelen requerir información muy puntual de jurisdicciones administrativas o incluso de instituciones educativas específicas a las que tienen que visitar o respecto de las cuales se ha suscitado una situación que requiere su atención. La mejor solución a estos requerimientos, que son siempre imprevisibles y suelen tener carácter de urgente, es un sistema de consulta como el descrito anteriormente y contar con un ministro o un asesor competente en su consulta. En todo caso, un buen sistema de consulta puede también ser directamente operado por un funcionario de las unidades de información que obtendrá de modo ágil la información que se requiere.

Sin embargo, hay un requisito básico para que esto funcione y este está dado por la existencia de un lenguaje común. Si se reproducen los diálogos hipotéticos presentados al inicio de esta sección, no hay sistema de consulta de datos que sirva. Así, existe un trabajo previo de construcción de entendimientos mutuos en el que toca una gran responsabilidad a las personas a cargo de los sistemas de información, tanto en lo que corresponde a saber escuchar y buscar entender lo que se solicita como en cuanto a saber comunicar lo que se tiene, incluidas las posibles limitaciones de la información.

- Junto a los ministros y viceministros se encuentran todos aquellos involucrados directamente en la **formulación y seguimiento de las políticas educativas**.

Es en este terreno donde se presentan los desafíos de información centrales para aquellos interesados en que la formulación de políticas sea, al menos en un grado mínimo, consciente de la evidencia existente⁴³. Es decir, es necesario combatir un mal que se grafica en la siguiente frase, usualmente atribuida a John Maynard Keynes: «No hay nada que guste menos a un político que estar bien informado. Esto torna muy complejo y difícil el proceso de toma de decisiones»⁴⁴.

⁴³ Usualmente se habla de *evidence-based policy making* (toma de decisiones de política basada en evidencia); sin embargo, la política necesariamente se fundamenta en muchas cosas más y, como lo sugieren Nutley *et al.* (2002: 2) resulta más realista esperar una *evidence-aware policy making* (toma de decisiones de política consciente de la evidencia).

⁴⁴ Esta frase es citada en muchas partes identificando a Keynes como su autor; sin embargo, no he podido encontrar una referencia precisa, por lo que se cita como **atribuida** (traducción propia).

La ironía de la frase apunta a un hecho real: muchas veces agentes políticos actuando como tales o bajo ropajes distintos (investigadores, agentes de desarrollo, etc.) se ven incomodados por la existencia de evidencia que suele mostrar la realidad de una forma compleja que va más allá de lo que sus nociones o prejuicios sugieren. Una de las maneras más fáciles de cometer errores es tener la certeza absoluta de que se está en lo correcto y, por lo mismo, no estar abierto a escuchar evidencia u opiniones que contradigan lo que uno piensa o, peor aun, que comprometan algunos intereses.

Esto se vincula al hecho de que nadie toma ni puede tomar jamás decisiones basadas exclusivamente en la evidencia y también a que, al decir de Frans de Waal, «alguna gente ansía dogmas» (2013: 89). Los seres humanos percibimos la evidencia desde un marco de creencias, opciones, prejuicios, intereses, etc. que enmarcan su lectura y terminan por darle un sentido. Las decisiones son el resultado de ese complejo juego de factores siempre presente en toda acción humana. Por eso es que a veces es muy difícil para algunas personas, ya que esto incluso puede comprometer problemas de autoestima e inseguridades personales, aceptar algo que contradiga sus creencias más sentidas, algo que las obligue a revisar lo que postulan, incluso cuando esto se inspire en muy buenas intenciones.

Entonces, comunicar información a este nivel pasa necesariamente por la **integración de la evidencia en discursos significativos**, es decir, requiere labores de análisis. Esta es una tarea que las áreas de información pueden desarrollar por su cuenta (si cuentan con los recursos para ello) y/o en conjunto con otros actores. En cualquier caso, necesariamente conjuga el manejo de la evidencia (y sus atributos) con la comprensión de los temas centrales de política y el conocimiento especializado sobre temas educativos y, por eso mismo, es potencialmente una labor muy compleja.

Algunas visiones más propias del mundo anglosajón (y presentes también en otros medios) tienden a sugerir que el análisis debe ser hecho por otros para evitar que las unidades de información pierdan «objetividad» e independencia; por lo mismo, sugieren que solo «se debe permitir que los datos hablen por sí mismos». Mientras que el riesgo de pérdida de independencia es real y debe ser manejado como tal, postular que los datos «hablen solos» es evidencia de un empirismo decimonónico insostenible en el siglo XXI. Los **datos** nunca hablan solos, pues su mera existencia ya involucra un esfuerzo humano de construcción en el que se expresan valores y formas de ver el mundo que los datos portan consigo. Recordemos que los «datos» nunca son «recolectados» (como si fuesen frutos silvestres) sino que siempre son producidos o generados desde acciones humanas intencionales. No reconocer esto solo sirve para hacer pasar las creencias propias como hechos «naturales» y, por lo mismo, para imponer, aunque sea de modo inconsciente, una determinada visión del mundo.

En cualquier caso, el análisis (independientemente de quién lo conduzca) es la forma de convertir la evidencia en algo útil en un proceso de toma de decisiones. El análisis sirve para identificar tendencias, registrar progresos y retrocesos, dimensionar brechas, etc.

Las labores de análisis más útiles a este nivel tienden a ser las que de modo conciso y directo se enfocan en un solo tema y tratan de especificarlo, dimensionarlo, hacerlo asible para la política.

Asimismo, como ya se anotó, el análisis (tanto el académico como el orientado a la formulación de políticas) supone **ensuciarse las manos** conociendo a fondo los detalles de la evidencia, sus limitaciones y su potencial, y ese es un terreno donde las unidades productoras de información tienen una posición privilegiada.

La elaboración de documentos de trabajo de extensión breve (como para que sean leídos fácilmente por personas que tienen muchas ocupaciones y urgencias que atender en el día a día de la gestión) puede representar contribuciones mayores a la formulación de políticas y también al debate público en general.

Por ejemplo, es posible imaginar una situación en la que un grupo de actores (inspirados por una profunda buena intención y siguiendo una agenda política signada probablemente por agentes externos de financiamiento) promueven un conjunto de iniciativas para atender un problema que, por alguna razón, consideran crucial. En ese marco, un pequeño documento de trabajo puede hacer una contribución a, por ejemplo, dimensionar con propiedad dicho problema, sea para ratificar su urgencia, sea más bien para sugerir que seguir por esa vía puede llevar a un significativo desperdicio de recursos que podrían destinarse a urgencias mayores.

- Finalmente, se encuentran todos aquellos vinculados a la **gestión de programas e intervenciones educativas**. Este es uno de los problemas mayores, ya que existen demandas crecientes en este terreno.

Usualmente, hay dos aspectos en los que los sistemas de información pueden contribuir a la gestión de políticas y programas específicos. El primero refiere a cuestiones generales que tienen que ver con la identificación de problemas que dichas políticas o programas han de enfrentar. Es decir, un primer desafío está dado por las necesidades de información propias del diseño inicial de las acciones. En este terreno, es de aplicación lo planteado en el punto anterior sobre la importancia que tiene el análisis de la información para dimensionar desafíos y medir progresos globales.

Un segundo aspecto pertenece a la gestión cotidiana de dichos programas y políticas, y aquí es donde suelen producirse desencuentros entre los sistemas de información y la gestión. La gestión cotidiana suele requerir información de detalle y correspondiente a períodos temporales breves. Por ejemplo, el número de estudiantes que se

atiende con un programa de alimentación escolar y la evaluación de posibles atenciones brindadas a población no destinataria de la intervención es algo que requiere monitoreo continuo, lo cual lleva a que sistemas de información pensados en dar cuenta de fenómenos agregados en, por ejemplo, períodos lectivos anuales, no sea suficiente.

Este desencuentro se origina en algo que no es un problema en sí mismo: en la diferencia entre aquellos propósitos propios de los sistemas nacionales de información educativa y aquellos que corresponden a la gestión. Esta diferencia puede abordarse de dos formas distintas: (i) modificando los propósitos de los sistemas nacionales de información de manera tal que se correspondan más próximamente con las necesidades de la gestión cotidiana; o (ii) mediante la instalación de dispositivos de generación de información específica para los programas que no sea parte de (aunque deberían ser compatibles con) las tareas de los sistemas de información nacionales.

Ambas opciones son perfectamente legítimas y no hay una que sea intrínsecamente «mejor» que la otra. En el primer caso, es preciso tener cuidado con que el sistema de información nacional tenga la capacidad y los recursos para absorber estas nuevas tareas y lo haga sin renunciar a otros fines que tocan a los aspectos más básicos de la función pública ya reseñados. En el segundo caso, el riesgo principal a conjurar es que los dispositivos de información propios de los programas se diseñen sin colaborar con los sistemas de información nacional, ya que esto puede traer posibles efectos perversos, como duplicaciones, «competencia» que merma la credibilidad de ambos, mal diseño de los instrumentos y procedimientos de generación y procesamiento de información (los expertos de programa no tienen por qué ser expertos en la generación de información) que devenga en esfuerzos que después no den los resultados que deberían u otros efectos no deseados o perversos que se pueden verificar en el tiempo.

Un par de ejemplos pueden ayudar a ilustrar estos riesgos:

En cierta oportunidad, en un país de América Latina, el conjunto de instrumentos usados en la recolección anual de estadísticas educativas contaba con un subconjunto que no se capturaba y solo se almacenaban los cuestionarios reportados por sí «algún día hay tiempo para digitalarlos». Este subconjunto correspondía a instrumentos complementarios sobre lengua nativa de los estudiantes (se trataba de un país multilingüe). La indagación que hice sobre esta situación puede ser descrita de la siguiente manera:

- (i) En un determinado momento, la oficina encargada de educación bilingüe intercultural se había aproximado a la oficina de estadística a expresarle ciertas necesidades de información.

- (ii) A partir de este requerimiento, ambas unidades se pusieron a trabajar de modo conjunto y diseñaron los instrumentos a utilizarse para su generación.
- (iii) Una vez que los instrumentos se diseñaron, se incorporaron al conjunto que se administraba en el ciclo anual del censo escolar.
- (iv) Luego de unos meses de iniciado el ciclo de operación del censo escolar, la información generada empezaba a retornar a la oficina de estadística (toda la información usual en medios magnéticos, pues se capturaba descentralizadamente, y la información de los instrumentos adicionales en papel, ya que no se había diseñado «aún» el módulo para su captura a incorporarse en el *software* de captura que se usaba para todo el censo escolar). Así, los formatos en papel llegaban a la oficina y se empezaban a almacenar de modo ordenado para su posterior digitación.
- (v) Como siempre, las unidades de información deben atender diversas demandas y, así, estos formatos se empezaron a acumular sin ser digitados y, lo más importante, **sin que nadie reclame la información**. Así, pasaron al menos dos o tres años.
- (vi) Lamentablemente, no cuento con información acerca de lo qué pasó después.

Este es un claro caso de buena voluntad y colaboración, pero de desajuste entre las necesidades del programa (incluso tal y como las expresaron las personas a cargo de él) y la operación de los sistemas de información. Es claro que los iniciales demandantes de la información o bien no la necesitaban, o bien la requerían en plazos distintos de los que concordaron las dos unidades, de tal forma que, una vez que se hubiese podido acceder a la información (con un pequeño esfuerzo adicional para la digitación), nadie se interesó en ello. Esto, por otra parte, tiene impactos en términos del desperdicio de recursos y la sobrecarga inútil impuesta sobre los informantes.

Un segundo ejemplo refiere a una situación en la que el sistema de estadística educativa de un país había colapsado y, por distintas razones que no es el caso explicar en detalle, los esfuerzos de reconstrucción no eran fructíferos. En ese contexto, la poca información que se generaba resultaba de dudosa calidad y, además, era atesorada por algunos funcionarios que restringían el acceso probablemente para no exhibir sus propias debilidades.

En ese contexto, una ONG, que en realidad era el brazo operativo de una agencia de cooperación para el desarrollo de ese determinado país, decidió «colaborar» desarrollando su propio esfuerzo de generación y procesamiento de información. El mecanismo principal que diseñó para ello consistió en establecer contacto con los funcionarios del sector educación de las unidades administrativas locales y solicitarles directamente la información a cambio de una retribución monetaria.

Los principales resultados de este esfuerzo pueden reseñarse de la siguiente manera:

- (i) Se instaló un sistema de información paralelo al que era sufragado por los contribuyentes, debilitando la credibilidad de este y sin hacer ninguna contribución a su mejora.
- (ii) Se comunicó, como mensaje implícito en la acción, que los funcionarios locales responsables del reporte de la información no tenían por qué cumplir con su obligación, sino que, al contrario, debían convertir esta en una fuente de beneficios económicos adicionales puesta al servicio de un actor privado.
- (iii) Los datos así producidos tenían también serias limitaciones ya que, en primer lugar, la ONG en cuestión no tenía capacidad operativa para llegar a todo el país, con lo que no se subsanaba la falta de información nacional; en segundo lugar, la información que sí se generaba, se producía con métodos dudosos y completamente desconocidos para el público; y, finalmente, dicha información terminó formando un acervo privado de un individuo que la gestionaba según su (seguramente buena) voluntad, al punto de que dicho acervo era conocido como «los datos de *Ted*» en alusión al nombre de la persona «dueña» de dicha información⁴⁵.

Esfuerzos como este desvían recursos, deslegitiman el sistema público y finalmente, en este caso, hacen una importante contribución a que los problemas de fondo que impiden la reconstrucción del sistema de información se vean fortalecidos. Probablemente, algunas de las acciones de la ONG en cuestión se veían favorecidas por lo que a todas luces era un uso ritual de la información. De hecho, en ese país, donde aproximadamente uno de cada cinco jóvenes de quince a diecinueve años no ha culminado la educación primaria, «los datos de *Ted*» permitían decir que la conclusión universal de la educación primaria ya había sido lograda en dieciséis de las dieciocho provincias que constituyen el país⁴⁶.

En cualquier caso, es importante observar que siempre es preciso el diálogo entre las partes a efectos de aclarar los posibles puntos de desencuentro, si estos

⁴⁵ Evidentemente, *Ted* es un nombre ficticio que usamos acá para no identificar al personaje, pues lo que ocupa a este texto es el tipo de práctica y no el individuo.

⁴⁶ Al indagarse sobre este sinsentido, la única explicación dada por el propio «dueño» de los datos fue: «hemos seguido el procedimiento de cálculo que nos indicó una misión del Banco Mundial». Es decir, se siguió mecánicamente un procedimiento de cálculo sin ningún entendimiento de lo que se hacía. El problema de la medición de la conclusión de la primaria era, en primer lugar, un problema de procedimiento (más allá de si los datos estaban bien o no) y las personas a cargo no lo entendían, solo cumplían o ejecutaban un ritual. Gracias a esto, el Plan Nacional de Educación para Todos, formulado en dicho país en aquellos años, estaba lleno de sinsentidos.

son un problema (o solo una situación derivada de legítimas diferencias de propósito), y cómo abordarlos. La falta de diálogo es la que lleva a frases como: «La información de ellos no me sirve para nada, está lista muy tarde y no tiene el detalle que necesito» o «Piden cualquier cosa y no entienden que la producción de información requiere tiempo para la operación de todo el ciclo de diseño, generación y procesamiento».

Al mismo tiempo, es preciso jerarquizar los propósitos. Esto es algo que la experiencia sugiere que es muy difícil de lograr en el mundo de los especialistas en educación, que siempre encuentran que **todo es importante**. Seguramente todo es importante, pero si esto se traduce en la imposibilidad de jerarquizar acciones y/o establecer prioridades, solo lleva a la inacción o a la acción inoperante, lo que puede ser peor, pues implica un desperdicio importante de recursos.

Jerarquizar problemas y necesidades permite definir mejor qué es lo que debe atender primariamente un sistema de información nacional y, por lo mismo, trazar el terreno en el que otras acciones pueden no solo ser necesarias sino también profundamente deseables.

1.10 NIVELES DE OBSERVACIÓN Y ANÁLISIS

A lo largo de lo que va de este texto, de modo implícito se ha estado haciendo siempre referencia a los sistemas de información de carácter nacional. En efecto, ese es el foco de este documento. Sin embargo, esto no tendría por qué ser así, las necesidades de información existen a diferentes niveles como diferentes son los ámbitos de acción de distintos actores.

La opción por los sistemas nacionales de información a la que este documento adhiere no tiene que ver con que ellos sean más o menos importantes que otros, sino que se explica por la necesidad, por un lado, de enfocarse en algo específico, sin pretender abarcar tanto que el esfuerzo pierda significado y se torne vacuo, y también, por otro lado, de centrarse en algo que corresponde a la experiencia y conocimientos de quien escribe, de modo tal que lo escrito pueda representar una contribución.

Ahora bien, esto plantea que existen diversos universos posibles para el desarrollo de los sistemas de información y, en ese sentido, es menester precisar algunos términos y ámbitos.

En primer lugar, es necesario considerar que los sistemas de información educativa públicos tienen (o al menos esto es lo que aquí se postula) como *raison d'être* servir a las personas a efectos de que puedan ejercer plenamente su derecho a la educación. Esto, por su parte, significa poner en el centro de las preocupaciones el aprendizaje por parte de las personas y no la prestación de los servicios educativos por sí mismos. Uno de los problemas más complejos que es imperativo abordar se origina en el hecho de que la cotidianeidad de los organismos públicos responsables por la educación (léase ministerios) está más marcada por las necesidades de la gestión de un sistema escolar, del aparato administrativo con el que se opera y de la contratación de personal (principalmente maestros) que por la atención a las necesidades de aprendizaje de las personas: **los ministerios suelen estar abocados a gestionar escuelas y contratar maestros, no necesariamente a contribuir a que las personas aprendan o a asegurarlo.**

Esto, en segundo lugar, no es el resultado de falta de visión o mala voluntad de los actores, sino que se vincula a la propia dinámica de su accionar: mientras el aprendizaje ocurre siempre en experiencias concretas de carácter local, de grupo e individual, la gestión nacional opera mayormente con agregados y con presiones que se expresan a ese nivel (como las demandas de los sindicatos docentes). Lamentablemente, si acaso, son

pocos los momentos en los que la acción de los agentes sociales ha sido lo suficientemente fuerte como para poner en la **agenda diaria** las necesidades de aprendizaje⁴⁷.

Asimismo, lo anterior se vincula al hecho de que resulta quimérico (y carente de todo sentido) pensar que en un ministerio nacional es posible tomar decisiones que efectivamente se han de traducir en acciones similares, simultáneas y con un mismo fin en los diversos contextos que conforman un país. La ilusión de creer que un ministerio puede regular el detalle del proceso de aprendizaje no lleva sino a la frustración⁴⁸.

De hecho, hoy existen tendencias en planificación y políticas educativas que subrayan la necesidad de que las **políticas y acciones nacionales** sean cada vez **menos prescriptivas** (es decir, dejen de intentar decirle a la gente lo que tiene que hacer) y se conviertan más bien en **mecanismos habilitadores** para que las personas puedan desarrollar autónomamente sus acciones de una manera tal que deriven en aprendizajes. Esto, por otra parte, no es sino una respuesta más adecuada al hecho de que una educación de calidad requiere necesariamente atender la diversidad y es mucho más consistente con una idea de libertad que, en más de una ocasión, ha cedido ante la lógica burocrática. Como sostiene John Stuart Mill:

El desacierto comienza cuando, en lugar de invocar la actividad y las facultades de los individuos, y de las instituciones, [el Estado] los sustituye con su propia actividad; cuando, en lugar de informar, aconsejar y, si es preciso, denunciar, los somete, los encadena al trabajo o los margina, actuando por ellos. El valor de un Estado, a la larga, es el valor de los individuos que lo componen; y un Estado que pospone los intereses de la expansión y elevación intelectual de sus miembros en favor de un pequeño aumento de la habilidad administrativa, o en otras palabras, un Estado que empequeñece a los hombres a fin de que sean, en sus manos, dóciles instrumentos –incluso cuando se haga para buenos propósitos– llegará a notar que con hombres

⁴⁷ Un buen ejemplo, aunque se trata de un caso singular, está dado por las movilizaciones estudiantiles en Chile iniciadas en 2006. Estas movilizaciones organizadas alrededor de una preocupación por **mejorar la calidad de la educación**, llevaron a un cambio en la agenda política del país, de modo que se iniciara un proceso que condujo al cambio de la ley de educación hasta entonces vigente (ley, por cierto, promulgada por la dictadura en sus últimas semanas). Hasta ese momento, el cambio legislativo no era prioridad para ninguno de los actores políticos o gremiales a nivel nacional.

Otro ejemplo se observó en mayo de 2010 en La Paz, donde padres de familia salieron a las calles a reclamar la reanudación de las actividades lectivas afectadas por una huelga magisterial.

⁴⁸ Se atribuye a un ministro de Educación francés de hace unas décadas haber dicho una frase del siguiente tipo: «En este momento todos los alumnos del país matriculados en el grado x, están revisando la unidad y, de la materia z». Esta frase denota una profunda vocación y cultura burocrática orientada a la sobre-regulación y el control, así como una total falta de entendimiento de cómo funcionan los asuntos humanos. Lamentablemente, este tipo de práctica no está hoy en desuso; en diversos ministerios hay personas que conciben su tarea como una abocada a decidir lo que otros deben hacer y, consiguientemente, a lidiar con el desajuste entre sus deseos y la realidad, para lo que se crea mecanismos de control o premios y sanciones («incentivos») que, finalmente, tampoco funcionan, ya que ignoran la lógica local de la acción. Un poco de teoría social contemporánea es altamente recomendable en estos casos.

pequeños ninguna cosa grande podrá ser realizada; y que la perfección del mecanismo al que ha sacrificado todo acabará por no servir de nada, por carecer del poder vital que, con el fin de que el mecanismo pudiese funcionar más fácilmente, ha preferido proscribir (Mill 1869: 207)⁴⁹.

Tal discusión, para efectos de este texto, nos conduce en dos direcciones. Por un lado, indagar sobre el propósito de los sistemas nacionales de información y el de sus instrumentos y herramientas; y, por otro lado, examinar lo referente a los vínculos entre los sistemas de información y la gestión en sus distintos ámbitos.

1.10.1. EL PROPÓSITO DE LOS SISTEMAS NACIONALES DE INFORMACIÓN Y DEL DESARROLLO DE SUS INSTRUMENTOS Y HERRAMIENTAS

Como ya se anotó, los sistemas nacionales de información tienen como propósito contribuir a los procesos de toma de decisión así como al debate público sobre educación, por lo que, usualmente, la información que se requiere es de tipo agregado.

La información agregada, por su parte, puede tener varias formas: (i) valores típicos y distribuciones nacionales, o subnacionales; (ii) análisis nacionales o subnacionales de la diversidad o heterogeneidad de diversas situaciones; y (iii) modelos que sirven para hacer proyecciones en el tiempo o simular potenciales impactos de determinadas medidas de política.

En el primer caso, la información al mayor nivel de agregación suele ser suficiente. De hecho, los sistemas de información educativa antes de la difusión del uso de microcomputadoras solo eran capaces de hacer esto. Por ejemplo, una escuela reportaba sus datos de matrícula a la autoridad local, esta agregaba manualmente dicha información y la enviaba a la autoridad provincial, que a su vez agregaba manualmente la información local y la remitía a la oficina central del país. Así, esta última podía contar en períodos de tiempo muy breves con información consolidada para servir a este tipo de propósitos.

El segundo caso corresponde al uso de la información de detalle con la que ahora se suele contar (por escuela, por sección, por estudiante) para generar análisis imposibles de conducir con resúmenes agregados. En efecto, análisis de dispersión o heterogeneidad son ahora posibles gracias a que se cuenta con lo que se suele llamar microdatos.

Una de las paradojas de los sistemas de información educativa actuales es que mediante una importante inversión en infraestructura tecnológica, que permite la captura de datos con una gran desagregación, estos producen únicamente (o a veces no llegan siquiera a producir) reportes agregados nacionales que se podrían generar de un modo más barato e incluso más rápido con la tecnología previa basada en la agregación de reportes con lápiz y papel.

Los modelos de simulación o proyecciones, a los que se refiere la tercera forma de agregar información, son una herramienta de trabajo particularmente sugerente. En

⁴⁹ Traducción propia.

todos los casos, un modelo no es sino un sistema de relaciones entre variables, de forma tal que cambios conocidos o simulados en una o varias variables se traducen en cambios esperados o proyectados en otras, dados determinados supuestos.

Un modelo muy simple consiste, por ejemplo, en lo siguiente:

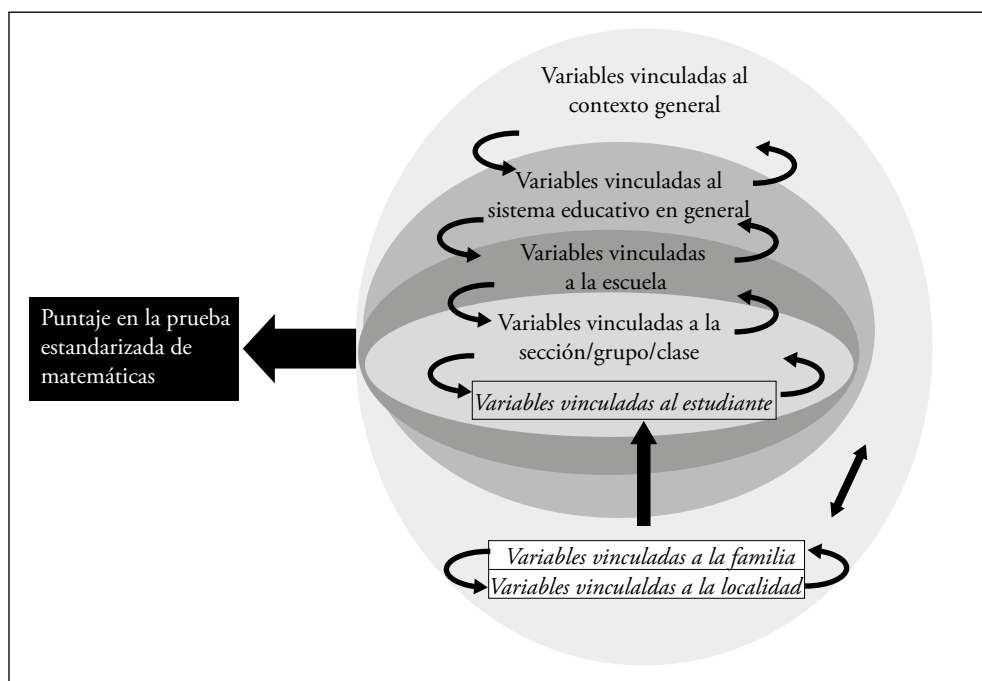
Donde:

$$Mat_x = Pob_x * CT_x$$

Mat_x : matrícula de personas de x años de edad
 Pob_x : población de x años de edad
 CT_x : tasa de cobertura total para la población de x años de edad⁵⁰.

Según este modelo simple, el volumen de matrícula de una edad dada (x) es una función del tamaño de la población de dicha edad y de la tasa de cobertura para la misma. De esta forma, cambios en la población o en la tasa de cobertura llevarán a cambios en el tamaño de la matrícula.

Gráfico 5. Modelo de análisis de logros académicos



Nota: se destaca con cursiva los factores que no son objeto directo de la acción educativa.

⁵⁰ En la siguiente sección se discutirá qué es una tasa de cobertura (y qué una tasa de matrícula), por ahora solo se presenta a efectos de ilustrar el ejemplo.

A partir de esta idea básica se puede desarrollar modelos mucho más complejos que, por ejemplo, incorporen en el lado derecho de la ecuación muchas más variables y asignen a cada una de ellas un **peso** (valor) y un **sentido** (signo) que defina su impacto sobre el lado izquierdo de la ecuación. Asimismo, es posible que cada uno de los elementos del lado derecho de la ecuación sea, a su vez, función de otro conjunto de variables. Así, uno puede llegar por ejemplo a un modelo como el del gráfico 5 (que es el modelo típico que se usa en los sistemas de medición de logro académico).

Nótese que cada uno de los conjuntos de «variables vinculadas a...» es un espacio en el que se puede incluir tantas variables como se considere necesario y las respectivas relaciones entre ellas.

En un escenario ideal, la investigación educativa puede permitir contar con un conjunto de evidencia empírica suficiente para construir modelos muy complejos que pueden tener distintos propósitos, como, por ejemplo, dar cuenta de los resultados en una prueba estandarizada de matemáticas.

Sin embargo, esta tarea no está exenta de problemas. Los más importantes se vinculan al hecho de que las distintas investigaciones dan resultados diferentes y a que no toda la evidencia disponible es consistente entre sí respecto de las relaciones entre las variables. Así, por ejemplo, existe un conjunto de investigaciones que sugieren que un menor número de estudiantes por clase tiene un impacto positivo en los aprendizajes; otras que señalan que los impactos dependen de rangos de dicho número; otras que no encuentran relación alguna entre ambas variables; e incluso algunas que encuentran evidencia de relaciones con impactos inversos a los que cabría esperar. ¿Cuál es la evidencia que uno incluye en el modelo? Evidentemente, la respuesta a esta pregunta es: depende. En efecto, la selección es siempre dependiente del criterio de quien construye el modelo y esto es así por la propia naturaleza de la labor cognoscitiva, es decir, no es una **limitación**, sino una **característica** del conocimiento. La limitación estaría dada por no documentar las elecciones hechas o por un uso acrítico del modelo que no se pregunta por esas decisiones y por el conjunto de supuestos que se encuentran detrás.

Otra área particularmente problemática está dada por el uso que se da o se pretende dar a los modelos. Supongamos que se cuenta con un modelo como el del ejemplo anteriormente descrito y este se encuentra bien documentado y recoge consistente y rigurosamente los resultados de un cuerpo significativo de la investigación educativa actual. ¿Cómo usarlo? El modelo puede perfectamente servir para simular resultados de diferentes medidas de política. Por ejemplo, se podría manipular la variable «número de estudiantes por clase» y ver cuáles serían los efectos esperados en los puntajes a obtenerse en las pruebas (siempre asumiendo que todos los supuestos del modelo se cumplen). Del mismo modo, se puede simular el impacto de distribuir textos escolares a cada estudiante, de que los docentes tengan un determinado nivel de certificación, etc.

Ahora bien, lo más importante de todo esto es entender que los resultados de esas simulaciones son **siempre un elemento más** a tener en cuenta en el proceso de tomar decisiones. Una de las cosas más nefastas que puede suceder es que se recomiende que las personas deben seguir acriticamente lo que el modelo sugiere. Las personas no pueden abdicar su responsabilidad por las decisiones que toman y entregarla a un modelo. Por ejemplo, si el modelo sugiere que el hecho de que los estudiantes cuenten con textos escolares no tiene un impacto significativo sobre las pruebas, entonces sería muy irresponsable derivar de ahí de modo directo que no es necesario invertir recursos en ello. ¿Por qué esto sería erróneo?

La respuesta a esta pregunta se explica por dos elementos centrales:

- (i) Los asuntos humanos se pueden modelar con importantes niveles de incertidumbre⁵¹, ya que estos nunca son idénticos y los elementos tomados de un contexto dado no son iguales a los de otro contexto. De hecho, **acciones similares en contextos diferentes probablemente conduzcan a resultados diferentes**⁵².
- (ii) El segundo elemento de respuesta tiene que ver con la relación entre el modelamiento explicativo (usualmente en búsqueda de causalidades) y lo que se valora en educación. No siempre lo que se considera valioso tiene un impacto sobre diversas variables que también importan, o incluso, lo primero puede tener impactos contradictorios y no por eso dejar de ser valioso.

Por ejemplo, es posible encontrar que escuelas autoritarias, donde el grado de disciplina es mayor, pueden lograr (controlando el efecto de otras variables)⁵³ mejores resultados en matemáticas, pero esto no debe llevar a concluir que hay que promover el autoritarismo en las escuelas. Hay cosas que valoramos socialmente así no tengan impacto en una variable a la que también se le asigna un valor particular.

De hecho, ciertas personas añoran los tiempos que graficara Goya en la pintura *La letra con sangre entra*, ya que asumen que la disciplina y la «motivación» logradas

⁵¹ Por cierto, este es también el caso de los eventos naturales, como se ha mostrado desde Heisenberg en 1926 y en muchos desarrollos posteriores en diversas áreas de las ciencias antes descritas como *exactas*.

⁵² Las acciones humanas derivan en resultados que dependen de estas y de su interacción con los contextos que son cambiantes. Véase por ejemplo Pawson y Tilley (1997) para entender la relación generativa con el contexto en el caso de los procesos de evaluación.

Por cierto, esto es lo que explica la famosa broma acerca de que los economistas destinan la mitad del tiempo a modelar y proyectar y la otra mitad a tratar de explicar por qué dichos modelos y proyecciones no fueron acertados.

⁵³ El concepto de control estadístico de variables se explica en el anexo 1. En alguna reunión internacional en la que se discutían ciertos asuntos técnicos vinculados al análisis de información, una representante de un país que no cabe identificar se opuso tenazmente al control estadístico de algunas variables ya que esto era percibido como no consistente con los principios dizque democráticos de su régimen político (que, por supuesto, era autoritario) que mandaban proscribir cualquier intento de «controlar» variables.

mediante el castigo físico son factores detrás del éxito académico que se asume (sin demostrar) tenían los sistemas educativos en el pasado.

Esto también es sugerido al observar otras variables. Por ejemplo, los resultados de PISA que se muestran en la siguiente sección (ver la página 110) sugieren que hay países con altos niveles de inversión por alumno que logran resultados similares a los de otros cuya inversión tiene menores niveles. Es probable que esto se explique por el hecho de que los primeros invierten en, por ejemplo, instalaciones deportivas, teatros y equipamiento para práctica musical, lo que posiblemente no tiene ningún impacto en los puntajes obtenidos en dicho estudio, pero sí un valor por sí mismo.

Esta discusión, por tanto, concluye en que los modelos sirven para **sugerir** o **suscitar** debates de política, no para **reemplazarlos**.

1.10.2 LOS VÍNCULOS ENTRE LOS SISTEMAS DE INFORMACIÓN Y LA GESTIÓN EN SUS DISTINTOS ÁMBITOS

Si los sistemas de información tienen como un propósito clave respaldar los procesos de toma de decisión, es claro que en cada uno de los diferentes niveles de operación del sistema las necesidades de información son diferentes, ya que los conjuntos de atribuciones que cada agente tiene son distintos.

Esto significa que no es posible asumir que la información que cumple un propósito en el plano nacional puede cumplirlo también a nivel, por ejemplo, local.

Así, mientras que la información sobre el desempeño docente es (o debería ser, como en cualquier otro empleo) de capital importancia para definir acciones de política de personal (promociones, despidos, incrementos salariales, etc.), esta información es importante solo para quien toma esas decisiones y entregarla a quien no puede hacer nada con ella puede llevar a crear o incrementar frustraciones debido a que no se cuenta con las facultades para afrontar un problema o situación dados.

Del mismo modo, la preocupación por la devolución de la información, especialmente a las escuelas, siendo sensata y muy bien intencionada, debe manejarse con cautela para no caer en situaciones que no ayudan mucho. Un ejemplo de esto es enviar a una escuela un pequeño boletín con información sobre su matrícula en el año lectivo precedente, cuando fue esa escuela la que proporcionó dicha información hace meses y ahora se encuentra preocupada por otros asuntos; otro ejemplo sería el de devolver indicadores que no se pueden calcular a nivel de escuela del mismo modo que se calculan a nivel agregado⁵⁴.

La diseminación de información, por tanto, debe partir de definir qué es lo que el usuario requiere y esto puede implicar diseños del sistema de información muy distintos, de forma tal que, eventualmente, el sistema nacional podría no estar en capacidad de responder a, por ejemplo, necesidades locales.

⁵⁴ Como las tasas de eficiencia interna interanual (promoción, deserción, repetición) discutidas más adelante.

Entonces, qué información se necesita es algo que debe definirse considerando las atribuciones que se tiene, de modo que se brinde la que pueda ser utilizada y, por lo mismo, tenga un valor para los usuarios. Esto (quién cuenta con qué información y en qué momento), aunado a la manera como se genera la información, significa que un sistema de información educativa a nivel de escuela o a nivel local no puede ser una simple réplica del sistema nacional.

Así, el presente acápite puede concluir con algunos puntos centrales a tener en cuenta:

- (i) El fenómeno educativo es observable en distintos ámbitos y niveles y con distintos propósitos, y es preciso tener en cuenta qué es lo que el sistema de información pretende abarcar a efectos de no generar falsas expectativas que terminen minando su credibilidad o de desinformar brindando información que carece de sentido o resulta equívoca en un nivel distinto de aquel para el que fue diseñada (como calcular tasas de repetición –siguiendo el modelo usual que revisaremos más adelante– a nivel escuela).
- (ii) Es de crucial importancia distinguir lo que es analizable a nivel del sistema de lo que corresponde analizar a nivel de las personas (que podrían no estar cubiertas por el sistema).
- (iii) Hay fenómenos que son observables a nivel individual (por ejemplo, niveles de desempeño), mientras que otros lo son a nivel del aula (por ejemplo, el clima de aula, el número de estudiantes en un grupo), de la escuela (por ejemplo, las características del equipo docente o del director) o mayores (por ejemplo, las características del *curriculum*).
- (iv) La observación y el registro de la información deben corresponderse con lo señalado en el punto anterior, mientras que los análisis pueden implicar la agregación o el tratamiento conjunto de registros individuales.
- (v) Un error típico que debe evitarse consiste en asumir que una persona o un equipo de personas que trabaja a nivel nacional en la construcción de indicadores educativos puede ser igualmente competente a la hora de construir indicadores a otros niveles que implican otro tipo de dinámicas y realidades. Un error similar se aplica a aquellos que trabajan en los sistemas de evaluación (medición de logros) cuando se cree que pueden desempeñarse con el mismo grado de éxito en cualquier otra instancia de evaluación educativa.

La similitud en los nombres (indicadores, evaluación) no implica que las realidades sean equivalentes. El conocimiento técnico especializado en las áreas mencionadas puede ayudar, pero no hay una traslación automática de las capacidades profesionales entre distintos ámbitos y, por lo mismo, siempre es crucial que el trabajo técnico se desarrolle en diálogo con los que cuentan con el conocimiento de los aspectos sustantivos a abordar.

SECCIÓN 2

SISTEMAS DE INDICADORES

Durante las últimas dos décadas se ha consolidado la tendencia a afirmar la necesidad de desarrollar indicadores para muchas cosas. Como muchas modas, esta es una que oculta cuestiones fundamentales, como el hecho de que en la vida diaria siempre contamos con «indicadores» de diferentes cosas o que el afán de construcción de indicadores impulsa a un paroxismo que lleva a producir información ininteligible.

Los indicadores no son nada más ni nada menos que «señales» respecto de cuestiones que son significativas para nosotros y, por lo tanto, nos permiten construir un juicio o valoración sobre esa situación que tiene sentido.

Así, un indicador no es necesariamente un cociente, una relación matemática entre datos cuantitativos o el resultado de una operación aritmética más o menos compleja. Es una señal cuyas propiedades dependen de aquello que él representa y de la necesidad de «señales» respecto de ello.

Un ejemplo cotidiano de conjuntos de indicadores está dado por el tablero de mando de un vehículo. Ahora bien, si tomamos un vehículo simple, como un automóvil de mediados de los años 1960, veremos que su tablero de control contiene información sobre velocidad (en kilómetros por hora), distancia total recorrida por el vehículo desde su fabricación (en kilómetros), combustible disponible (en fracciones del total posible) y algunas luces de alerta (bajo nivel de combustible, descarga de batería y presión de aceite)⁵⁵.

Si comparamos este tablero con el de un auto contemporáneo, veremos que este presenta mucha más información que aquel: tanto un conjunto de indicadores luminosos (estado del sistema ABS⁵⁶, de las bolsas de aire [*airbags*], de las luces de neblina, del

⁵⁵ La información anterior corresponde al tablero de mando de un Volkswagen Beetle de 1966.

⁵⁶ Sistema de freno antibloqueo.

cinturón de seguridad, de la presión de los neumáticos, etc.) como, junto a estos, información sobre velocidad (en kilómetros por hora), distancia total recorrida por el vehículo desde su fabricación (en kilómetros), distancia recorrida desde un punto o momento definido por el usuario (en kilómetros), combustible disponible (en fracciones del total posible) y frecuencia de rotación del eje del motor (en revoluciones por minuto).

A esto se puede agregar que muchos automóviles contemporáneos, al contar con un computador a bordo, proporcionan alguna de esta misma información siguiendo otros esquemas. Por ejemplo, la cantidad de combustible no solo como fracción del total, sino medida en litros y a partir de ahí como un valor estimado (dado el comportamiento típico de conducción actual definido según algunos parámetros) de la cantidad de kilómetros que se puede recorrer con tal cantidad de combustible. Asimismo, se incorpora crecientemente el uso de instrumentos de posicionamiento geográfico global (GPS) que junto a los mapas y el *software* correspondiente permiten contar con una ayuda visual y auditiva para la conducción.

Por supuesto que el mayor número de indicadores presentes en el tablero de un automóvil contemporáneo (en comparación con uno de hace cincuenta años) resulta pequeño si se compara con el de la cabina de mando de un avión comercial, o inmenso si se compara con el de una pequeña cortadora de césped con motor.

¿Qué es lo que sugieren estos ejemplos para nuestra discusión sobre indicadores?

En primer lugar, hay una variedad inmensa de posibilidades de indicadores para conducir un vehículo. Estas posibilidades son dependientes de:

- **Las posibilidades tecnológicas.** El indicador de estado del ABS no podía estar en el tablero del Volkswagen 1966, pues si bien dicho sistema se había inventado en 1929, su uso estaba limitado a algunos tipos de aviones, automóviles de carrera y muy pocos vehículos comerciales fabricados en los años 1960. Los sistemas ABS modernos para automóviles y camiones empiezan a introducirse en 1978.
- **Lo que socialmente se considera necesario.** Por ejemplo, si bien las bolsas de aire fueron inventadas en los años 1960, su uso empieza a ser normado como obligatorio mucho después (desde 1984 en los Estados Unidos). Al considerarse un elemento necesario, se torna también tal contar con un indicador de su estado de operación. De la misma forma, en la actualidad el uso de GPS empieza a ser obligatorio en varios países de Europa.
- **La manera como se conduce el vehículo.** Una aeronave puede tener un conjunto mucho más complejo de instrumentos e indicadores que un automóvil por razones que incluyen el hecho de que no es pilotada por una sola persona (entonces cada tripulante puede hacerse cargo de distintos instrumentos) y porque, salvo momentos como el despegue y el aterrizaje o situaciones puntuales de mucha turbulencia o potencial peligro, en una aeronave los tripulantes tienen tiempo

para hacer varios chequeos de instrumentos durante el vuelo. Si el tráfico aéreo fuera similar al tráfico de automóviles en tierra, eso sería imposible y habría que ampliar la tripulación y/o reducir el número de instrumentos a algo manejable por personas que no pueden dejar de mirar lo que pasa en su entorno inmediato. Por su parte, el tablero de una cortadora de césped con motor es completamente diferente de los otros dos, ya que está ajustado a las tareas que son propias de este tipo de máquina.

- **Los indicadores no son siempre números, ni cocientes, ni el resultado de fórmulas.** Por ejemplo, las señales luminosas presentes en el tablero de un automóvil indican la presencia de una anomalía⁵⁷ y tal información es **suficiente para formular un juicio**. Por ejemplo, si se enciende una luz que indica que un neumático está con menor presión de aire de la debida, eso es suficiente para considerar si cabe detenerse en la próxima estación de servicio a revisar y/o resolver el problema. Para tomar dicha decisión no es necesario conocer en ese momento cuál de los neumáticos presenta el problema, ni el nivel exacto de presión de aire; de hecho, contar con esta información adicional podría ser contraproducente, ya que leerla supone destinarle mayor atención que la requerida por un indicador luminoso simple, y esto puede implicar un riesgo de seguridad en la conducción.

Ahora bien, todos estos elementos de información cobran sentido en el conjunto del que son parte. Tal es el caso, en realidad, de todas las acciones humanas que no pueden ser entendidas fuera de un contexto.

Ejemplos de esto son, por ejemplo, las palabras que usamos. Ninguna palabra puede ser entendida de modo aislado, siempre es leída desde un universo de significados. Por ejemplo, ¿qué significa la palabra **actual**? Una misma palabra como **actual** (si es que puede decirse que es la misma, en todo caso esa misma secuencia de grafías eslabonadas en una sola unidad) tiene un significado diferente en castellano y en inglés. Por lo tanto, no puede ser entendida si no tenemos en cuenta (de modo explícito o implícito) cuál es el marco de significados al que pertenece.

De hecho, al formular la pregunta sobre su significado en el párrafo precedente, lo más probable es que el lector haya asumido el significado que esta palabra tiene en castellano, ya que la encontró inmersa en un texto en dicho idioma, y por lo tanto de modo «natural» se asumió un significado específico. Lo mismo puede suceder con

⁵⁷ El hecho de que la luz no esté encendida indica la ausencia de dicha anomalía. Es decir, las luces brindan información dicotómica no cuantificada sobre el estado de algún dispositivo. El estado del dispositivo (incluso si pudiese ser medido de modo numérico) es tratado como una variable nominal u ordinal (dependiendo del caso) con dos categorías: OK y No OK, Activado y Desactivado (si es nominal); Normal o Recomendable, Menos/Más de lo normal o Recomendable (si es ordinal). Sobre los niveles de medición o escalas de las variables, véase el anexo 1.

elementos de información. Un elemento definido en un contexto y trasladado a otro sin ningún control (es decir, sin tomar en cuenta su interrelación con el contexto en el que se le definió) puede dar lugar a sinsentidos. Tal es el caso de, por ejemplo, usar la razón número de alumnos por docente como una medida del tamaño de la clase o grupo, cuando esta es una medida creada para dar cuenta de necesidades de inversión relativa en recursos humanos, dada la matrícula.

De igual forma, los demás actos humanos solo pueden ser descritos si captamos su sentido dado un marco de significado particular. Un ejemplo clásico de la sociología brindado por Weber consiste en describir lo que hace una persona a la que vemos usando un hacha para cortar pedazos de madera. **Entender** lo que esa persona está haciendo requiere ir más allá de la descripción de las acciones físicas que desarrolla, ya que, en realidad, ella puede estar haciendo cosas muy distintas como: cortar leña para calentar su casa; desarrollar las tareas propias de su empleo remunerado; llevar a cabo una actividad física recreativa; entrenarse físicamente con un objetivo específico; desahogar sus iras y frustraciones; etc. La descripción «cortar leña» es absolutamente insuficiente para **entender** lo que la persona hace como un algoritmo de cálculo es insuficiente para definir un indicador.

Todo esto nos sugiere que la construcción de indicadores no suele ser un proceso aislado de la definición de un procedimiento de cálculo, sino más bien un proceso en el que se piensa (o se asume, como en el ejemplo de la palabra **actual**) un universo de referencia en el que la construcción del indicador o indicadores tiene sentido. Informar si las bolsas de aire de un auto se encuentran operativas o no es algo que hace parte de una preocupación por la seguridad vial en el marco de las convenciones actuales sobre este tema, y su inclusión o no en un tablero de instrumentos es decidida no solo en función de si es posible hacerlo, sino con relación a su **pertinencia**, dado el universo de preocupaciones que tocan al tema de la conducción de un vehículo hoy. Así, un indicador es definido como parte de un conjunto de información que se considera valioso y que conforma una unidad en la que cada una de las partes tiene un sentido.

2.1 ¿POR QUÉ UN SISTEMA DE INDICADORES Y NO UN LISTADO?

Los párrafos precedentes sugieren la respuesta a esta pregunta. Los indicadores no conforman un sistema por la mera agregación de elementos singulares a la manera como las papas en un saco forman un saco de papas. Ellos son piezas de un conjunto que tiene sentido. Por eso, se suele hablar de sistemas de indicadores y no de indicadores a secas.

El sistema de indicadores proporciona el marco de significados en el que cada pieza de información cobra sentido y –como el tablero de instrumentos de un automóvil– cristaliza una visión de lo que se considera importante, necesario, valioso y significativo, por lo mismo, se asocia a una agenda de preocupaciones: en nuestro caso las de aquellos que participan en acciones públicas. El sistema de indicadores debe pues enriquecer su comunicación, la negociación entre ellos y la toma de decisiones (European Commission 1999), ya que, como se ha señalado desde el inicio, la construcción de información, siendo una tarea con un alto componente técnico, solo cobra sentido con relación al conjunto de preguntas de política que con dicha información se busca contribuir a responder⁵⁸.

Así, una mirada cándida a este tema puede derivar en situaciones como la siguiente: dado que la OECD cuenta con un sistema de indicadores educativos robusto⁵⁹ y que este parece ser bueno para los países desarrollados, ¿por qué simplemente no lo adoptamos y listo? Esta es una opción válida en el sentido de que permite ahorrar algunos esfuerzos; sin embargo, el éxito de dicha empresa no está garantizado a menos que la agenda de política a la que se quiere servir sea similar a la de la OECD.

Es decir, la OECD es una organización internacional y su sistema de indicadores tiene como foco contar con información internacionalmente comparable. ¿Se quiere servir a ese mismo propósito en un país determinado? Asimismo, el sistema de indicadores educativos de la OECD atiende a una agenda política de dicha organización, la misma que define cuáles son sus preocupaciones prioritarias: ¿el país en cuestión tiene una agenda similar? Si se consulta la publicación emblemática de la OECD sobre estadísticas educativas (OECD 2013a), se encontrará mucha información muy interesante, pero ni una palabra sobre, por ejemplo, los niveles de acceso y conclusión de la educación primaria. Este, que es un tema importante para la mayor parte de los países del mundo, no lo es para la OECD, ya que aproximadamente el 97% de su población de quince

⁵⁸ Véase también Rutkowski (2008).

⁵⁹ Su manual de indicadores educativos se encuentra en OECD (2004).

años se encuentra estudiando la educación secundaria, es decir, accedió y culminó la primaria y luego prosiguió sus estudios⁶⁰.

Lo mismo sucede con, por ejemplo, la imperiosa necesidad de desarrollar los aspectos relativos a los docentes en los sistemas de información educativos. Uno podría basarse nuevamente en los importantes trabajos hechos por la OECD (Santiago 2002), sin embargo, notará que estos esfuerzos tienen un claro foco de preocupación vinculado a la dinámica del mercado laboral docente y este foco no es necesariamente idéntico al que ocupa a otros países no miembros de la OECD.

Debe quedar claro que un sistema de información muy robusto puede no ser transferible a otras circunstancias a menos que haya cierto acuerdo en los aspectos sustantivos centrales que lo sustentan.

Un sistema de indicadores, como ya se anotó, denota una visión de un tema y una agenda respecto del mismo y, además, los indicadores singulares están entroncados en dicha visión y en tal agenda. Por lo mismo, es preciso ser explícito respecto de la visión que se adopta, tanto por transparencia como para brindar los elementos necesarios para poder hacer uso de la información. Nuevamente, un problema mayor está dado por afirmaciones del tipo: «¿De veras necesitamos un marco conceptual para nuestros indicadores?, ¿no basta con listarlos y ya?», ya que esta actitud usualmente supone que, por defecto, se adopta una visión que se asume como válida para todos.

Las visiones que subyacen a un sistema de información pueden ser el resultado de opciones explícitas o de la asunción acrítica de una determinada perspectiva que se da por sentada y no se somete a revisión. La primera opción permite mejorar, la segunda produce ensimismamiento en un sistema de creencias (de grupo, organización, cultura, etc.) que, usualmente a partir de prejuicios etnocéntricos y prácticas antidemocráticas, se asume como poseedora de validez universal⁶¹.

⁶⁰ Este valor excluye a México (si se le incluyese descendería de modo sustantivo).

⁶¹ Ciertamente resulta agobiante tener que lidiar con afirmaciones soberbias (y profundamente ignorantes) como: «En los países subdesarrollados los censos son de baja calidad» o «Es muy difícil hacer algo como esto en un país como x, donde ni siquiera saben cuánta población tienen». Lo peor es que mucha gente convive con estas generalizaciones absurdas que se asumen como válidas por venir de posiciones privilegiadas en términos de la distribución global de poder; es decir, la falacia del argumento *ad verecundiam* se convierte en una falacia del tipo *ad crumenam* (algo es cierto porque el que lo dijo tiene mayores recursos o dinero).

2.2 EL SISTEMA DE INDICADORES

De la discusión previa se desprende que, si estuviésemos ante la posibilidad de diseñar un sistema de información educativa *ex nihilo*, sería necesario tratar de pensar el conjunto de los elementos de dicho sistema (incluyendo el marco de indicadores necesario) de manera conjunta.

Esta situación no se presenta en la realidad, ya que los sistemas de información, con sus defectos y virtudes, existen y operan. Justamente dentro de una de las acciones más nefastas que existen está aquella practicada por algunas nuevas autoridades que se lanzan a «reconstruir» todo arrasando con las experiencias previas, lo que no permite ningún aprendizaje institucional. Por lo tanto, lo que cabe es combinar un ejercicio mental orientado al diseño del sistema que, al mismo tiempo, sea capaz de recuperar lo existente, potenciarlo y rectificar lo que quepa rectificar.

Así, si bien las condiciones reales hacen imperativo aprender de lo ya vivido, esto requiere enriquecerse por una mirada renovada que se plantee de modo permanente las preguntas claves que hacen al diseño general del sistema y no que se aborden los desafíos nuevos como eventos aislados.

Lo anterior se traduce en que el diseño de un sistema de indicadores educativos donde no se cuente con uno, o su mejora si ya existiese, no debería partir de preguntarse qué más se puede hacer con lo que se tiene, sino qué es lo que se necesita; y, luego, ver cómo lo existente se ajusta a lo que se necesita o, en su defecto, cuál es el camino que hay que seguir para ello.

Existe una amplia bibliografía sobre sistemas de indicadores educativos que se puede consultar. Cada uno de los textos existentes (incluido este) debe ser leído buscando elementos que den sugerencias o alimenten la reflexión propia, no que la sustituyan, ya que cada esfuerzo ha de corresponderse con sus propios desafíos.

La Unesco ha producido diversos materiales a lo largo del tiempo. En todos los casos, una preocupación siempre presente en estos ha sido la comparabilidad internacional de los datos, lo que no tiene por qué ser una preocupación central para el trabajo nacional⁶². Entre estos materiales habría que destacar los siguientes:

- *Manual of Education Statistics (Manual de estadísticas educativas*; Unesco 1961).
- *A Study of Wastage at School (Un estudio sobre el desperdicio de recursos en las escuelas*; Unesco 1972), que debe destacarse por su impacto de largo plazo sobre las

⁶² Sin embargo, ello tampoco debe desdeñarse, pues el análisis comparado, si es desarrollado en serio (es decir, si no se limita a la elaboración de *rankings* –sobre esto véase la cuarta sección–), puede ser de mucha utilidad para fines nacionales. Sobre la lectura nacional de la información internacional, véase Tiana *et al.* (2004).

maneras en las que da cuenta de las trayectorias estudiantiles (modelos de cohortes reconstruidas).

- CINE en sus revisiones de 1976, 1997 y 2011 (Unesco-UIS 2013a; Unesco 1976, 1999a).
- *Education for All. The Year 2000 Assessment. Technical Guidelines (Guía técnica para la evaluación del progreso de la Educación para Todos 2000*; International Consultative Forum on Education for All 1998).
- La publicación electrónica *World Education Indicators (Indicadores mundiales de educación)* de 1999 (Unesco 1999b).
- Las recomendaciones sobre el tema aprobadas por diversas conferencias generales desde 1958.
- Las guías técnicas sobre indicadores publicadas en 2009 (Unesco-UIS 2009a).

En gran medida, el trabajo de la Unesco (que es reseñado en el anexo 2 junto a otros hitos en este ámbito) ha sentado las bases de lo que conocemos en el terreno de las estadísticas educativas con todas sus virtudes y límites.

A estos esfuerzos globales es necesario añadir el trabajo específico desarrollado por la Orealc-Unesco Santiago. En particular debe destacarse lo siguiente:

- Los trabajos iniciados por Ernesto Schiefelbein destinados tanto a mejorar la base empírica de los modelos de cohortes reconstruidas (Schiefelbein y Grossi 1984) como a presentar un mejor análisis de la situación educativa de la región en el marco del entonces llamado Proyecto Principal de Educación (Unesco Santiago 1988, 1990, 1992, 1996, 2001)⁶³.
- Los desarrollos alcanzados como agencia a cargo de los elementos técnicos del proyecto de indicadores de la Cumbre de las Américas, llamado Proyecto Regional de Indicadores Educativos (PRIE)⁶⁴. En particular, para el punto que nos ocupa, véase PRIE (2003) y Muñoz Izquierdo *et al.* (2004).
- El esquema conceptual y los desarrollos metodológicos cristalizados en Unesco Santiago (2007, 2008).

Simultáneamente, organizaciones como la IEA (International Association for the Evaluation of Educational Achievement) han desarrollado sus propios esfuerzos, especial-

⁶³ La Unesco Santiago ha producido dos informes de «situación educativa» posteriores al Proyecto Principal de Educación: el primero fue presentado ante una reunión ministerial a inicios de 2007 (Unesco Santiago 2008) e hizo parte de un esfuerzo mayor de conceptualización de los problemas de calidad de la educación, por lo mismo, debe ser consultado en conjunto con Unesco Santiago (2007); el segundo ha sido publicado en 2013 (Unesco Santiago 2013).

⁶⁴ Ver: <www.prie.oas.org>.

mente en el terreno de la medición de logros de aprendizaje, y la OECD en el de la estadística, la medición de logros de aprendizaje y la investigación educativa en general.

LA IEA es una organización independiente conformada tanto por instituciones nacionales de investigación como por agencias estatales y es la depositaria de más de cinco décadas de experiencia en el terreno de la medición de logros estudiantiles a nivel internacional, conduciendo estudios como Timms (Estudio Internacional de Tendencias en Matemáticas y Ciencias, que se centra en competencias en dichas áreas en estudiantes en cuarto –y desde 2011 quinto o sexto– y octavo grados), Pirls (Estudio Internacional de Progreso en Comprensión Lectora, conducido entre estudiantes en cuarto grado, aunque desde 2011 tiene una opción de aplicación en quinto y sexto grado), ICCS (Estudio Internacional sobre Educación Cívica y Ciudadana entre estudiantes de octavo grado) y TEDS-M (Estudio Internacional sobre la Formación Inicial del Profesorado de Matemáticas), entre otros⁶⁵.

La OECD ha producido el manual de indicadores ya mencionado (OECD 2004), además de sus volúmenes anuales de estadísticas comparadas (OECD 2013a) y diversos materiales vinculados al desarrollo de su PISA (que mide competencias de la población de quince años de edad) relativos a la medición de logros⁶⁶.

El trabajo de la OECD está presidido por un entendimiento particular de las realidades educativas que esta institución condensa en el modelo de la tabla 1.

Tabla 1. Marco para la construcción de indicadores educativos, OECD, 2012

	1. Productos y resultados de la educación y el aprendizaje	2. Elementos de política y contextos que definen los resultados	3. Antecedentes o limitantes que contextualizan la política
I. Individuos que participan en educación y aprendizaje	1.I Calidad y distribución de los productos de aprendizaje individuales	2.I Actitudes, compromiso y comportamientos individuales	3.I Características y antecedentes de los individuos
II. Entorno en el que se desarrolla la instrucción	1.II Calidad de los servicios de instrucción	2.II Prácticas pedagógicas y de aprendizaje y clima de aula	3.II Condiciones de aprendizaje de los alumnos y de enseñanza de los docentes
III. Proveedores de servicios educativos	1.III Productos de las instituciones educativas y su desempeño institucional	2.III Entorno y organización escolar	3.III Características de los proveedores de servicios y de las comunidades
IV. Sistema educativo en conjunto	1.IV Desempeño general del sistema educativo	2.IV Arreglos institucionales a nivel de sistema, asignación de recursos y políticas	3.IV Contexto nacional: lo educativo, social, económico y demográfico

Fuente: OECD (2012a: 17); traducción propia. Las palabras *outputs* y *outcomes* han sido traducidas como 'productos' y 'resultados', respectivamente.

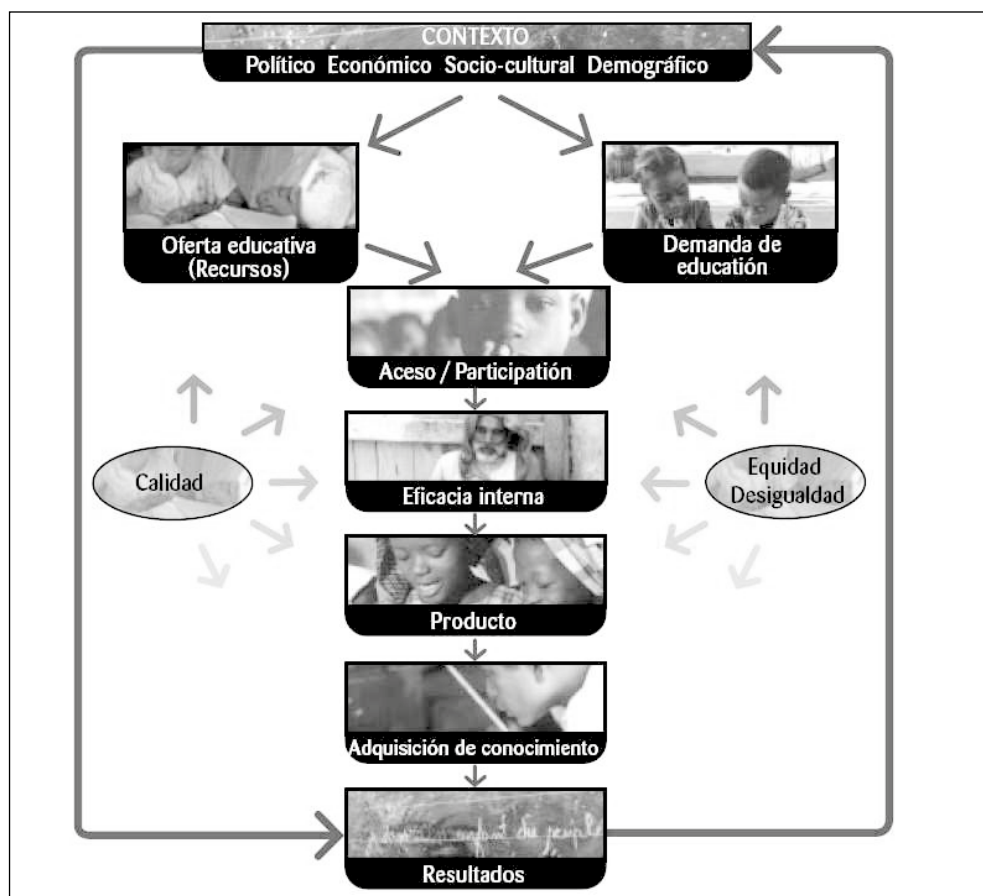
⁶⁵ Sobre las actividades de la IEA, véase <<http://www.iea.nl>>.

⁶⁶ Ver el documento ya mencionado de OECD y Unesco-UIS (2003) y el sitio web de PISA (<<http://www.pisa.oecd.org>>), así como los estudios de la OECD de competencias en adultos (OECD y Canadá-Statistics Canada 2000; OECD 2013b).

Del mismo modo, Mercosur Educativo inició un esfuerzo sistemático en el terreno de los indicadores educativos en 1997⁶⁷ y también un número importante de países han hecho esfuerzos que se traducen en documentos, en la forma como reportan su información en los sitios web de los ministerios y en sus contribuciones a los debates regionales sostenidos en diversos talleres regionales y reuniones técnicas de los últimos años⁶⁸.

Mutatis mutandis, todos los casos reseñados en este acápite hacen referencia a aspectos observables a nivel macro del sistema educativo y utilizan –excepto Guadalupe *et al.* (2002) y Unesco Santiago (2008)– un sistema de indicadores inspirado en una forma de ver la operación de los sistemas educativos que se representa en el esquema del gráfico 6.

Gráfico 6. Marco de indicadores y modelo analítico de la Unesco, 1999



Fuente: Unesco (1999b).

⁶⁷ Véase: <http://www.mercosur.int/t_ligaenmarco.jsp?title=off&contentid=854&version=1&channel=secretaria>.

⁶⁸ Como ejemplo de esfuerzos nacionales, véase Argentina-Ministerio de Educación, Ciencia y Tecnología (s. f.) y el capítulo 1 de Guadalupe *et al.* (2002).

La principal variante de este modelo, introducida por la OECD y por aquellos más directamente vinculados a los estudios de medición de resultados de aprendizajes, consiste en considerar la naturaleza multinivel de los fenómenos educativos. Esto es, supone que el aprendizaje, siendo una experiencia individual (primer nivel), se da en un contexto inmediato de pares y docentes (segundo nivel: aula) enmarcado en una dinámica institucional (tercer nivel: escuela y su entorno) que, a su vez, hace parte del contexto mayor (cuarto nivel: contexto nacional). El gráfico 5 (página 90) y la tabla 1 (página 103) muestran claramente estos niveles.

Sin embargo, la pertinencia actual y, por ende, el potencial de este sistema de indicadores requieren ser objeto de revisión crítica, ya que dicho sistema fue diseñado en un contexto diferente, con preocupaciones de política distintas de las que ocupan hoy la agenda educativa tanto a escala global como en América Latina y el Caribe en particular.

El hecho de que el documento mencionado de la Unesco Santiago, *Situación educativa de América Latina y el Caribe: garantizando la educación de calidad para todos* (2008), justamente se haya distanciado de este esquema ilustra que dicho volumen constituye un esfuerzo de reflexión crítica y renovación que fue positivamente acogido por los ministros de la región y, además, hace parte de un contexto general en el que otras iniciativas también resultan concurrentes a efectos de desarrollar visiones de la educación que sean consistentes con los principios y valores adoptados por la comunidad internacional –por ejemplo Pigozzi (2006)–⁶⁹. Este tema será abordado en detalle en el acápite siguiente (2.3, «Calidad de la educación») de esta misma sección.

Finalmente, cabe distinguir los elementos de información que son necesarios a efectos de dar cuenta de los sistemas educativos en general y de las aproximaciones centradas en el monitoreo del progreso hacia metas educativas. En efecto, mientras los primeros tienen una vocación holística y buscan dar cuenta de distintos aspectos que importan, dada una visión del sistema, las segundas tienen un elemento de mayor pragmatismo y se centran en corroborar niveles de progreso respecto de determinados objetivos y metas.

En este último caso, los sistemas de indicadores son subsidiarios de las metas propuestas –como es el caso de OEI (2010)– y la tarea principal consiste en definir el conjunto de indicadores que representen **parsimoniosamente** el estado de los fenómenos respecto de los cuales se han planteado las metas. En estos casos, la tarea de definición de los indicadores es dependiente de cuán específica sea la definición de las metas.

Por ejemplo, definir como meta educativa que la educación primaria debe ser universalizada es menos específico que plantear que la conclusión de la educación primaria debe ser universal, ya que en el primer caso no se especifica qué es exactamente lo que debe ser universal (por ejemplo, el acceso o la conclusión).

⁶⁹ En este marco llama la atención que Unesco Santiago (2010) se retrotraiga hacia el esquema conceptual anterior de una manera bastante rígida y con un texto particularmente críptico.

Del mismo modo, plantear que la conclusión debe ser universal es menos específico que decir que todos los niños deben terminar la educación primaria, ya que en este segundo caso se acota explícitamente la población de referencia a la que dicha universalización refiere. En el primer caso, es perfectamente posible (y deseable) pensar que el derecho a culminar la educación primaria lo tienen todas las personas y, por lo tanto, los desafíos a enfrentarse implican tareas que tocan no solo a la educación regular de niños, sino también a la educación de jóvenes y adultos, en tanto y en cuanto sea necesario corregir rezagos y marginaciones derivadas del no acceso o no conclusión en el pasado.

Siguiendo con el mismo ejemplo, decir que todos los niños deben terminar la primaria es menos específico que afirmar que todos los menores de, digamos, quince a dieciocho años deben hacerlo. Para efectos de medir los niveles efectivos de conclusión, la población de referencia debe contar con una edad mínima que le haya permitido culminar el nivel. No es posible medir la conclusión de la educación primaria entre, por ejemplo, niños de ocho años. Esto lleva a que medidas efectivas de la conclusión solo sean posibles para una población de al menos una edad dada, pero esto plantea el problema de que dichas medidas, en realidad, dan cuenta de un fenómeno presente, que es resultado del comportamiento pasado del sistema. Para mostrar el **comportamiento presente**, se necesitan medidas **prospectivas** que, tomando en cuenta la situación presente, anticipen una situación futura (por ejemplo, cuántos de los niños en edad de empezar la educación primaria hoy habrán de culminarla)⁷⁰.

En el otro caso mencionado, los sistemas que dan cuenta de modo general de la situación educativa y de la construcción de estadísticas e indicadores son dependientes del esquema conceptual que se utilice para responder acerca de lo que se considera importante. Estos esquemas conceptuales son contingentes a los debates académicos y políticos. Por ejemplo, en los años 1970, pocos países consideraban que la medición estandarizada de logros de aprendizaje era un tema clave, ya que se asumía que la aprobación de un determinado grado de estudios se traducía, *grosso modo*, en los aprendizajes prescritos.

En este sentido, los sistemas de estadísticas e indicadores son permeados por los debates académicos (¿cuál es el mejor indicador para dar cuenta del fenómeno *x*?), las agendas políticas y las visiones ideológicas que definen qué es lo que se considera relevante. Así, como lo señala Robinson (2011; ver también Ted 2006, 2013), existe una visión común según la cual medir los aprendizajes logrados en matemáticas y comunicación o lenguaje es más importante o central que, por ejemplo, medir la capacidad creativa de los niños. Esta opción trasunta una determinada forma de ver las cosas que es distinta de una visión más humanista y menos instrumental o academicista de la educación. Lo mismo se aplica a las visiones de temas como los relativos a los docentes, que están

⁷⁰ Estos son los temas que ocupan a Guadalupe y Louzano (2003) y Guadalupe y Taccari (2004).

más centradas en la dinámica del mercado laboral docente (Hanushek y Pace 1995; Hanushek *et al.* 2004; Murnane 1991; OECD 2005) o en lo que ellos hacen o dejan de hacer como actores (Akyeampong y Stephens 2000; Day *et al.* 2006; Kennedy 2004).

Sin embargo, a pesar de las diferencias entre sistemas destinados a monitorear metas y sistemas, es importante tener en cuenta que los objetivos y metas son establecidos en función de visiones de lo que se considera deseable y, por lo mismo, ambas aproximaciones tienen un sustrato común. La verificación del cumplimiento de metas no escapa a la discusión general, ya que la propia definición de metas se deriva de ella.

En cualquier caso, la definición y el seguimiento de metas es de primera importancia en la función pública (la acción pública busca garantizar el derecho a la educación expresado, por ejemplo, en asegurar que todas las personas hayan culminado la educación primaria o en que su acceso a las oportunidades educativas no se encuentre restringido mediante prácticas discriminatorias) y, para efectos de este texto, los ejemplos de metas que habrán de tratarse corresponden, principalmente, a las fijadas a nivel internacional en el marco de la iniciativa Educación para Todos (EPT)⁷¹.

⁷¹ Nótese que justamente la iniciativa EPT surge en 1990 con una vocación a favor de una «visión ampliada de la educación» que se expresa en que las metas establecidas representan aspiraciones concretas e ineludibles, aunque no agotan la agenda educativa (Buchert 1995; World Conference on Education for All 1990). Dado que aún no se aprueban los nuevos SDG, aquí se ha optado por mantener la referencia, a modo de ejemplo, en las metas de la EPT. En todo caso, respecto de la discusión sobre las nuevas metas internacionales, recomiendo revisar (Learning Metrics Task Force 2013a).

2.3 CALIDAD DE LA EDUCACIÓN

En los últimos años, las referencias a la **calidad de la educación** como un área de particular interés para las políticas públicas y el quehacer ciudadano han cobrado creciente importancia. De hecho, es posible afirmar que la preocupación por la «calidad» educativa se ha vuelto un lugar común en el discurso educativo⁷².

Un claro ejemplo de esto se encuentra en la iniciativa de la EPT que, a partir de la Conferencia Mundial de Educación de Dakar en 2000, se plantea un objetivo específico relativo a la mejora de la «calidad de la educación» en sus distintos aspectos, aunque no define con precisión qué ha de entenderse por esto. En atención a este objetivo y a la necesidad de precisar su contenido y alcance, el informe mundial de seguimiento de la EPT del año 2005 se abocó, justamente, a analizar **el imperativo de la calidad** (Unesco 2004a)⁷³.

Otro ejemplo se encuentra en el informe de progreso educativo 2006 del Programa de Promoción de la Reforma Educativa en América Latina y el Caribe (Preal 2006), que destaca que la región avanza en **cantidad**, pero no en **calidad** (sobre esta dicotomía volveremos más adelante) y, sobretodo, el importante desarrollo de sistemas de «medición de la calidad educativa» en muchos países del mundo y estudios internacionales sobre el mismo tema⁷⁴.

En este marco de discusiones, la Orealc-Unesco Santiago desarrolló entre 2003 y 2007 un proceso de reflexión destinado a esclarecer los términos de este debate en consonancia

⁷² El presente acápite se basa en un gran número de discusiones sostenidas con el conjunto de integrantes del equipo profesional de la Orealc-Unesco Santiago entre los años 2003 y 2007 y parte de un texto preliminar no publicado que fue presentado en la reunión preparatoria del Comité Intergubernamental del Prealac (Santiago de Chile, 6 y 7 de diciembre de 2005). Así, los siguientes párrafos (hasta antes del gráfico 8) retoman, con la generosa autorización de Orealc-Unesco Santiago, textos que originalmente redacté para dicho documento (a efectos también de dejar registro impreso de un documento que, de otra forma, estaría hoy perdido). Asimismo, las ideas aquí expuestas son las que sustentaron la estructura y contenidos de Unesco Santiago (2008) y son armónicas con lo desarrollado en Unesco Santiago (2007). Por lo tanto, ellas pertenecen al equipo profesional de la Orealc-Unesco Santiago en dichos años y, en particular, ha sido posible gracias al aporte de Rosa Blanco, Ricardo Hevia, Carolina Hirmas, Ana Luiza Machado y todo el equipo del SIRI.

⁷³ El equipo responsable del informe mundial de seguimiento de la EPT es un equipo independiente al que la Unesco, como agencia que lidera la EPT, aloja institucionalmente. Por esta razón, es importante tener en cuenta que los informes mundiales que este equipo produce no son «informes de la Unesco» sino de dicho equipo independiente.

⁷⁴ A la fecha, cerca de 150 países en el mundo miden niveles de aprendizaje de sus estudiantes de alguna manera mediante evaluaciones muestrales, censales, exámenes de certificación, estudios internacionales, etc. (Learning Metrics Task Force 2013a: 45).

con los mandatos y principios de la Unesco, sentando las bases para la descripción y el análisis de la situación educativa regional desde una perspectiva que atienda a las características del debate actual y brinde, como es rol de la Unesco, elementos de juicio que potencien la capacidad de los Estados miembros para abordar sus propios desafíos. Simultáneamente, en la sede central de la Unesco se iba desarrollando un debate similar cuyos resultados se expresan en Pigozzi (2006).

2.3.1 EL CONCEPTO DE CALIDAD DE LA EDUCACIÓN

El término «calidad de la educación» está compuesto de conceptos cuyo contenido y alcance no pueden ser tenidos por dados y, por lo mismo, implican potenciales problemas que requieren ser esclarecidos a efectos de asegurar que el contenido que se quiere que la frase transmita sea claro y pertinente. Sin dicho esclarecimiento, las labores necesarias de generación y análisis de información para el seguimiento, monitoreo y evaluación de la «calidad de la educación» resultarían poco claras.

A pesar del consenso generalizado sobre la necesidad de proporcionar una educación de calidad, existen diferentes enfoques y discusiones sobre lo que esto significa. Por tal razón, el acuerdo prácticamente unánime que existe sobre la necesidad de mejorar la educación podría, en realidad, ser una coincidencia superficial que vela discrepancias sustantivas de enfoque que subyacen a distintas posturas (McCowan 2013). Dicho de modo sumario, los discursos van desde aquellos que identifican la calidad con los resultados de aprendizaje, tal y como pueden medirse mediante pruebas estandarizadas, hasta aquellos (como la postura de la Unesco) que afirman una visión amplia e integral de la misma, aunque entre estos existen diversas posturas que priorizan distintos elementos⁷⁵.

La primera vertiente de pensamiento destaca la centralidad de los logros de aprendizaje, ya que estos aparecen como el **resultado** o **producto** concreto más importante que debe esperarse de la operación del sistema educativo. Esto revela, en primer lugar, que el sistema educativo es visto e interpretado en base a una analogía que lo asemeja a una operación de producción (vista desde una perspectiva mecánica e instrumental), donde los procesos educativos y los recursos que a estos se destinan son subsidiarios al **producto**. De ahí que lo central sea observar este último. Por otra parte, se tiende a asimilar la observación de los resultados a lo observable mediante pruebas estandarizadas de logro académico en áreas que se consideran claves⁷⁶. Este último elemento introduce

⁷⁵ Para una discusión amplia de las diversas perspectivas, véase Unesco (2004a). Don Adams (1993: 7 y ss.) presenta un sumario diferente en el que reconoce que el término calidad es usado con al menos seis significados diferentes: reputación; calidad de los insumos y procesos; calidad de los procesos; calidad de los contenidos; calidad de los resultados e impactos; y «valor agregado». Véase también la discusión en Acedo *et al.* (2012), así como en Unicef y Unesco (2007).

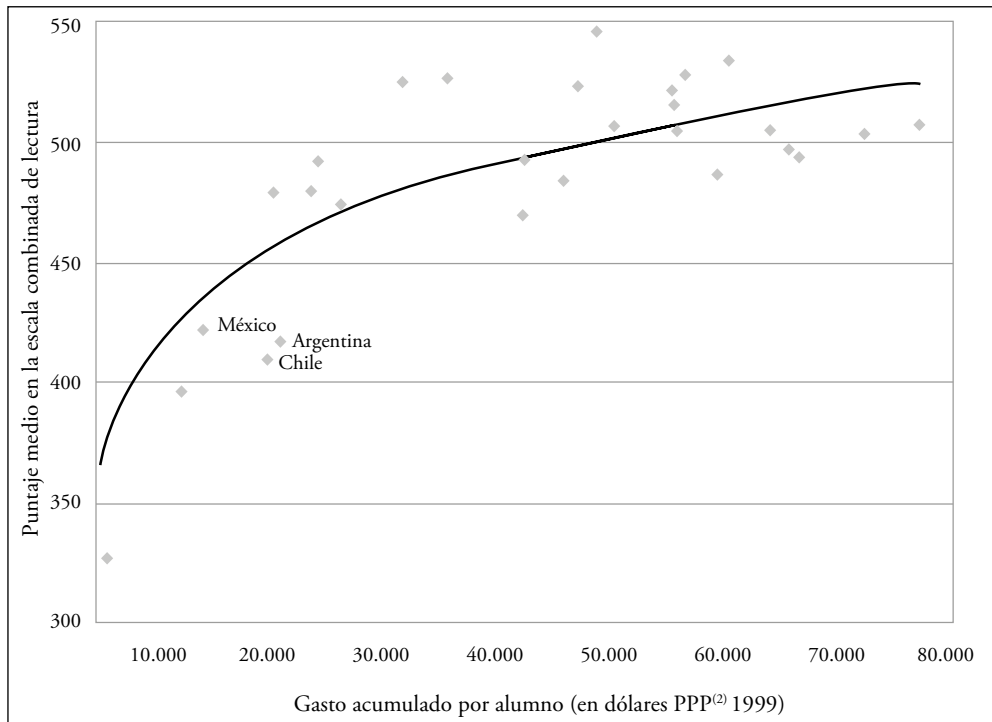
⁷⁶ Ejemplos de esta visión pueden encontrarse en los informes del estudio PISA de la OECD, por ejemplo OECD (2014), o en Unesco Santiago (2010).

un componente de reduccionismo en el enfoque que no suele traducirse en análisis certeros.

La segunda vertiente de pensamiento, por su parte, destaca la complejidad del fenómeno educativo y busca desarrollar una idea de calidad que no se sobresimplifique mediante el uso de una analogía mecánica y tome en cuenta los distintos elementos que se consideran valiosos. Un ejemplo permite ilustrar cabalmente la diferencia entre ambos enfoques.

El gráfico 7 presenta los puntajes promedio obtenidos por los países en la escala combinada de lectura usada por PISA en su primera ronda (2000+⁷⁷) y compara estos puntajes con los niveles de gasto en educación acumulados por estudiante hasta la edad de quince años (edad de la población objetivo del PISA).

Gráfico 7. Puntajes medios en la escala combinada de desempeño lector y gasto educativo acumulado por alumno hasta la edad de quince años, PISA, 2000+⁽¹⁾



Notas

⁽¹⁾ Versión simplificada del gráfico presentado en Unesco (2004a: 119), elaborado a partir de los datos contenidos en OECD y Unesco-UIS (2003).

⁽²⁾ PPP: corrección por paridad de poder de compra (*Purchasing Parity Power*) que el dólar tiene en cada país. Así, no se trata de dólares nominales, sino corregidos según este factor para el año dado.

⁷⁷ El signo + indica que algunos países condujeron el estudio en el año 2001.

El gráfico 7 muestra que existe una relación entre ambas variables, de suerte que mayores niveles de inversión acumulada se corresponden con puntajes medios más altos, aunque dicha relación se debilita a partir del punto en que se alcanza cierto umbral de gasto⁷⁸.

Si seguimos la lógica de la primera vertiente antes mencionada, el gráfico 7 permite concluir que a mayor volumen de insumos (aproximado por los niveles de gasto), mayor el producto logrado. Asimismo, muestra algunas comparaciones dignas de destacar. Como se señala en el «Resumen regional para América Latina y el Caribe» (del informe mundial de seguimiento de la EPT), el gráfico 7 lleva a concluir que

[...] los alumnos de México, Chile y Argentina obtienen puntuaciones similares en competencias de lectura, cuando el gasto por alumno en México es de 12.189 dólares, mientras que Chile y Argentina alcanzan respectivamente la suma de 17.820 y 18.893 dólares, calculados en paridad de poder adquisitivo. Esto suscita el problema de la utilización eficaz de los recursos (Unesco 2004b: 3).

Considerando que los autores plantean este tema en el marco de la discusión sobre la calidad, resulta adecuado releer este párrafo en los siguientes términos, que se corresponden a un castellano más directo: «México es más eficiente que Chile y Argentina pues logra la misma calidad educativa con menores niveles de inversión»⁷⁹.

La frase parece inocente y perfectamente apegada a lo que la evidencia sugiere; sin embargo, hay un problema derivado del foco reduccionista en los resultados tal y como son representados con una prueba dada. Este problema es muy simple: la población objetivo de PISA era la población de quince años matriculada en la educación secundaria. En números gruesos, esta representaba en México la mitad de la población total de dicha edad, cuando en Chile y Argentina dicha proporción era aproximadamente 87% y 75%, respectivamente. Es decir, el similar resultado mostrado por México incorpora el hecho de que en este país la mitad de la población de quince años estaba excluida de la educación secundaria, cuando en Chile y Argentina el 13% y el 25%, respectivamente,

⁷⁸ Recientemente, Vegas y Coffin (2015) han publicado un análisis similar al que sugiere este gráfico.

⁷⁹ Como se puede notar, el texto citado habla de «eficacia» y no de «eficiencia»; sin embargo, esto puede explicarse por problemas de traducción y no por diferencias en el sentido de la afirmación. De hecho, esta tendencia a usar en español el término «eficacia» en vez de «eficiencia» ha estado presente en diversas traducciones de documentos estadísticos de la Unesco desde hace mucho tiempo. Por ejemplo, en las especificaciones técnicas de los indicadores del UIS (Unesco-UIS 2009a) se habla de «eficacia interna» cuando se describe indicadores vinculados a la **eficiencia** interna, como la tasa de repetición por grado (Unesco-UIS 2009a: 14); y del coeficiente de «eficacia» (Unesco-UIS 2009a: 16). Nótese que estas especificaciones reproducen con algunos cambios las contenidas en Unesco (1999b), que es un texto producido y traducido hace más de una década. Probablemente estemos ante un caso en el que exista una interferencia del francés en la traducción al español, ya que en francés es usual hablar de *efficacité interne* y no de *efficience interne* en el caso de la educación (como lo hace la versión francesa del citado documento [Unesco-UIS 2009a], mientras la versión inglesa sí usa la palabra *efficiency*).

se encontraban en esa condición en el año 2000. Entonces, la mayor «eficiencia» de México podría perfectamente explicarse por sus niveles de exclusión. Puestas así las cosas, ¿los tres países tienen sistemas educativos de igual calidad? ¿Es México más eficiente? ¿Se puede ser de igual calidad con diferentes niveles de exclusión, es decir, la exclusión o la falta de equidad no son criterios que importen para definir la calidad? Nótese que las poblaciones estudiadas por PISA en los tres países son comparables, dado que se corresponden con la misma definición de la población objetivo, pero esta definición deja de lado el problema de las diferencias en los niveles de acceso y abre la puerta para que un análisis no muy feliz de las cosas termine afirmando un despropósito. Es también evidente que un análisis más cuidadoso podría verse informado por otras variables (como los niveles de acceso), pero si el discurso político que enmarca el estudio subraya que lo que importa es «la calidad y no la cantidad», entonces, el uso de información sobre acceso podría no ser tan evidente.

El ejemplo muestra que el reduccionismo mecánico implicado en las visiones de la calidad educativa basadas en una analogía insumo-producto (incluso en su versión mejorada de contexto-insumos-procesos-productos-resultados) lleva a situaciones absurdas, como postular que la calidad es independiente de la equidad, lo que está cándidamente implicado en exagerar el énfasis en la «calidad» por encima de la «cantidad», lo cual, como muestra este ejemplo, resulta algo completamente equívoco.

Así como se hace esta observación sobre estos países latinoamericanos, sería perfectamente posible comparar otros países en el gráfico 7. Por ejemplo, si se observa a los países con desempeños cercanos a la media de la OECD (500 puntos), se notará que hay algunos con niveles de inversión acumulada por estudiante poco mayores a veinte mil dólares PPP así como países con niveles de inversión cuatro veces mayores. ¿Son estos últimos más «ineficientes»? ¿No será que invierten en elementos que no tienen por qué traducirse en mayores puntajes en lectura y que hacen parte de un buen entorno escolar? ¿O será que, simplemente, este tipo de juicios se considera oportuno cuando se habla de países latinoamericanos y no en el caso de países de la OECD?⁸⁰

⁸⁰ El mencionado documento de Vegas y Coffin plantea también, sobresimplificando el tema, que un incremento de 14 puntos en la escala de matemáticas de PISA se logra con una inversión de mil dólares PPP adicionales por alumno. Entre 2001 y 2012, el Perú vio mejorar su rendimiento medio en PISA de entre 318 y 336 (intervalo de confianza al 95%) a entre 376 y 392 (intervalo de confianza al 95%); es decir, mejoró entre 40 y 74 puntos, cuando en dicho período el gasto público en educación secundaria por alumno se incrementó, en términos reales, en aproximadamente seiscientos dólares nominales. Es decir, si el cálculo anterior hecho para matemáticas se aplicase de la misma forma en lectura (lo que no es así, pero valga como ilustración), el Perú solo «debió» aumentar en 8,4 puntos en la escala de PISA, pero su incremento fue entre cinco y nueve veces mayor.

Hay que tener mucho cuidado con los mensajes que se envían a partir de un modelo estadístico: alguno podría concluir que no hay forma de mejorar hasta que, digamos, tripliquemos el gasto por alumno, cuando hay casos claros de educación de buena calidad a menores costos. Ese argumento puede velar la necesidad de no solo gastar más, sino también de gastar **mejor**.

Este ejemplo y la reflexión previa, sin embargo, solo nos sirven como punto de partida, ya que revelan la necesidad de romper la equivalencia de uso corriente entre «calidad de la educación» y «niveles de rendimiento académico (medidos mediante pruebas estandarizadas en un par de áreas cognitivas consideradas centrales) que se pueden considerar aceptables»⁸¹, pues esta equivalencia no brinda una noción de calidad de la educación que sea pertinente.

A efectos de avanzar en esta dirección es necesario considerar algunos puntos clave. En primer lugar, rastrear el uso actual de **la palabra calidad**. Su uso corriente se origina en el mundo administrativo y empresarial y no es uno que tenga un significado unívoco. De hecho, en 1994 la Organización Internacional de Estándares (ISO) definió la calidad como «la totalidad de los atributos y características de un producto o servicio que lo hacen capaz de satisfacer necesidades explícitas o implícitas» (ISO 8402, 1994)⁸²; y en 2005 revisó esta definición adoptando la siguiente: «grado en el cual un conjunto de características inherentes satisface los requerimientos» (ISO 9000, 2005)⁸³.

Estas definiciones ponen el acento en que la «calidad» no es un atributo o propiedad intrínseca del producto o servicio, sino que nace de la relación entre este y los «requerimientos» (de un cliente). El ajuste entre un producto o servicio y las expectativas que existen alrededor de él resulta ser el elemento clave de estas definiciones.

Siguiendo esta lógica, en el caso de la educación la calidad se vincularía a su **capacidad para satisfacer expectativas** que tienen que ver tanto con objetivos explícitos (por ejemplo, metas políticamente establecidas) como con concepciones sobre la educación. Así, para propósitos ilustrativos, una visión que prioriza los impactos económicos de la educación entendería que esta es de calidad si, por ejemplo, eleva la productividad⁸⁴.

⁸¹ Opiniones en este sentido se han vertido en más de una ocasión, como, por ejemplo, en los debates del II Foro Hemisférico de Evaluación de la Educación (Brasilia, junio de 2005), donde un número significativo de participantes manifestó insistentemente la necesidad de contar con una visión ampliada de la «calidad», es decir una visión no limitada al logro académico. Del mismo modo, las labores del Grupo de Trabajo Internacional sobre Medición de la Calidad de los Aprendizajes tomaron esta necesidad como su punto de partida y las mismas consideraciones se acogieron en las consultas y en el trabajo en general de la Learning Metrics Task Force (2013a, 2013c).

⁸² El texto original dice: «[...] the totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs».

⁸³ La ISO 9000 revisa y sustituye a la ISO 8402 que, por tanto, fue dada de baja (ver ISO 2009).

⁸⁴ Este ejemplo no es accidental, da cuenta de la lógica central de los esfuerzos de la OECD en el terreno de la medición de aprendizajes tanto de los jóvenes (PISA) como de los adultos (según lo planteado en los años 1990 por el International Adult Literacy Survey (IALS) y más recientemente en el Programme for International Assessment of Adult Competencies (Piaac). Por ejemplo, tanto IALS como Piaac definen a su población objetivo como aquellas personas entre 16 y 65 años de edad, lo que podría llevar a pensar que los mayores de 65 años no son considerados adultos. En realidad, de lo que se trata es de enfocarse en la población económicamente activa, por lo que las competencias de las personas mayores de 65 años resultan, desde este punto de vista, irrelevantes. Esta situación ilustra, una vez más, cómo las decisiones que definen el alcance y las características de los sistemas de información siempre trasuntan determinadas visiones acerca de lo que importa.

Al mismo tiempo, es importante clarificar el segundo término de la frase calidad de la **educación**. En efecto, ¿qué se ha de entender por «educación»? Esto es un terreno aun más complejo. Consideremos de modo esquemático los siguientes puntos:

- Existe un conjunto muy amplio de ideas sobre la educación que se originan en la reflexión que históricamente las sociedades han hecho sobre sí mismas. Estas ideas impactan en las visiones y creencias tanto explícitas como implícitas que cada persona tiene sobre la educación.
- Un primer grupo de ideas sobre la educación se vincula al pensamiento más clásico occidental que la ve como el desarrollo del potencial humano en tanto y en cuanto los seres humanos son concebidos como seres que **aprenden** tanto a ser individuos como a ser integrantes de una colectividad. Esta tradición de pensamiento destaca la distancia entre los humanos y otros seres vivos en los que muchos comportamientos básicos se encuentran «preprogramados» en la propia especie, mientras que los humanos «aprenden» o construyen su ser individual al tiempo que «aprenden» y reproducen los valores, creencias y demás elementos de su ser en comunidad (la cultura). En ese sentido, la relación entre educación y aprendizaje es la clave del concepto de educación.
- Otro grupo de ideas tiende más bien a vincular la educación a necesidades humanas y sociales identificables en términos de la vida social. Por ejemplo, se subraya el **rol** de la educación en la socialización de las personas (dado su papel en la internalización de normas sociales). Del mismo modo, se asigna a la educación otras «tareas» derivadas de necesidades sociales, como el desarrollo de la productividad y de pautas de vida ciudadana, el cuidado del medio ambiente, etc. En este caso, la educación tiende a ser vista como subsidiaria a otras esferas de la vida social y su «rol» suele ser definido en **función** de estas.

Nuestro discurso diario sobre la educación suele combinar, no siempre de manera muy consistente, elementos que se originan en los distintos planteamientos que acabamos de reseñar. Se necesita, por tanto, hacer un esfuerzo por dilucidar qué parte de estos elementos puede componer una visión consistente de las cosas.

Para ello, las visiones pedagógicas, como la teoría social contemporánea, brindan elementos de juicio que permiten desarrollar una visión amplia de la educación que, hasta cierto punto, reconcilie elementos que vienen de diversas vertientes, pero sin el contrasentido de ver a la educación como central en la vida humana al tiempo que es definida como subsidiaria a varios de sus elementos concebidos desde una perspectiva funcional o instrumental.

Un primer elemento clave en este terreno lo constituye la proclamación de la educación como un derecho humano fundamental. Al hacerlo, las sociedades contemporáneas

optan por valorar a la educación como un fin en sí mismo y no porque goza de utilidad para otros fines. Esta valoración obedece a que la educación está enraizada en la propia definición de la condición humana, ya que es entendida como requisito para el desarrollo pleno de nuestra humanidad, así como para garantizar otros derechos humanos y las libertades fundamentales. Si bien otros efectos de la educación pueden ser valorados (como, por ejemplo, su impacto en la productividad), estos ocupan un lugar secundario en el argumento: «La educación seguiría siendo igual de importante incluso si estos efectos no existiesen».

Una visión de la educación centrada en el derecho plantea diversas dimensiones imbricadas en su ejercicio y garantía. Así, el acceso a la educación aparece como una primera condición a garantizar, pero notando que no se trata de cualquier educación, sino de una que al ser relevante y pertinente pueda exitosamente favorecer el desarrollo del potencial individual y colectivo de las personas⁸⁵.

Entonces, el derecho de todas las personas a una educación de calidad radica en que gracias a ella nos desarrollamos como seres humanos y podemos ejercer otros derechos. Ahora bien, hacer efectivo el derecho a la educación pasa a su vez por asegurar otros derechos, tales como el derecho a la participación y a la no discriminación, a la propia identidad y a la igualdad de oportunidades.

Los primeros (participación y no discriminación) implican que ninguna persona debería sufrir ningún tipo de restricción para participar en las diferentes actividades de la vida humana. En el ámbito educativo, esto quiere decir que todas las personas deberían educarse, independientemente de su origen social y cultural y de sus características personales, y deberían vivir su experiencia educativa (ya que los elementos vivenciales son centrales al aprendizaje) en espacios comunes.

Sumado a lo anterior, lograr la individuación del sujeto en la sociedad implica asegurar el derecho a la propia identidad, respetando y valorando a cada uno como es y promoviendo su autonomía, autogobierno y la construcción de un proyecto de vida.

Adicionalmente, alcanzar una educación de calidad para todos, que promueva el máximo desarrollo, aprendizaje y participación de cada persona, solo será posible si se asegura el principio de igualdad de oportunidades, proporcionando a cada quien lo que necesita en función de sus características y necesidades individuales, para que esté en igualdad de condiciones de aprovechar las oportunidades educativas.

La reflexión anterior puede ser vista en consonancia con los desarrollos del pensamiento social contemporáneo. Este ha logrado formular enfoques de los asuntos humanos que permiten poner en el centro de la reflexión las acciones humanas (y no las «estructuras sociales»), lo que implica tomar en consideración aspectos básicos

⁸⁵ Estas ideas son desarrolladas *in extenso* en el capítulo segundo de Unesco Santiago (2007), así como en Pigozzi (2006).

relativos al carácter social de la vida humana. En efecto, la comprensión de los asuntos humanos requiere entender los sentidos o significados que las acciones sociales tienen y estos suponen un universo de relaciones humanas en las que son constituidos. Esta línea de pensamiento, iniciada por Husserl y la fenomenología, con su énfasis en la intersubjetividad y el *Lebenswelt*, ha devenido tanto en planteamientos relativos a la racionalidad dialógica propia del acento en la acción comunicativa (Habermas 1984) como en visiones de lo social en las que la acción humana es el momento en el que lo social se torna concreto y, por tanto, se reproduce; por ejemplo: las nociones de *habitus* de Bourdieu (2000), dualidad de la estructura de Giddens (1984) o *agency* en el realismo crítico, en particular en Archer (2000).

Estas reflexiones pedagógicas y de la teoría social sustentan la necesidad de una mirada de la educación caracterizándola desde la centralidad de la acción humana (*agency*) que define al fenómeno educativo al tiempo que opera en condiciones que la educación contribuye a crear. Asimismo, impulsa el pensamiento educativo en el sentido de subrayar capacidades humanas centrales vinculadas a la reflexión (y no al cálculo instrumental) y a la construcción colectiva de sentidos. En este marco, la educación se torna un elemento central de la vida humana en tanto es el espacio propio de la construcción de valores y sentidos que tocan a la convivencia y no puede ser reducida a un mero instrumento para satisfacer fines (que se asumen valiosos dado un principio de **utilidad**), cuya definición parece escapar a la acción humana y pertenecer a estructuras cuya lógica se asume inmanente y, por tanto, ahistórica.

En este marco de ideas, **la educación es un espacio clave de humanización** de las personas, tanto con relación al desarrollo de su potencial como individuos y como seres en comunidad como en cuanto a la construcción de una cultura que efectivamente concrete valores socialmente sancionados.

El desarrollo del individuo y de la especie humana es el principal sentido de la educación y, en tanto mecanismo al servicio de la humanización, tiene preeminencia sobre otras consideraciones políticas, económicas, sociales o culturales. Su potencial para contribuir a que las personas alcancen sus propios objetivos (económicos, sociales y culturales) y al logro de una sociedad mejor es lo que sustenta el valor primordial de la educación.

Sin embargo, los contenidos de su tarea humanizadora no son ahistóricos: cada colectividad y tiempo los expresa como ciertos mandatos o finalidades (algunos bastante instrumentales) para la educación. Estos varían en el tiempo y de un contexto a otro y dependen de qué tipo de posicionamientos ideológicos logran hegemonía cultural en un momento dado. Así, por ejemplo, actualmente se espera que la educación juegue un papel relevante en la concreción de los principios de lo que a inicios del siglo XXI se considera una **vida civilizada**:

- (i) educar en y para los derechos humanos y para el desarrollo de las personas como ciudadanos (sujetos libres y consiguientemente responsables);
- (ii) contribuir al desarrollo de sociedades más equitativas, justas y democráticas, reduciendo las desigualdades, promoviendo la resolución pacífica de los conflictos y mejorando las condiciones de vida de las personas y de las sociedades; y
- (iii) fortalecer la identidad cultural en un marco de diálogo con otras culturas y la valoración de la diversidad⁸⁶.

En adición a lo anterior, más allá de las diferencias en cuanto a enfoques, existe coincidencia en considerar tanto que el desarrollo cognitivo es un importante objetivo explícito de todos los sistemas educativos como el reconocimiento del papel de la educación en el estímulo del desarrollo creativo y emocional de las personas, la contribución a los objetivos de paz, ciudadanía y seguridad, la promoción de la igualdad y la transmisión de los valores culturales, tanto universales como locales a las generaciones futuras⁸⁷.

Por otro lado, también tiende a remarcarse la coincidencia de tres principios que pueden enumerarse de la siguiente manera: necesidad de pertinencia (ajuste a las características particulares de cada uno); necesidad de equidad en el acceso y los resultados; y necesidad de asegurar los derechos inalienables de las personas. Tales principios informan y orientan los contenidos y procesos educativos aceptados socialmente como válidos, además de representar objetivos sociales más amplios que se busca atender a través de los sistemas educativos⁸⁸.

⁸⁶ Estas consideraciones se encuentran desarrolladas en Blanco (2005).

Es claro también que lo aquí descrito son postulados que hoy se consideran de validez universal pero que tienen una raíz europeo-occidental muy fuerte –incluso cuando es posible encontrar en culturas no europeas ciertas tradiciones dialógicas muy significativas, como las identificadas en Sen (2005)–. Su valor universal está dado por el nivel de hegemonía cultural logrado por Occidente, el mismo que se hizo patente en la aprobación universal de la *Declaración universal de los derechos humanos* en 1948, no accidentalmente al término de una conflagración mundial en la que los Estados que encarnaban las promesas emancipadoras del Occidente moderno (en sus distintas vertientes) lideraban la coalición vencedora. Sin embargo, cabe también anotar que en la actualidad (seis décadas después de dicha declaración) el discurso hegemónico en Occidente es profundamente instrumental, por ejemplo cuando se refiere a «la educación para elevar la productividad y ser más competitivos», es decir, la educación es considerada como una forma de construir «capital humano» (Becker 1993).

⁸⁷ Un ejemplo emblemático de estas preocupaciones se encuentra en la génesis de lo que hoy se conoce como Bachillerato Internacional (el Colegio Internacional de Ginebra). Véase Maurette (1948).

⁸⁸ El mismo debate en América Latina y el Caribe ha producido un sinnúmero de desarrollos y propuestas acerca de la necesidad de desarrollar una visión integral y amplia de la calidad de la educación. Si bien las perspectivas pueden variar, la recurrencia de algunos componentes claros a la hora de considerar sus ejes y dimensiones induce a un camino convergente referido a lo que se considera esencial para abordar el tema de la calidad educativa. Así, los atributos que reafirman una educación de calidad se asocian principalmente a la relevancia de los objetivos y contenidos, la eficacia para alcanzar los objetivos, la pertinencia de los contenidos y métodos curriculares, la equidad en cuanto a las oportunidades y la eficiencia en la utilización de los recursos. Para una mayor discusión, ver, entre otros: Muñoz Izquierdo *et al.* (2004); Schmelkes (1996); Toranzos (1996). A nivel global puede consultarse: Unicef (2000).

Estas discusiones generales sobre la **calidad de la educación** impactan o contribuyen a constituir la manera como las sociedades y las personas la entendemos y, por lo mismo, terminan expresándose también en documentos doctrinarios o de política que los países de modo singular o en grupo sancionan.

Un ejemplo de interés es el Proyecto Regional de Educación para América Latina y el Caribe (Prelac) adoptado por los ministros de Educación en 2002 (Unesco Santiago 2002b)⁸⁹. A pesar de su nombre, el Prelac no fue concebido como un proyecto, sino más bien como un marco orientador o doctrinario que sienta bases conceptuales para enfrentar los desafíos existentes y lograr las metas de la EPT. En ese sentido, dicho marco acoge un conjunto de reflexiones que tocan directamente a lo aquí discutido.

Es posible destacar dos elementos claves del Prelac: (i) su preocupación por los sentidos y contenidos de la educación, es decir, por el significado; y (ii) su entendimiento de las personas como agentes que constituyen (producen y reproducen) el mundo de lo social (en este caso centrado en el tema educativo), tanto en su acción individual libre y responsable como en la configuración de colectivos locales (que se expresan, por ejemplo, en una cultura escolar) o en otros más amplios (traducidos en, por ejemplo, los niveles de responsabilidad social por la educación).

Lo hasta aquí señalado deriva en una definición amplia de la **calidad de la educación** coherente con estas perspectivas contemporáneas. Estas subrayan que una educación de calidad es aquella capaz de **dotar de sentidos y asegurar el desarrollo de potenciales humanos básicos en las condiciones particulares en las que nos toca vivir, promoviendo la educación como derecho humano y el desarrollo pleno mediante el aprendizaje en los planos cognitivo, social, ético, emocional y creativo de todos y cada uno.**

Esta visión, por su parte, tiene correlatos sobre elementos centrales de la práctica educativa. Así, la construcción de sentidos y el desarrollo de capacidades supone un entronque con las condiciones particulares de las personas (pertinencia) y con los desafíos de desarrollo que deben afrontar (relevancia); tiene, a su vez, una relación no incidental sino sustantiva con la forma como las sociedades operacionalizan algunos objetivos en la política educativa (eficacia) y consagran principios de operación propios de la sociedad contemporánea (que incluyen la eficiencia en el uso de recursos públicos). Asimismo, se vincula a una noción contemporánea del universo de derechos civiles, políticos y sociales que constituyen la vida de los ciudadanos (equidad).

Resulta imprescindible notar que no todos estos elementos son igualmente observables, ni tienen concreciones equivalentes en el mundo contemporáneo. Esto se vincula al hecho de que las sociedades contemporáneas organizan las prácticas educativas en **sistemas educativos** (Archer 2013) y estos son el objeto prioritario aunque no exclusivo de las reflexiones y de la acción pública en el terreno de la educación. Los sistemas educa-

⁸⁹ Para su mayor expresión conceptual, véase el capítulo segundo de Unesco Santiago (2007).

tivos son el dispositivo fundamental de la acción educativa en el terreno público. Es a través de este tejido institucional que los países desarrollan sus principales esfuerzos para la consecución de los objetivos colectivos en educación⁹⁰. Por lo tanto, es preciso señalar que las reflexiones hasta aquí planteadas no se refieren a la **calidad de la educación** en general, sino a la **calidad de los sistemas educativos**⁹¹.

2.3.2 DIMENSIONES DE LA CALIDAD DE LOS SISTEMAS EDUCATIVOS

A efectos de traducir lo hasta aquí señalado en pautas que informen el desarrollo de los sistemas de información, es preciso operacionalizar esta discusión definiendo los ámbitos de observación que los sistemas de información pueden cubrir, de modo que, a partir de tal definición, sea posible proceder a desarrollar los mecanismos, instrumentos y dispositivos de operación de dichos sistemas.

La discusión anterior ha identificado cinco componentes o dimensiones fundamentales de la calidad de los sistemas educativos, a saber: (i) relevancia; (ii) pertinencia; (iii) equidad; (iv) eficacia; y (v) eficiencia. Estas dimensiones no son todas de la misma naturaleza, como se señala en Unesco Santiago (2007). Mientras que las tres primeras tocan a la naturaleza de la educación como derecho humano fundamental, las dos últimas introducen un nivel de análisis distinto, que se corresponde con la naturaleza de la acción pública en educación. Veamos con cierto detalle cada una de estas dimensiones.

- (i) La **relevancia** del servicio que brindan los sistemas educativos toca a los sentidos básicos de la educación y, por lo mismo, es también un atributo general de las prácticas educativas.

Dado el foco en la observación de los sistemas educativos, la relevancia se vincula a los contenidos y formas en los que se estructuran los servicios educa-

⁹⁰ Así, si bien la educación abarca también acciones humanas que no están comprendidas por el sistema educativo (ya sea en el entorno familiar, laboral, público o privado), la reflexión sobre un sistema público de información educativa se ocupa primariamente de la educación estructurada en un sistema educativo. Esto no debe confundirse con tener un foco exclusivo en las personas atendidas por el sistema; todo lo contrario, un criterio fundamental de observación y juicio sobre el desempeño de los sistemas educativos se vincula directamente a su capacidad para garantizar el derecho a la educación de **todas** las personas, por lo que la inclusión es una preocupación central, lo que supone la capacidad para atender las características específicas de cada individuo. Tampoco debe confundirse con un foco exclusivo en las modalidades centrales que ocupan a los sistemas educativos (como la educación de niños en escuelas), ya que el sistema educativo abarca diversas modalidades de prestación de los servicios (para jóvenes y adultos, no convencionales, etc.).

⁹¹ La **calidad de la educación** podría perfectamente referirse a diversas dimensiones o niveles en los que se da el fenómeno educativo, sea a nivel de escuela, de la gestión educativa o, por ejemplo, del impacto de los medios de comunicación en la educación de las personas. El hecho de que la educación sea un fenómeno que toca de modo tan central los asuntos humanos hace que tenga manifestaciones en ámbitos de la vida de muy diversa naturaleza y alcance.

tivos que dichos sistemas brindan. En este sentido, el sistema educativo será de calidad en tanto brinde un servicio cuyos contenidos sean relevantes. Tal relevancia es entendida como el ajuste de la oferta de servicios por parte de los sistemas educativos a las necesidades del mundo actual, tanto en términos de las competencias para la vida social como del desarrollo individual, a efectos de que la persona actúe en el mundo como un sujeto libre y responsable.

En este sentido, los cuatro pilares del aprendizaje del siglo XXI definidos por la Unesco (1996) acerca de **los contenidos y prácticas de la educación para construir sentidos acerca de nosotros mismos, los demás y el mundo en el que vivimos**⁹² se convierten en el referente normativo que puede orientar la observación y la medición.

La observación y el análisis de los *curricula* nacionales y la normativa vigente se convierten en un paso obligatorio para lograr una descripción del estado de la situación de los sistemas educativos en este respecto; sin embargo, esto solo es un punto de partida. La medición directa de lo que las personas logran en este terreno mediante la operación de los sistemas educativos (incluyendo elementos cognitivos y actitudinales) es un elemento de información de inmenso valor⁹³.

- (ii) La **pertinencia** del servicio que brindan los sistemas educativos es otro punto de partida central para dar cuenta de la calidad de los mismos. En tanto y en cuanto el sistema sea capaz de establecer un nexo fluido entre lo socialmente definido y las condiciones particulares de las personas, es posible contar con un sistema de calidad. Así, la pertinencia es definida como la capacidad de ajuste de la oferta de servicios de los sistemas educativos a las condiciones particulares de las personas y colectividades en sus correspondientes contextos culturalmente constituidos, de modo que se asegure un apropiado enlace entre la oferta educativa y su cultura, necesidades, concepciones y prácticas.

En este sentido se requiere, entre otras cosas, una evaluación de juicio de la capacidad de adaptación o flexibilidad de los *curricula* nacionales en conjunto con la gestión pedagógica existente.

- (iii) La **equidad** tiene una relación estrecha tanto con nuestras nociones contemporáneas de lo que es objeto de derechos y de la acción pública (orientada a la garantía de derechos) como con los atributos de dicha acción, por lo que se vincula de modo directo a los sistemas educativos en tanto mecanismos de la misma. Así, la equidad es definida con relación a la capacidad del sistema

⁹² El desarrollo de la identidad (ser) como sujeto que hace parte de un mundo social (vivir juntos), de la capacidad de aprender (conocer) y de la capacidad de actuar (hacer).

⁹³ Sin embargo, esto requiere operativos de una gran complejidad (como los vinculados a la medición de los logros cognitivos y las actitudes) y la identificación de la contribución de las trayectorias escolares a los logros (o no logros) que se observen.

educativo de asegurar universalmente el principio de igualdad de oportunidades con relación al derecho a la educación (por ejemplo, niveles de acceso y resultados), independientemente de las condiciones preexistentes (sociales, culturales o de otra índole).

La inclusión de variables relativas al perfil (socioeconómico, cultural, etc.) de las personas hace posible la realización de un análisis amplio que contribuya a la descripción y diagnóstico de esta dimensión.

- (iv) La **eficacia** puede ser observada de un modo directo por su relación consustancial con los objetivos del sistema educativo. En efecto, tales sistemas en tanto objeto de la acción pública se plantean de modo explícito objetivos de política cuyo logro puede ser verificado. Así, la eficacia se refiere a la capacidad de los sistemas educativos para lograr sus objetivos, cumplir sus obligaciones y desempeñar sus funciones. En este sentido, la imbricación entre resultados y acciones humanas que en ellos derivan lleva a la necesidad de evitar privilegiar una mirada exclusiva a los «resultados» y considerar aspectos vinculados a los sujetos, las prácticas y la gestión curricular que hacen parte de la interacción.
- (v) La **eficiencia**, por su parte, compromete un atributo central de la acción pública: que se ejecute honrando los recursos que la sociedad destina a tal fin, por lo que la obligación de ser eficiente toca a la garantía de un derecho ciudadano clave. Por lo tanto, la eficacia es definida con relación al financiamiento destinado a la educación, a la responsabilidad en el uso de este y a los modelos de gestión de recursos e institucional.

Los países de América Latina y el Caribe han acumulado una extensa experiencia en el análisis de esta dimensión dado que algunas de las variables asociadas con ella describen situaciones particularmente comunes en Latinoamérica. Repetición, deserción, años transcurridos en el sistema educativo para lograr egresar de un nivel educativo dado y gasto en educación, entre otros, constituyen contenidos tradicionales en los análisis realizados en la región, aunque algunos de ellos deben ser revisados con detenimiento (como se hará más adelante). La información referida a los modelos de gestión también merece un esfuerzo adicional a la hora de incluirla como parte de esta dimensión.

Junto a estas dimensiones, es necesario considerar tanto la **dinámica** propia de los sistemas educativos como los desafíos a los cuales se enfrentan dados los cambios permanentes en el mundo del que somos parte, los mismos que derivan en nuevas y cada vez más complejas demandas. Esto introduce una perspectiva de reflexión vinculada a la inclusión de información que surge de la observación del desempeño de los sistemas educativos en el tiempo. Dados los atributos cambiantes de la sociedad contemporánea, la observación de los **esfuerzos** y **progresos** de los sistemas educativos no es accidental,

ya que se vincula tanto a la necesidad de avanzar (en términos de las dimensiones anteriores) como a la de prestar atención de modo continuo al surgimiento de nuevos requerimientos y desafíos.

En este sentido, resulta necesario que la descripción y el análisis consideren el progreso, la estabilidad o el retroceso del sistema y los esfuerzos emprendidos para la mejora de los servicios educativos en su conjunto.

Finalmente, es importante mencionar la razón por la que hasta este punto no se incluyen los aspectos que se vinculan a los **impactos** de corto, mediano y largo plazo que la educación tiene sobre el conjunto más amplio de la vida de las personas y las colectividades, como es el caso de las condiciones de salud, empleo, ingresos, conciencia ciudadana, etc. Estos aspectos no han sido incluidos ya que la dinámica temporal de estos impactos no solo es compleja y de largo aliento sino que tampoco es equivalente a las de los componentes aquí identificados, que se centran en dar cuenta de la situación **actual** de los sistemas educativos. Por lo mismo, si bien estos análisis son de singular importancia –véase Muñoz Izquierdo *et al.* (2004)–, su introducción como una dimensión adicional no se corresponde con el foco indicado. Esto no debe entenderse como desdén, sino como el reconocimiento de un distinto énfasis político (los desempeños presentes) y de la necesidad de abordar estudios sobre los impactos de largo plazo de la educación desde una perspectiva informada por la investigación a la que los sistemas de información sustentan⁹⁴.

La complejidad de las interrelaciones presentes entre los diferentes actores del sistema y los procesos educativos que se dan tanto en el interior como externamente a dicho sistema conlleva una dificultad adicional: las fronteras y asociaciones entre los diversos elementos identificados previamente tienden a ser difusas o, dicho de otra manera, no resulta posible encontrar fronteras nítidas y definitivas entre cada dimensión y sus elementos.

El modelo descrito no es, ni pretende ser, una fotografía o reproducción que refleja de la realidad ya que, para empezar, esto es epistemológicamente un sinsentido (el conocimiento es una reconstrucción del mundo, no un «reflejo» de este, como postulaba cierta pseudo-filosofía de inicios del siglo XX); tal modelo es, más bien, un mecanismo de apoyo o un soporte analítico y heurístico. Por lo mismo, si bien es posible identificar las relaciones características entre las dimensiones propuestas y sus componentes citados, esta identificación tiene el propósito de organizar la generación, análisis y uso de la información, no pretendiendo ser una descripción rígida de realidades que no lo son.

⁹⁴ De hecho, la información generada por los sistemas de información es la que hace posible estos análisis. Asimismo, cabe anotar que en discusiones preliminares de estas ideas (durante reuniones de viceministros y talleres regionales de estadísticas educativas en América Latina y el Caribe), la legítima preocupación por los impactos de mediano y largo plazo fue expresada en más de una ocasión. Nótese también que el esquema más tradicional desarrollado por la Unesco, y cuya representación gráfica fuera reproducida en el acápite previo de esta sección, sí toma en cuenta esta dimensión (ver el gráfico 6, en la página 104).

Así, tanto los aspectos vinculados al aseguramiento de derechos como el ajuste de los servicios educativos a los cuatro pilares del aprendizaje del siglo XXI y también los sentidos que se construyen mediante la educación tienen una relación más próxima con la dimensión vinculada a la **relevancia**.

De la misma manera, la flexibilidad de los sistemas educativos, las prácticas inclusivas y la atención a la diversidad se vinculan de manera más estrecha a la **pertinencia** de los servicios brindados por el sistema educativo.

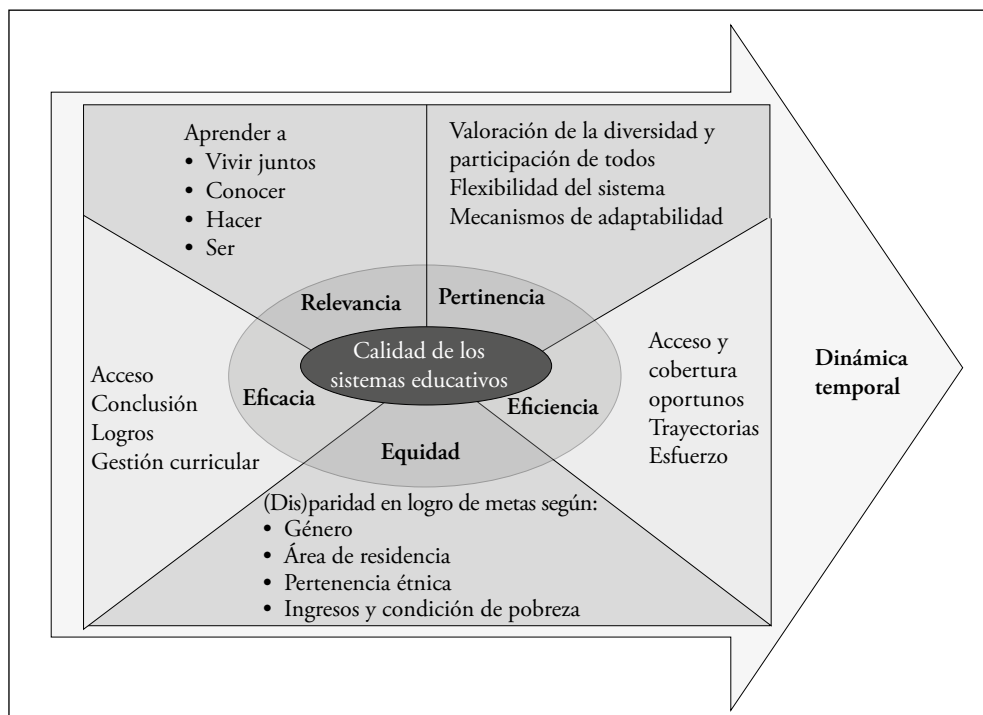
La dimensión **equidad**, por su parte, puede ser objeto de análisis prestando particular atención a la disponibilidad de recursos (y su distribución), así como al aseguramiento de la igualdad de oportunidades, con relación a los niveles de aprendizaje y al acceso tanto a los servicios educativos como a las oportunidades que la experiencia educativa crea.

Por su parte, las variables con vínculos más estrechos con la **eficacia** son aquellas referidas a la observación de la consecución de los objetivos propuestos y de los entornos de los procesos de enseñanza-aprendizaje: entornos donde estos procesos existen, así como los actores del mismo (docentes y alumnos), lo mismo que las interacciones entre ellos y con sus pares, lo que puede ser definido como gestión curricular.

La dimensión **eficiencia**, finalmente, puede ser observada con relación a los temas referidos a financiamiento y uso de recursos, gestión y responsabilidad social.

El gráfico 8 trata de representar de modo sintético y esquemático los elementos aquí expuestos, los cuales, como se anotó al inicio del acápite, corresponden al esfuerzo desarrollado por la Orealc-Unesco Santiago entre 2003 y 2007 (cuyos principales resultados se encuentran expuestos en Unesco Santiago 2007, 2008).

Gráfico 8. Visión de la calidad de los sistemas educativos desarrollada por la Orealc-Unesco Santiago



Nota

Elaboración a partir de Unesco Santiago (2007, 2008) y de borradores no publicados de documentos de discusión mencionados al inicio de este acápite.

2.3.3 LA CORRESPONDENCIA ENTRE EL ENFOQUE PROPUESTO Y LOS MODELOS DE ANÁLISIS MÁS EXTENDIDOS SOBRE LA MEDICIÓN DE LA CALIDAD DE LA EDUCACIÓN

Para empezar, la llamada **medición de la calidad de la educación** es, en realidad y fundamentalmente, la medición de ciertos resultados de aprendizaje en términos cognitivos y mediante pruebas estandarizadas.

La discusión anterior debería ser suficiente para establecer la siguiente inecuación:

$$QEd \neq LO \neq CSk_s \neq Score_{st}$$

Donde:

QEd: calidad de la educación

LO: logros de aprendizaje

CSk_s: habilidades cognitivas en la materia *s*

Score_{st}: puntaje en pruebas estandarizadas.

Esta inequación busca hacer explícitos los límites más evidentes de esta identidad y se refiere a que, si bien nadie realmente postularía en serio que existe tal identidad, sí existe el postulado que afirma que el último elemento de la misma es una representación lo suficientemente buena del primero.

Esta afirmación se basa en un principio básico de **parsimonia** que, efectivamente, busca cierta economía y realismo en la empresa del conocimiento: al no haber ninguna posibilidad (al menos humana) de un conocimiento omnisciente, entonces es necesario buscar aproximaciones que sean suficientes⁹⁵.

Justamente, las páginas previas muestran que este no es el caso, sino más bien que la inequación postulada invoca un reduccionismo que puede llevar a desaciertos interpretativos garrafales (como el previamente identificado a propósito de la relación entre los puntajes en PISA y el gasto acumulado por estudiante).

Siendo la medición de logros estudiantiles a través de pruebas estandarizadas un elemento de información de primera importancia (más adelante discutiremos este punto), es claro que usarla como sinónimo de medición de la calidad de la educación es extremadamente reduccionista y, además, completamente innecesario y riesgoso, dado el punto de vista del tipo de racionalidad que dicho razonamiento implica.

En segundo lugar, los modelos analíticos de uso corriente tanto en los sistemas de medición de logro como aquellos asociados a sistemas de indicadores educativos en general no contienen una visión amplia de calidad como la aquí descrita.

En tercer lugar, la reflexión aquí reseñada sobre la complejidad y amplitud del concepto de calidad de la educación, y de los sistemas educativos en particular, tiene consecuencias sobre los elementos centrales de lo que se conoce como modelo CIPP (contexto-insumos-procesos-producto). Como se desprende de la discusión anterior, el modelo CIPP, tomado de la administración y la gestión industriales, tiene una lógica interna propia de la racionalidad instrumental en la que lo central es la P final, el producto, y todo lo demás es visto como condición de este. Así, la calidad de un sistema

⁹⁵ La idea de parsimonia ha estado permanentemente presente en el desarrollo de la ciencia, aunque se asocia de modo particular a la lógica del fraile franciscano inglés William of Ockham, quien en el siglo XIV planteara: «Frustra fit per plura quod potest fieri per pauciora», es decir, que resulta fútil hacer con más lo que se puede hacer con menos (1340, libro I, cap. 12). Esta es una de las formulaciones del principio que ha venido en llamarse la «navaja de Occam» (usando la forma castellana del apellido). En la misma línea, pero enfatizando otro aspecto, Antoine de Saint-Exupéry señaló en *Terre des Hommes* que «la perfección se logra no cuando ya no resta nada que añadir, sino cuando ya no hay nada más que eliminar» (1939: 57; traducción propia). De un modo análogo, Einstein señaló en 1933 –en la Conferencia Herbert Spencer en Oxford– que «difícilmente se puede negar que el objetivo supremo de toda teoría es establecer sus elementos básicos irreductibles, tan simples y limitados en número como sea posible, sin que esto signifique limitar la adecuada representación de un dato empírico singular» (1934: 165). En un interesante ejercicio de parsimonia aplicada, esta frase es usualmente citada como: «Todo debe hacerse de la manera más simple posible, pero no más simple que ello». En cualquier caso, la frase de Einstein advierte respecto de los riesgos de exagerar la aplicación de la navaja de Occam, es decir, de sobresimplificar.

(cualquiera que este sea) en un modelo CIPP es juzgada por los atributos del producto final y la manera como los insumos y procesos intermedios contribuyen o no a dicho producto, dadas determinadas condiciones de contexto.

Desde este punto de vista, el modelo es lógicamente consistente y, por lo tanto, resulta coherente y ha sido posible tomarlo como eje orientador de empresas exitosas de generación y análisis de información. De hecho, sus limitaciones no se encuentran en ese plano, sino en su no correspondencia con una forma de enfocar los asuntos humanos que es radicalmente distinta de la sugerida por el debate educativo contemporáneo, al que se ha querido responder en la discusión precedente⁹⁶.

El foco puesto en las prácticas y los significados es sencillamente **diferente** al del modelo CIPP; aquel responde a una concepción de ser humano y del conocimiento que es, a su vez, diferente de la implícita en la racionalidad instrumental y en los modelos funcionales y/o sistémicos asociados a esta. Entonces, es necesario preguntarse si la no correspondencia epistemológica e incluso axiológica entre lo que los sistemas educativos crecientemente se plantean como ejes orientadores de su accionar y los modelos de información y evaluación basados en el modelo CIPP acarrea consecuencias sobre la pertinencia de usar dichos modelos. Mediante una *reductio ad absurdum*, la respuesta más directa es que sí, esta correspondencia sí tiene consecuencias, pues si ese no fuera el caso tendríamos la situación absurda según la cual la reflexión epistemológica y axiológica no impactaría de ninguna forma sobre la generación de conocimientos.

Los marcos conceptuales usados para la construcción de sistemas de información educativa han priorizado una visión «mecanicista» de la educación, basada en una analogía productiva. Estas aproximaciones, como cualquier otra en la formulación de conocimientos, no carecen de implicancias epistemológicas complejas que resultan difíciles de sostener cuando se trata de entender los sentidos y la naturaleza de las acciones humanas que configuran el fenómeno educativo.

Desde ese punto de vista, referirse a la matrícula final (matrícula inicial más saldo neto de transferencias a lo largo del año) como «existencias» (*stocks*), no solo resulta una curiosidad semántica sino denota una visión que desdeña a los agentes del sistema en tanto tales a favor de priorizar una idea «sistémica» racionalista y fundamentalmente simplista. Lo mismo se ilustra con el tratamiento de los docentes (agentes centrales en el proceso educativo) como «insumos».

Del mismo modo, conceptos como el de **demanda educativa**, usados en realidad para referirse a la población que teóricamente es objetivo de un programa educativo dado, resultan innecesarios o potencialmente equívocos.

⁹⁶ Una interesante crítica de la analogía de producción en educación puede encontrarse en Hodas (1993); véase también Robeyns (2006).

Detengámonos en este punto a efectos de presentar una ilustración que sirva de epílogo a esta discusión: ¿Cuál es la demanda educativa de, digamos, el cuarto grado de primaria? Las respuestas a una pregunta tan simple como esta pueden tener formas tan disímiles como:

- (i) toda la población (de cualquier edad) para la cual el tercer grado es su último grado aprobado;
- (ii) la anterior población menos aquellos que tienen una edad superior a la considerada a efectos de definir la obligatoriedad (si esta es definida para un grupo de edad);
- (iii) el primer grupo menos aquellos que no solicitan inscribirse en un programa educativo del grado siguiente (es decir, los que efectivamente **demandan** el servicio de cuarto grado);
- (iv) el grupo anterior más aquellos que podrían solicitar la matrícula en el grado siguiente de contar con una oferta de dicho servicio en su comunidad;
- (v) la población de la edad prevista normativamente para cursar ese grado;
- (vi) la anterior menos aquellos que teniendo la edad no cumplen los requisitos necesarios (haber aprobado el grado previo);
- (vii) los que han culminado exitosamente (es decir, aprobaron) el grado tercero el año lectivo precedente.

Se podría entrar en una discusión interminable sobre estas posibilidades y, sin embargo, no encontrar ninguna solución satisfactoria, máxime cuando este concepto en realidad solo es usado para definir una **demanda teórica** (equivalente a la opción [v] antes indicada). Si ese es el caso, ¿por qué llamarla **demanda** y no simplemente usar el nombre que la describe de forma más directa y con menos ambigüedad: **población de la edad que normativamente corresponde al cuarto grado**?

2.4 CONSTRUCCIÓN DE INDICADORES SOBRE ALGUNOS TEMAS CENTRALES DEL DEBATE EDUCATIVO

El presente acápite busca discutir la construcción de indicadores sobre algunos temas centrales presentes en el debate educativo. La relación no pretende ser exhaustiva, sino ilustrativa de algunos puntos de interés, y busca también discutir algunas nociones comunes que conducen a serias dificultades en términos de la capacidad de los sistemas de información educativa actuales para informar de modo efectivo en los debates públicos y los procesos de toma de decisiones.

2.4.1 COBERTURA Y PERFILES DE MATRÍCULA

Como ya se anotó, el diálogo técnico sobre el desarrollo de indicadores educativos no es un ejercicio puramente académico, sino una respuesta a las necesidades públicas de información principalmente vinculadas a las políticas.

Cobertura

En el caso de la cobertura, podemos organizar la reflexión considerando que el objetivo político ideal es que el 100% de los niños en edad de ingresar al sistema educativo lo hagan y permanezcan en él hasta culminar los niveles que se considere pertinentes (sean o no obligatorios) y, preferentemente, sin atrasarse en sus estudios, ya que el atraso escolar (sea por repetición o por ingreso tardío) conspira contra la probabilidad de culminación de los estudios⁹⁷. Así, un sistema educativo «perfecto» (con relación a este único aspecto) es uno en el que el 100% de los niños en edad de ingresar lo hacen y, a partir de ahí, progresan a lo largo de todo el ciclo educativo sin atrasarse ni desertar hasta que concluyen oportunamente sus estudios.

Dicho lo anterior, ¿a qué se orientan las preguntas sobre cobertura? El primer elemento a considerar tiene que ver con qué se entiende por cobertura. Una definición de diccionario afirma que es la «cantidad o porcentaje abarcado por una cosa o una actividad» o

⁹⁷ En efecto, el atraso escolar produce dos fenómenos concurrentes que conspiran contra la posibilidad de culminación de estudios: (i) crea un clima de desajuste entre el estudiante y su entorno en el que el sentimiento de fracaso es un factor preponderante; y (ii) incrementa el costo directo y el costo de oportunidad de estudiar, dado que se requieren más años (por repetición) o se alcanza, sin haber culminado los estudios, una edad en la que hay mayores presiones o demandas por generar ingresos o asumir otras responsabilidades, especialmente en los sectores de la población que se encuentran en situación de desventaja por motivos socioeconómicos. Así, el atraso escolar es un fenómeno que requiere ser visto con atención y detalle si se quiere asegurar la culminación de los estudios de modo equitativo en un país. Unesco Santiago (2008: 122 y ss.) mostró esta asociación negativa entre el atraso (y la ineficiencia expresada en el atraso) y la conclusión.

la «extensión territorial que abarcan diversos servicios, especialmente los de telecomunicaciones»⁹⁸. Estas acepciones de diccionario destacan el elemento definitorio que tiene que ver con el **alcance de los servicios**, con aquello que estos abarcan, sea en cantidades absolutas, relativas (porcentajes) o en otras propiedades como el espacio.

Una traducción directa de estas acepciones al terreno de la educación sugiere una afirmación como la siguiente: «La cobertura educativa es el grado en el que el sistema educativo atiende a la población».

Así, una primera forma de medir la cobertura es simplemente considerar la cantidad de personas que el sistema educativo atiende. Por ejemplo, en el País A, la cantidad de personas matriculadas en al menos un programa educativo en el año 2012 es de 3.109.765 personas. El sistema educativo, por tanto, sirve o «cubre» a poco más de tres millones de personas. Llamemos a esta primera aproximación al tema **cobertura global**.

Aunque parezca simple, este número por sí mismo es importante, ya que brinda a las personas una primera idea de la magnitud del sistema educativo. Sin embargo, no nos dice mucho más que eso. Por lo mismo, vale la pena expresar ese mismo número como una fracción de la población total del país a efectos de tener una aproximación un poco más fina. Así, la **cobertura global** puede expresarse tanto en términos absolutos como relativos. Si el País A tenía en 2012 una población total de 10.545.238 habitantes, entonces la **cobertura global** expresada en términos relativos sería igual a 0,257 de la población total. Esta magnitud puede, a su vez, expresarse en formas más familiares como 25,7% o 2.570 por cada 10.000 habitantes. En términos formales tendríamos:

Donde:

CG : cobertura global

Mat_t : personas matriculadas en al menos un programa educativo en el período t

Pob_t : población total en el período t

$$CG = \frac{Mat_t}{Pob_t} * x$$

x : factor que puede ser igual a 1, 100 o 10.000 para usar los ejemplos dados (fracción, porcentaje, cada 10.000 habitantes).

Se puede calcular para cada jurisdicción para la que se cuente con información de matrícula y población, en la medida en que estas sean consistentes (que el lugar de matrícula pueda ser asumido como lugar de residencia).

Nótese que en la definición hablamos de la «cantidad de personas que el sistema educativo atiende» y en la fórmula esto se expresa como el número de personas que se encuentran matriculadas en, al menos, un programa educativo. Sin embargo, es

⁹⁸ Definiciones tomadas de la vigésima segunda edición del *Diccionario de la Real Academia Española* (RAE 2001).

posible que no se cuente con esta información sino con la matrícula total de todos los programas educativos, que no es exactamente lo mismo, ya que una misma persona puede estar matriculada en más de un programa educativo, por lo que eso llevaría a una posible sobrestimación del valor calculado⁹⁹. Es responsabilidad de los sistemas de información determinar las mejores formas de obtener los valores que se corresponden con las definiciones (la matrícula no es igual al número de personas, pero puede ser una buena aproximación si se cuenta con un mecanismo para depurar posibles dobles matrículas) y estimar y documentar los posibles sesgos. Como siempre, se necesita un procedimiento que sea lo «suficientemente bueno» (usar la matrícula, por ejemplo) y al mismo tiempo conocer el sesgo, para justamente saber hasta qué punto se cuenta con algo que efectivamente es lo «suficientemente bueno» combinando criterios que permitan usar eficientemente los recursos disponibles.

Por ejemplo, si el procedimiento descrito (usar la matrícula total) introduce una sobreestimación no mayor a, digamos, 1% del valor calculado, de modo que el 25,7% antes indicado en realidad sería un valor que se encuentra entre 25,4 y 26,0%¹⁰⁰ y corregir este sesgo supone un esfuerzo muy grande movilizándolo muchos recursos, tal vez valga concluir que se puede vivir cómodamente con ese nivel de sesgo (siempre y cuando tengamos una buena estimación de su tamaño) y destinar esos recursos a otras mejoras en el sistema.

Ahora bien, la **cobertura global** es una medida muy gruesa de la cobertura, ya que no cabe esperar que la matrícula en un sistema educativo sea dependiente únicamente de la población total, sino que importa tomar en cuenta la estructura de edades de la población.

De esta forma, se puede calcular la cobertura para cada edad simple y/o por grupos de edades. Esta medida existe hoy en día y la Unesco la llama **tasa de escolarización por edad específica**, nombre completamente críptico que no ayuda en nada a comunicar su significado. Por esta razón, hemos optado ya desde hace un tiempo por llamarla **cobertura total (por edades)**¹⁰¹. En términos formales tendríamos:

⁹⁹ Este sesgo no existiría si en vez de usar la información de matrícula se utilizase la información de un censo de población (asumiendo que dicho censo pregunte lo que aquí se necesita). Esto muestra cómo la forma en que se captura la información (la unidad de observación para la que se hace el registro) tiene un impacto sobre el resultado que se computa. Lo mismo se aplica a una encuesta de hogares, pero en este caso el error muestral debe ser tomado en cuenta. Por otra parte, en ambos casos (censos y encuestas) es muy importante revisar cómo se formula la pregunta y en qué período.

¹⁰⁰ Nótese que «1% del valor calculado» es igual a 1% de 25,7, es decir, 0,3 puntos porcentuales. Esto debe mencionarse, pues existe cierta tendencia a entender una afirmación como la anterior como equivalente a «un punto porcentual», que no es lo mismo.

¹⁰¹ Con este nombre aparece en Guadalupe *et al.* (2002); Guadalupe (2002); PRIE (2003); Unesco Santiago (2008).

Donde:

CT_e : cobertura global de la edad (o grupos de edades) e

Mat_e : población de la edad (o grupos de edades) e matriculada en al menos un programa educativo

Pob_e : población total de la edad (o grupos de edades) e

x : factor que puede ser igual a 1, 100 o 10.000 (fracción, porcentaje, por cada 10.000 habitantes).

$$CT_e = \frac{Mat_e}{Pob_e} * x$$

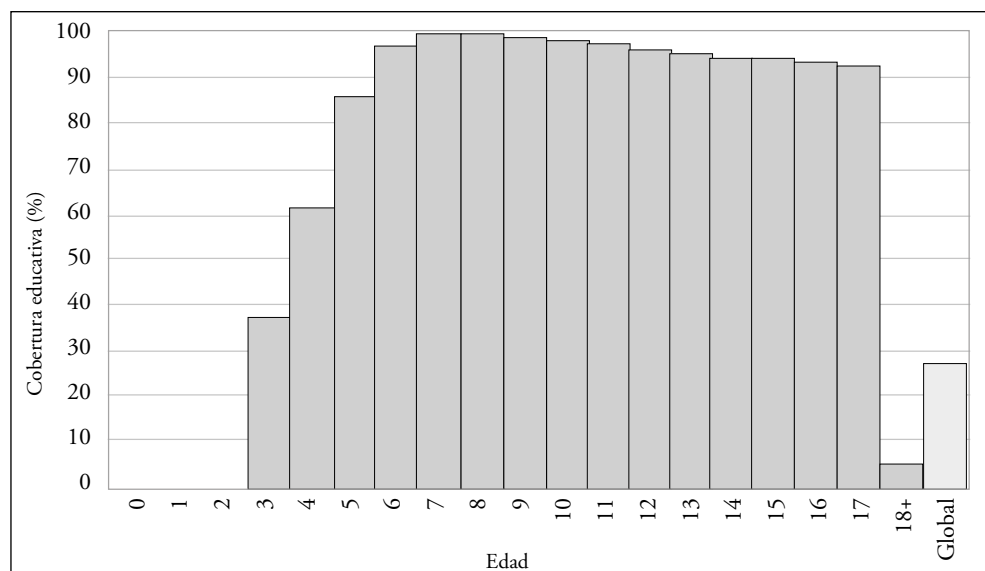
Se puede calcular para cada jurisdicción para la que se cuente con información de matrícula y población, en la medida en que estas sean consistentes (que el lugar de matrícula pueda ser asumido como lugar de residencia).

La observación hecha en el caso anterior acerca de la posible utilización, y riesgo de sobrestimación, de la información de la matrícula como medida del número de personas también es de aplicación en este caso.

Asimismo, resulta claro que el complemento de esta medida es exactamente el número (absoluto o relativo) de personas que **no** son atendidas por el sistema educativo. Es decir, esta medida permite determinar el volumen absoluto y relativo de personas que se encuentran fuera de la escuela (en realidad, para ser más preciso: que no acceden a ningún programa educativo de cualquier modalidad).

En el ejemplo mencionado, podríamos tener lo que se muestra en el gráfico 9.

Gráfico 9. Cobertura educativa global y total por edades simples, País A, 2012 (en porcentajes)



Nota

Valores hipotéticos presentados con fines de ilustración.

Sin embargo, esa medida tampoco es suficiente, ya que el hecho de que las personas sean atendidas por el sistema no es equivalente a afirmar que dicha atención se da de la forma deseada o esperada; por lo mismo, es posible hacer una precisión mayor a este enfoque: no basta con la cobertura total, vale la pena ver también la cobertura en el **nivel** (es decir, si las personas son servidas en el nivel educativo que se espera dada su edad) y en el **grado** correspondiente (dada su edad). Esta diferencia es también importante, pues una persona puede estar en el nivel que «le corresponde» pero con marcado atraso escolar, que la tasa en el nivel no mostraría o, más bien, velaría. Así, se puede calcular también la **cobertura en el nivel** y la **cobertura en el grado** o **cobertura oportuna**.

La **cobertura en el nivel** es la medida tradicionalmente conocida como **tasa neta de matrícula**, nombre que también es poco ilustrativo del contenido que mide¹⁰².

Por su parte, la **cobertura oportuna** cuenta con dos medidas que son su complemento: la **cobertura tardía** (cuando la persona se encuentra «atrasada», es decir matriculada en un grado inferior al que le correspondería dada su edad) y la **cobertura precoz** (cuando la persona se encuentra «adelantada», es decir matriculada en un grado superior al que le correspondería dada su edad). Ambas situaciones de extraedad son distintas y cabe registrarlas de modo independiente.

En términos formales tendríamos:

	Donde:
	CN_e : cobertura en el nivel de la edad e
$CN_e = \frac{Mat_e^n}{Pob_e} * x$	Mat_e^n : población de la edad e matriculada en el nivel n
	CO_e : cobertura oportuna de la edad e
$CO_e = \frac{Mat_e^g}{Pob_e} * x$	Mat_e^g : población de la edad e matriculada en el grado g
	CTa_e : cobertura tardía de la edad e
$CTa_e = \frac{Mat_e^{g+}}{Pob_e} * x$	Mat_e^{g+} : población de la edad e matriculada en grados superiores a g
	CP_e : cobertura precoz de la edad e
$CP_e = \frac{Mat_e^{g-}}{Pob_e} * x$	Mat_e^{g-} : población de la edad e matriculada en grados inferiores al grado g
	Pob_e : población total de la edad e
	x : factor que puede ser igual a 1, 100 o 10.000 (fracción, porcentaje, cada 10.000 habitantes).

Se pueden calcular para cada jurisdicción para la que se cuente con información de matrícula y población, en la medida en que esta sea consistente (que el lugar de matrícula pueda ser asumido como lugar de residencia).

¹⁰² Más adelante se presenta una reflexión sobre los indicadores de matrícula tradicionales, en la que se aventura una explicación sobre su génesis y, asociada a esta, sus nombres. Para las definiciones usuales de indicadores hoy existentes consultar Unesco-UIS (2010b).

Una vez más, debe tenerse presente que los indicadores están definidos a partir de datos de la población, aunque la información sobre matrícula puede ser lo suficientemente buena para el propósito.

La tabla 2 y el gráfico 10 ilustran la situación observable en el País B y muestran claramente cómo cada indicador mide fenómenos distintos que deben ser vistos de modo concurrente para contar con una mejor reconstrucción de los aspectos vinculados a la cobertura.

Tabla 2. Cobertura educativa total, oportuna y tardía, por edades simples, País B, 2012 (en porcentajes)

Edad	Cobertura total (%)	Población no atendida (%)	Cobertura oportuna (%)	Cobertura tardía (%)
3	11,5	88,5	11,5	0,0
4	32,0	68,0	32,0	0,0
5	64,2	35,8	64,2	0,0
6	82,9	17,1	77,3	5,6
7	91,8	8,2	61,7	30,1
8	96,2	3,8	59,0	37,2
9	91,8	8,2	56,9	34,9
10	85,7	14,3	56,1	29,6
11	83,0	17,0	55,2	27,8
12	65,9	34,1	39,4	26,6
13	52,5	47,5	38,6	13,9
14	47,5	52,5	38,4	9,1
15	45,5	54,5	38,0	7,5
16	43,8	56,2	37,1	6,8
17	43,1	56,9	37,0	6,2

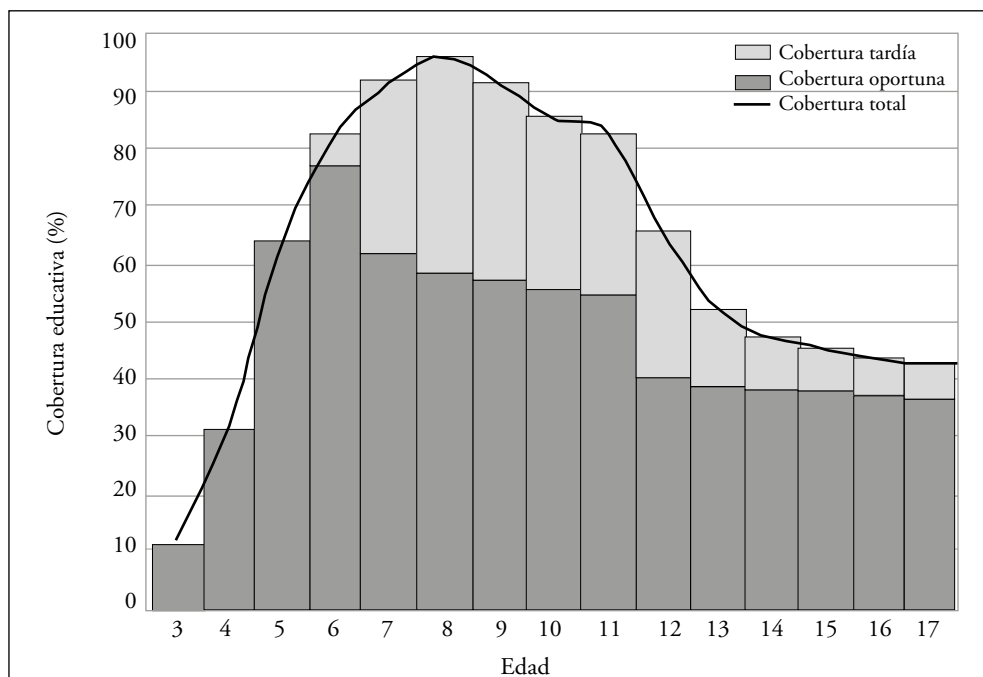
Nota

Valores hipotéticos presentados con fines de ilustración. A efectos de contar con un ejemplo simple, se ha asumido que en este país hipotético no hay matrícula precoz.

La información contenida en la tabla 2 y en el gráfico 10 no requiere de entrenamiento previo para su lectura e interpretación. Cualquier usuario verá de modo directo qué quiere decir que exista un porcentaje de la población que no es atendido; verá también que, si bien la matrícula oportuna a la edad de ingreso a la educación obligatoria (seis años en primer grado de educación primaria) no parece ser muy baja (una cobertura oportuna del 77%), el atraso escolar aumenta con las edades de los estudiantes hasta que la cobertura en edad solo llega a un 37% en la edad final aquí considerada; y asimismo podrá observar que la cobertura en edad cae con una mayor pendiente entre los seis y siete años y entre los once y doce años, es decir, debe haber una importante tasa

de repetición en el primer grado y problemas en el tránsito a la educación secundaria (aunque estos son aspectos de la dinámica del sistema que veremos más adelante).

Gráfico 10. Cobertura educativa total, oportuna y tardía, por edades simples, País B, 2012 (en porcentajes)



Nota
Valores hipotéticos presentados con fines de ilustración.

Por otra parte, la cobertura oportuna o cobertura en edad es un indicador directo y no sesgado de la probabilidad real de una cohorte dada de haber alcanzado un grado educativo dado en el tiempo establecido. Así, para una población nacida en el año t que debe ingresar a la escuela en el año $t + i$, la tasa de cobertura en edad muestra la probabilidad conjunta de haber ingresado, aprobado cada grado y sido promovido al siguiente sin atraso. Es decir, muestra el grado de éxito del sistema en cumplir esta parte de sus responsabilidades (asegurar el ingreso oportuno y el progreso a través del sistema sin contratiempos). Por esta razón, la tasa de cobertura oportuna o de cobertura en edad es un excelente indicador resumen del desempeño de los sistemas educativos en cuanto a estos aspectos¹⁰³.

¹⁰³ Así, en una formulación inicial que desarrollamos con José Rodríguez y Juan Pablo Silva en el Ministerio de Educación del Perú hace más de quince años (Guadalupe *et al.* 2002: cap. 1), este indicador fue llamado índice global de escolarización, nombre no necesariamente muy feliz, pero que apuntaba a destacar su sentido como medida directa, insesgada y comprensiva de los principales aspectos vinculados al acceso al sistema educativo.

En el caso mencionado, para la cohorte que en 2012 tenía quince años, la probabilidad de haber tenido una experiencia escolar exitosa –en términos de flujo (ingresar oportunamente y progresar sin atrasos)– ha sido de 38,0% y los objetivos de política consisten en hacer que esta probabilidad se aproxime a 100%, lo que, entonces, muestra la magnitud del desafío pendiente.

Por otra parte, si se cuenta con información de suficiente calidad, es posible calcular estos indicadores para cualquier desagregación para la que se cuente con datos de matrícula y población y los cálculos no se verán afectados por las migraciones o los traslados.

Resulta evidente que estos indicadores dan información relevante sobre la cobertura del sistema de una manera que las actuales tasas de «participación» no pueden mostrar y que, incluso, pueden velar.

Este último punto lleva a considerar el hecho de que en el lenguaje común de los sistemas de información educativa no se hace referencia a la cobertura, sino que se suele hablar de **escolarización** o **participación**.

El primer término, escolarización, resulta fácil de manejar, pero trasunta un cierto énfasis en la prestación del servicio más que en los sujetos de este (es decir, una suerte de autocentramiento que, como ya se sugirió, hace que las entidades públicas aparezcan más preocupadas por su propia operación –gestionar escuelas y contratar maestros– que por contribuir a que las personas aprendan).

Por su parte, el segundo término, participación, siendo el más extendido, es el que transmite un significado menos claro. De hecho, es posible identificar una afinidad con el concepto usado con relación al mercado de trabajo (**participación laboral**), el mismo que da cuenta de decisiones de personas vinculadas a ser parte de la población activa de un país. Esto, si bien cabría ser de aplicación en el caso de la educación de adultos, no resulta tan ilustrativo cuando se trata de la educación de niños. Por cierto, llama la atención que una búsqueda sistemática de definiciones claras de lo que quiere decir «participación» en educación en distintas publicaciones que hablan de esta sea una empresa infructuosa¹⁰⁴.

Es posible barruntar que la preferencia por la imprecisión del término esté asociada al hecho de que las medidas que usualmente se agrupan bajo esta categoría (ratios brutos y tasas netas de matrícula), justamente por ser agrupadas como si refirieran a fenómenos similares, no tienen un sentido preciso y, de alguna manera, se evita usarlas para los temas que importan (como la cobertura), aunque personas poco cuidadosas terminan acuñando usos profundamente discutibles, como hablar de «cobertura bruta».

Detengámonos aquí a revisar las actuales tasas de «participación».

¹⁰⁴ Siendo la educación de adultos la excepción, ya que es claro que, en este caso, existe un esfuerzo por subrayar el elemento de voluntad implícito en «participar» de un programa educativo. Por esta razón, los «estudiantes» de programas de educación de adultos son usual y deliberadamente llamados «participantes».

La tabla 3 muestra la información necesaria a efectos de calcular las «tasas de participación» para el País A.

Tabla 3. Población en edad de escolarización de primaria y secundaria y matrícula, País A, 2012

Nivel y edad	Población total	Matrícula total	Matrícula en el nivel que corresponde a la edad	Matrícula en el grado que corresponde a la edad
Primaria	1.300.000	1.303.283	1.277.700	
6 años	216.687		209.590	209.590
7 años	216.601		214.710	206.512
8 años	216.588		215.561	205.224
9 años	216.623		213.984	203.105
10 años	216.697		212.707	200.504
11 años	216.804		211.148	194.897
Otras edades		25.583		
Secundaria	1.297.924	1.269.931	1.198.386	
12 años	216.840		193.441	193.441
13 años	216.749		201.385	192.375
14 años	216.555		202.295	191.764
15 años	216.297		202.233	190.957
16 años	215.919		199.943	190.455
17 años	215.563		199.089	190.156
Otras edades		71.545		

Nota:

Valores hipotéticos presentados con fines de ilustración.

En este sistema educativo, que tiene una primaria de seis grados y una secundaria también de seis grados, la ratio bruta de matrícula en primaria es del 100,3% y la de secundaria del 97,8%, mientras que la tasa neta de primaria bordea el 98,3% y la de secundaria el 92,3%.

Las **ratios brutas de matrícula** son el cociente entre la matrícula total de un nivel y una población de referencia (la que en teoría podría o debería estar matriculada en dicho nivel). Como se desprende de la definición, el numerador no es necesariamente un subconjunto del denominador, ya que puede incluir a cualquier estudiante, independientemente de su edad. Así, este indicador debería, para empezar, ser llamado **razón** o **ratio** y no tasa, ya que compara unidades pertenecientes a dos conjuntos de naturaleza diferente.

Así, la pregunta que cabe hacerse tiene que ver con el propósito de calcular dicha ratio de esa forma. Una ratio como esta permite expresar un volumen en función de otro, es decir, es una **medida relativa del volumen de la matrícula**, el mismo que es expresado usando unidades determinadas, en este caso, el tamaño de una población de referencia. Así, una ratio bruta de matrícula en educación primaria del 100,1% quiere decir que, por determinados motivos, el sistema educativo está atendiendo en educación primaria un volumen de matrícula 0,1% superior al que cabría esperar en circunstancias «perfectas».

¿Qué conclusiones de política pueden derivarse de esto?: (i) ¿que hay que expulsar a ese porcentaje adicional que hace que el sistema cueste más de lo que debería?; (ii) ¿que el sistema está en condiciones de atender a todos los que debería atender –hay capacidad de oferta– y, por lo tanto, hay que preocuparse por que los matriculados sean los que deben y no en construir escuelas o contratar más docentes?

La primera conclusión es fácil de descartar, ya que si hay estudiantes que no «deberían» estar en el nivel eso se explica porque: (i) ingresaron precozmente; (ii) ingresaron tardíamente; (iii) se matricularon más de dos veces en el mismo grado¹⁰⁵; o una combinación de (i) y (iii) o de (ii) y (iii).

La segunda potencial conclusión requiere un poquito más de esfuerzo para ser respondida con propiedad. ¿Es cierto que un sistema que atiende un número de estudiantes que es superior al volumen de su población objetivo cuenta con la capacidad para atender a dicha población objetivo?

En aras de mantener la discusión en los términos más simples posibles, imaginemos un país que cuenta con un 90% de su población viviendo en zonas urbanas o periurbanas, mientras que el 10% restante habita en zonas rurales con significativas dificultades para el tránsito hacia y desde las áreas urbanas. Imaginemos también que en las áreas urbanas existe una importante dotación de servicios educativos que permite que todos los niños en edad de cursar la educación primaria lo estén haciendo, aunque también se observa ciertos niveles de atraso escolar. Por su parte, asumamos que en las áreas rurales no existe ninguna prestación de servicios educativos y que, dadas las dificultades para el tránsito, los niños en edad escolar habitando en dichas áreas no pueden acceder a los servicios educativos en las zonas urbanas.

¹⁰⁵ Nótese que este fraseo podría haberse sustituido por «repetieron uno o más grados», pero se ha evitado la palabra repetición ya que, en determinados contextos, esta solo incluye a la matrícula en el mismo grado en dos períodos lectivos **consecutivos** y/o cuando la persona reprobó la vez anterior. El fraseo usado permite incluir a aquellos que se matriculan en el mismo grado independientemente de la secuencia en períodos lectivos (pueden haber interrumpido sus estudios por uno o más períodos lectivos) e independientemente de sus resultados académicos en el grado en cuestión (pueden haber abandonado los estudios sin haber reprobado). El tema de la edad, que es el que está en discusión, es dependiente del tiempo y no del resultado académico (reprobación, retiro o abandono temporal) o de la formalidad en el registro del flujo de la matrícula (repetidores y/o reentrantes).

En un caso como el mencionado se podría observar una tasa neta de matrícula en educación primaria de aproximadamente 90% (todos los habitantes de áreas urbanas acceden y estos son el 90% del total) y una ratio bruta de matrícula que, como resultado de la repetición y el ingreso tardío, digamos, alcanza un valor de poco más de 100%.

Sin embargo, como ya se anotó, toda la matrícula se encuentra concentrada en las áreas urbanas y no existe oferta educativa en las zonas rurales. Por lo tanto, concluir que dicho valor de 100% (o superior) de la ratio bruta de matrícula indica que el sistema educativo cuenta con la capacidad de oferta suficiente, dada su población, sería una conclusión errónea que sumiría a una fracción de la población en la marginación.

Esto se explica porque el acceso a los sistemas educativos está mediado por diferentes factores que se traducen en que para distintos grupos de población el esfuerzo necesario para llegar a la escuela (costos de transporte o de eventual migración, aunque sea temporal, entre otros) sea diferente. Este esfuerzo es lo que se conoce como **costos de transacción**.

Por lo tanto, las **ratios brutas de matrícula** no miden la capacidad de oferta a menos que se pueda razonablemente asumir que los costos de transacción vinculados al acceso al sistema son iguales para toda la población, supuesto que es muy difícil de asumir, excepto en un ejercicio teórico o en una discusión sobre escenarios hipotéticos. Es decir, encontrar costos de transacción iguales en el mundo real sería una curiosidad digna de destacar y que habría que demostrar con mucho rigor.

Entonces, la segunda conclusión de política (no se necesita más escuelas o maestros) sería tan nefasta como la primera (expulsar estudiantes), aunque es menos evidente a simple vista y, de hecho, es la conclusión usual cuando se trata de «análisis» basados en estas ratios.

Lamentablemente, hay muchos ejemplos en los que la dificultad para encontrar significado y utilidad para las **ratios brutas de matrícula** lleva a considerar que, en efecto, son un indicador de capacidad de oferta (Unesco-UIS 2009a: 10).

Las ratios brutas de matrícula miden el volumen de la matrícula y lo expresan en términos de una unidad de referencia dada. De hecho, brindan la misma información que el número de estudiantes matriculados (que es el volumen absoluto y no relativo) y es posible pensar en usar otras unidades de referencia que puedan tener sentido para determinados propósitos o, por lo menos, no conducir al equívoco que el procedimiento actual implica.

Por ejemplo, en Unesco Santiago (2008) no solo no se incluye ni una sola ratio bruta de matrícula, sino que en el único caso en el que esto hubiese sido posible, o incluso deseable¹⁰⁶, se optó por recalcular la ratio usando otro valor de referencia. Así,

¹⁰⁶ En el caso de niveles educativos en los que no se aspira a una cobertura universal o donde no cabe esperar un ajuste mayor entre edades y ciclos lectivos (ambos criterios suelen ser de aplicación en el caso de

en la sección relativa al acceso a la educación terciaria (Unesco Santiago 2008: 93 y ss.), la matrícula de dicho nivel se expresa en términos de los habitantes del país, de modo que se obtiene una proporción por cada cien mil habitantes¹⁰⁷. De esta manera, se usó información sobre volúmenes donde correspondía y de una forma que no se presta a los equívocos que las actuales ratios brutas de matrícula conllevan.

Por otra parte, es importante notar que cuando se mide volúmenes no siempre «más» es sinónimo de «mejor» o «menos» de «peor». A efectos de contar con un ejemplo simple, veamos un caso real para el que calculamos la ratio bruta de matrícula en un único grado y no en todo el nivel.

En 1993, la ratio bruta de matrícula de primer grado de primaria en el Perú era 162,4%. Cinco años después, en 1998, dicho valor era de 129,3%. ¿Cómo interpretar ese cambio?, ¿buenas o malas noticias?, ¿mejoró o empeoró la situación?

Veamos la tabla 4, que nos ayuda a entender el problema.

Tabla 4. Matrícula en primer grado de primaria y población de referencia, población de seis años, Perú, 1993 y 1998

Año	Población de seis años	Matrícula en primer grado	Matrícula oportuna en primer grado	Matrícula de otras edades
1993	564.812	917.485	493.772 (87,4%)	423.713
1998	572.410	740.210	530.730 (92,7%)	209.480

Nota

La información de población fue tomada de Guadalupe (2001: 26) y la de matrícula oportuna fue computada con la de los censos escolares 1993 y 1998 (descargados de <<http://escale.minedu.gob.pe>> en agosto de 2001).

La tabla 4 muestra claramente que entre 1993 y 1998 algo pasó con la población en extraedad en el primer grado. En efecto, la población matriculada oportunamente (de seis años y matriculada en primer grado) aumentó en poco más de cinco puntos porcentuales (asumiendo que los datos usados tienen la misma calidad en estimar la matrícula omisa en el evento censal), mientras que la matrícula en extraedad disminuyó dramáticamente.

El cambio ocurrido fue resultado de la introducción, en 1995, de lo que entonces se llamó el «programa de articulación» entre el preescolar y la primaria, que trató como un ciclo único a tres grados que comprenden los dos primeros de primaria y el último del

la educación terciaria), una medida de volúmenes puede tener sentido, pero no lo tiene en los otros casos, en los que las preguntas de política tienen que ver con objetivos de cobertura y oportunidad de la matrícula.

¹⁰⁷ Nótese que en este caso numerador y denominador sí pertenecen al mismo universo (residentes en el territorio de un determinado país) y, por lo mismo, se cuenta con una tasa. En la misma publicación se modifica este indicador de modo adicional a efectos de calcularlo no solo en base al criterio territorial, sino al criterio de residencia permanente, de modo que se pueda contabilizar a los nacionales que estudian en programas ofrecidos en otros países para que la movilidad internacional de los estudiantes de educación terciaria sea tomada en cuenta.

preescolar) y eliminó la repetición dentro de dicho ciclo. Esto significa que en 1993, la matrícula de primer grado estaba «abultada» por la repetición y, a partir de la eliminación de la repetición, la matrícula empezó a tener un tamaño próximo a lo que cabe esperar en situaciones «ideales».

No cabe aquí discutir el efecto de la promoción automática sobre el logro de los aprendizajes¹⁰⁸. En cualquier caso, para lo que se viene argumentando en esta sección, la eliminación de la repetición se traduce en una mejora en el flujo de la matrícula y en las características de esta (menos atraso escolar). Desde ese punto de vista, la caída de la tasa de matrícula en el grado es una buena noticia.

Ahora bien, la caída del volumen de la matrícula no siempre es una buena noticia. Puede también ser el resultado de otros fenómenos (por ejemplo, una decisión política que lleva a la exclusión de las niñas de las escuelas luego de que se había logrado ampliar sus niveles de acceso).

Entonces, las ratios brutas pueden subir por buenas razones (aumento de la atención) o por malas razones (mayores niveles de repetición) y no lo muestran. Así, tomarlas de modo aislado, asumiendo que «más es mejor», es un error metodológico potencialmente muy peligroso. Lamentablemente, hay muchos ejemplos de esto que incluyen:

- El índice de desarrollo humano del Programa de las Naciones Unidas para el Desarrollo (PNUD). En efecto, el componente de educación de este índice estuvo compuesto, entre 1995 y 2011, por las tasas de alfabetismo (que discutiremos más adelante) y las ratios brutas de matrícula combinadas de educación primaria, secundaria y terciaria. La «Guía para el lector» donde se abordan las explicaciones metodológicas al respecto se encuentra en UNDP (2009: 203 y ss.).
- El índice de la EPT producido por el equipo del Informe Mundial de Seguimiento de la EPT. Este índice considera en su cálculo los índices de paridad de las ratios brutas de matrícula.

Por lo indicado anteriormente (que no cabe asumir que «más» es «mejor»), resulta imposible interpretar de modo unívoco cuál es el significado de la (dis)paridad de estas ratios brutas. Que haya más niños que niñas puede querer decir que

¹⁰⁸ Algunas personas sostienen que la promoción automática no resuelve el problema de que los niños logren o no los aprendizajes esperados. Siendo esto seguramente cierto (la promoción automática en sí misma no tiene una implicancia académica inmediata, al menos para el grado que se concluye), habría que preguntarse si la repetición sí sirve a ese fin. Es decir, si la repetición tampoco garantiza los aprendizajes –como parece ser el caso en muchas circunstancias con efectos devastadores adicionales en la equidad (Demeuse *et al.* 2001)–, entonces, ¿para qué incurrir en la repetición? Si la promoción automática y la repetición son igualmente ineficaces con relación a los aprendizajes, ¿por qué elegir la repetición, si es que: (i) es más costosa para los padres y la comunidad; (ii) eliminarla liberaría recursos que podrían destinarse a mejorar los aprendizajes; (iii) la repetición aumenta los costos de oportunidad de estudiar; (iv) la repetición afecta la autoestima y los procesos de socialización de los alumnos; etc.?

aquellos logran mayores niveles de acceso que estas o que estas presentan menores niveles de repetición que los primeros. Del mismo modo, la paridad puede ser el resultado de fenómenos muy distintos, donde las diferencias en acceso y repetición terminan compensándose en un valor medio carente de significado.

- Un innumerable conjunto de trabajos poco cuidadosos, que se basan en encontrar asociaciones entre la «expansión de la matrícula» medida mediante la ratio bruta de matrícula y, por ejemplo, la dinámica del crecimiento económico –por ejemplo, Wolff (2000) o Benavot (1992)¹⁰⁹–. En algunos casos, estos trabajos son valiosos y sugieren reflexiones o presentan hallazgos muy importantes, pero estos obedecen más a otros factores que a la equívoca equiparación de la expansión del acceso con el crecimiento de la matrícula sin problematizar el fenómeno.

Del mismo modo, existe una tendencia a considerar que, a falta de datos de matrícula por edades, las ratios brutas de matrícula pueden usarse como un *proxy* de los niveles de acceso medidos, por ejemplo, con las llamadas **tasas netas de matrícula**.

Para que una medida sea considerada *proxy* o aproximación de algo debe medir un fenómeno que se encuentre claramente correlacionado con el fenómeno de interés, de modo que ambos covaríen. Esta covariación puede tener distintas formas (lineal, no lineal; directa, inversa; etc.) pero sin dicha covariación una medida no puede ser un *proxy* de algún fenómeno. Por ejemplo, cuando se toma el valor del producto per cápita (que es un indicador relativo –al tamaño de la población– de la riqueza total producida en un país) como *proxy* de niveles de vida, se está asumiendo que existe una relación directa entre ambas cosas (a mayor riqueza en el país, más altos los niveles de vida). Por las razones ya expuestas, las ratios que miden volumen no covarían con el acceso (a menos que se controle un conjunto adicional de variables, lo que no es el caso) y, por lo mismo, no son un *proxy* de este.

Lo anterior, por tanto, muestra que tampoco es adecuado usar las ratios brutas de matrícula como un sustituto de las llamadas **tasas netas de matrícula** cuando estas no se encuentran disponibles, ya que miden dos cosas completamente diferentes.

En conclusión, las ratios brutas de matrícula miden volúmenes y estos se ven afectados por el acceso y por la extraedad (en el ingreso o por el flujo). ¿Existe alguna forma de «limpiar» el efecto de la extraedad?

Algunas personas consideran que esto es posible tratando de modo conjunto las ratios brutas y las tasas netas de matrícula. Dado que estas últimas sí son una tasa (numerador y denominador refieren a lo mismo), excluyen del rango de cálculo a la matrícula y a la población en edades fuera de lo que se considera propio de un nivel educativo dado.

¹⁰⁹ Este segundo texto agrega al problema el no detenerse a identificar explícitamente si las tasas de matrícula que usa son netas o ratios brutas.

Pero la extraedad no es un fenómeno que se defina con relación a un nivel educativo dado, sino con relación a un **grado** dado. Por ejemplo, en un sistema educativo con una primaria de seis grados y edad de ingreso de seis años, el caso de un niño de once años matriculado en primer grado, si bien está matriculado en el nivel que le corresponde, no puede esconder el hecho de que tiene cinco años de atraso.

Por lo tanto, la operación

$$OA_n = GER_n - NER_n$$

Donde:
 OA_n : atraso escolar en el nivel n
 GER_n : ratio bruta de matrícula en el nivel n
 NER_n : tasa neta de matrícula en el nivel n .

resulta equívoca, ya que subestima el atraso.

Por su parte, las tasas netas de matrícula, como ya se indicó anteriormente, miden lo que hemos llamado cobertura en el nivel. En ese sentido, la tasa neta de matrícula es un indicador directo de este fenómeno y no tiene ningún problema conceptual (excepto un nombre que no ayuda a entender qué es lo que mide).

Sin embargo, es un indicador insuficiente del fenómeno de la cobertura, ya que no considera a las personas que pueden estar matriculadas en otro nivel educativo y tampoco la oportunidad de la matrícula, que es un fenómeno de singular importancia. Por estas razones, hemos sugerido que se incorpore a un conjunto más amplio de indicadores de cobertura en el que, con una nomenclatura ajustada, su sentido y alcance resultan mucho más precisos.

Una lectura no informada de estas tasas puede, por ejemplo, llevar a concluir que en América Latina y el Caribe (región que tiene una tasa neta de matrícula de aproximadamente 92% en educación primaria y 73% en educación secundaria –según datos agregados correspondientes a 2012–)¹¹⁰ la primaria es prácticamente universal, mientras que cerca de un cuarto de la población en edad de estudiar secundaria está excluida del sistema. Los equívocos en estas lecturas están en que «primaria universal» solo quiere decir que los niños que deben cursar primaria acceden a esta y no necesariamente la culminan; por su parte, el otro 27% de la población en edad de estudiar secundaria no necesariamente está fuera del sistema, una parte de ella (no sabemos cuántas personas a partir de estas tasas) se encuentra en primaria o en programas de otros niveles de educación.

Cabe indicar también que en los últimos años se ha venido usando una tasa neta de matrícula «ajustada» para la educación primaria. El ajuste consiste en incorporar en el

¹¹⁰ El dato regional agregado más reciente está disponible en el Centro de Datos del UIS, al 31 de marzo de 2015.

numerador la matrícula de personas con edad para estar en primaria que se encuentran matriculadas en educación secundaria (y que, por lo tanto, accedieron a la primaria previamente). De esta forma, la tasa «ajustada» se aproxima (pero no es idéntica, pues excluye otros niveles educativos) a una tasa de cobertura total tal y como se ha definido en páginas anteriores. Estos ajustes se encuentran presentes en Unesco-UIS (2009a) y en el informe mundial de seguimiento de la EPT desde el año 2006.

Cabe, sin embargo, hacer algunas precisiones sobre este «ajuste»:

- ¿Por qué llamar ajuste a un cambio en la naturaleza del indicador? Esto puede llevar a más confusión, ya que sugiere que la tasa original y la tasa ajustada miden lo mismo, solo que la «ajustada» lo hace de una «mejor» manera. Si este fuera el caso, ¿no habría que descartar la «peor» manera?
- El «ajuste» es insuficiente, pues excluye por razones no documentadas la matrícula en educación preescolar u otros niveles educativos. ¿Por qué? ¿Cuál es el sustento conceptual para proceder en tal sentido?
- Si se hiciera el «ajuste» correctamente, es decir, incluyendo toda la matrícula y no solo algunos niveles seleccionados sin un criterio explícito, el indicador sería completamente innecesario, ya que existe otro indicador que hace exactamente eso: la llamada **tasa de escolarización por edad específica** que, como se ha mencionado antes, debería abandonar ese nombre y ser simplemente llamada tasa de cobertura total calculable por edades simples o por rangos de edades¹¹¹.
- Una obligación básica de quienes trabajan con información es documentar claramente los cambios y orientar a los usuarios sobre el uso que se hace de los datos.

Si estos indicadores tan usuales tienen tantos límites, ¿no resultaría más conveniente y claro simplemente redefinir algunos indicadores, usar otros y promover un uso directo e inequívoco de indicadores que sean relevantes? Este es uno de los esfuerzos que se desarrolló en Unesco Santiago (2008). Sin embargo, llevar esto a la práctica no siempre es posible, ya que a muchas personas les resulta difícil aceptar que hay que corregir cosas que han venido haciendo de la misma manera por cinco, diez, veinte años, sin nunca advertir un problema que, como se ha mostrado, no era muy complicado.

Ahora bien, es importante anotar que las ratios brutas de matrícula no son el resultado de la incompetencia de sus creadores. De ninguna manera. Son el resultado de un momento dado en el que la expansión de los sistemas educativos, en términos de

¹¹¹ Cabe anotar que el indicador se puede calcular de cualquiera de esas maneras sin que esto signifique que son «varios» indicadores, ya que el cambio en la población objeto del cálculo no redefine lo que se está midiendo. Si este fuera el caso, la medición para cada subconjunto singular de la población (edad simple, grupos de edad, sexo, países, regiones, áreas o jurisdicciones dentro de un país, otros subgrupos poblacionales) sería un indicador diferente, lo que es un completo sinsentido.

matrícula, era una preocupación capital y, por lo tanto, en el cual medir volúmenes que deben ser tenidos en cuenta para, por ejemplo, estimar necesidades de inversión era algo trascendental. La mayor parte de los indicadores educativos más conocidos hoy (las tasas de matrícula y los vinculados al estudio de los flujos de matrícula) obedecen exactamente a esa preocupación¹¹².

Por lo mismo, en defensa de las llamadas tasas de participación se puede argüir que no miden cobertura (justamente por eso es que hay que proponer medidas y nombres alternativos para hacerlo) y que más bien se abocan a dar cuenta de la matrícula.

En efecto, el análisis de la cobertura pone en el centro de la preocupación a la población en su conjunto, no a los matriculados, y por eso es de capital importancia a efectos de prevenir una mirada superficial a fenómenos básicos de exclusión.

Lo anterior no quiere decir que la matrícula no merezca un análisis propio, pero el estudio del comportamiento de la matrícula por sí mismo, si bien se encuentra vinculado con la cobertura, no es un análisis de esta.

Para terminar este acápite sobre cobertura educativa, vale llamar la atención sobre una limitación mayor para el cálculo de los indicadores de cobertura: la posible falta de consistencia entre la información de matrícula por edades y la información poblacional (asumiendo que ambas existen).

Estos potenciales problemas de consistencia están vinculados a varios aspectos:

- La información de matrícula puede estar sesgada ya que en los casos en los que es generada mediante el reporte estadístico y no mediante la carga de datos a partir de los registros administrativos (aunque estos también pueden tener problemas similares) puede incluir inexactitudes sobre la cantidad total de matriculados y sobre la distribución de estos por edades. Ciertamente, es de interés de los sistemas educativos, más allá de los fines estadísticos, contar con sistemas de registro más adecuados que eviten estas inexactitudes, así como tener presente que estos problemas afectan a cualquier dato originado en las mismas fuentes.
- La información de población por edades simples suele corresponder a proyecciones realizadas para los períodos intercensales. Estas proyecciones se hacen sobre hipótesis de fecundidad, mortalidad y migración que pueden no haberse verificado en el tiempo. Asimismo, en países cuya población es pequeña en número, los errores de las proyecciones pueden resultar demasiado elevados, de suerte que no es recomendable proceder en este sentido. Algo similar ocurre en períodos de marcados cambios demográficos o de fluctuaciones de corto plazo en variables demográficas (como resultado, por ejemplo, de algún tipo de evento catastrófico)

¹¹² En ese sentido, no debe llamar la atención que los modelos de flujo fueran presentados en los años 1970 como herramientas para estudiar el «desperdicio» en los sistemas educativos (Unesco 1972).

que no suelen ser bien capturados por esquemas de proyección que priorizan contar con tendencias suavizadas. Esto, por cierto, afecta de la misma manera a cualquier indicador que tome información demográfica para ser calculado.

En cualquier caso, estas limitaciones, que no son solo de aplicación a los indicadores aquí tratados, revelan la necesidad de que los sistemas nacionales de estadística educativa contemplen mecanismos y procedimientos de verificación tanto de la consistencia interna como, y sobre todo, de la veracidad de los datos. Estos procedimientos incluyen el uso de mecanismos muestrales o de fuentes de datos externas al sistema educativo.

Perfiles de matrícula

Para analizar los perfiles de la matrícula, las llamadas tasas de participación también son muy limitadas e introducen restricciones y fuentes de sesgo innecesarias: solo son calculables para jurisdicciones en las que se cuenta con información de población y matrícula equivalente y eso se complica a escalas menores donde o no se cuenta con proyecciones de población o la ubicación de la matrícula no se corresponde necesariamente con el lugar de residencia de los estudiantes.

La información de matrícula, por sí misma, puede ser usada para estudiar perfiles de una manera muy provechosa. Por ejemplo, es perfectamente posible calcular:

Donde:

$$MO_g = \frac{Mat_e^g}{Mat_g} * x$$

$$MT_g = \frac{Mat_{e+}^g}{Mat_g} * x$$

$$MP_g = \frac{Mat_{e-}^g}{Mat_g} * x$$

MO_g : matrícula oportuna en el grado g
 Mat_e^g : matrícula de la edad e en el grado g
 MT_g : matrícula tardía en el grado g
 Mat_{e+}^g : matrícula de edades mayores a e en el grado g
 MP_g : matrícula precoz en el grado g
 Mat_{e-}^g : matrícula de edades inferiores a e en el grado g
 Mat_g : matrícula total en el grado g
 x : factor que puede ser igual a 1, 100 o 10.000 (fracción, porcentaje, cada 10.000 habitantes).
 Se puede calcular para cualquier nivel de desagregación.

En efecto, en situaciones en las que la cobertura total no sea un problema mayor o el interés sea solo observar la composición de la matrícula, es posible concentrarse únicamente en el análisis de esta y obtener resultados convergentes con los sugeridos por las tasas de cobertura calculando para cada grado la matrícula oportuna (matriculados en edad en el grado g sobre el total de matriculados), precoz (matriculados con edad menor a la oficial o teórica en el grado g sobre el total de matriculados) y con atraso (matriculados tardíos en el grado g sobre el total de matriculados).

Estos indicadores tendrían la ventaja adicional de no requerir información sobre población y, por lo tanto, podrían ser calculados a los menores niveles de desagregación posibles, es decir, para cada programa educativo¹¹³.

Ahora bien, si se contase con información poblacional también sería posible comparar el perfil de la matrícula con el de la población de las edades correspondientes mediante la comparación directa de la matrícula en cada grado con la población de la edad correspondiente a dicho grado, es decir, con la forma óptima que dicho perfil debería tener.

Este perfil es resultado de la composición de la matrícula: una determinada matrícula oportuna o en edad, un porcentaje dado de matrícula prematura y una magnitud dada de matrícula tardía o matrícula en extraedad.

Asimismo, el perfil puede ser descrito en términos de dos indicadores: **ratio matrícula/población** y **pendiente**.

El primer indicador, la ratio matrícula/población, muestra el grado de ajuste o desajuste entre el tamaño de la matrícula total y el de la matrícula «teórica» (población de la edad correspondiente al grado) y una vez calculado para cada grado puede agregarse por niveles educativos como el promedio ponderado por la población (lo que es aritméticamente equivalente a calcular la ratio bruta de matrícula pero, en este contexto, tiene un significado claro y preciso como medida relativa del tamaño de la matrícula). En términos formales, esta ratio se expresa de la siguiente manera:

Donde:

RMP : ratio matrícula/población

Mat_g : matrícula total en el grado g

Pob_e : población total de la edad e

$$RMP = \frac{Mat_g}{Pob_e}$$

No resulta conveniente expresar esta ratio como porcentaje a efectos de evitar confusión con las tasas de cobertura.

Se puede calcular para cada jurisdicción para la que se cuente con información de matrícula y población, en la medida en que estas sean consistentes (que el lugar de matrícula pueda ser asumido como lugar de residencia).

Por su parte, el segundo indicador, pendiente, da cuenta de cómo se va transformando el perfil en el paso de cada grado y, por lo mismo, permite identificar los momentos del ciclo escolar donde se observa las caídas más marcadas y se calcula simplemente como el cociente entre el indicador calculado para un grado dado y el cociente del grado precedente.

¹¹³ En Brasil se suele calcular, en este mismo espíritu, la *taxa de distorção idade-série* es decir, la tasa de distorsión edad-serie.

2.4.2 EFICIENCIA

La eficiencia de los sistemas educativos está directamente vinculada a dos preocupaciones: (i) el uso adecuado de los recursos que la colectividad destina a la educación (a través de impuestos o de gastos directos), es decir, asegurar que estos no se desperdicien; y (ii) el flujo o trayectoria de las personas a través del sistema educativo, de modo que su experiencia sea exitosa.

La primera preocupación ha estado a la base del desarrollo de un conjunto de indicadores entroncados en lo que se conoce como modelos de cohortes, indicadores que se han difundido y son muy usados desde hace varias décadas.

La segunda preocupación ha sido menos aparente, pues un conjunto más limitado de los indicadores existentes se preocupan en sí mismos por los elementos educativos implicados en las trayectorias escolares.

Asimismo, el uso corriente suele distinguir entre dos grandes espacios de la eficiencia: el interno y el externo. En este esquema, la eficiencia interna refiere a los **procesos** internos del sistema educativo, que a su vez pueden ser vistos con relación a un ciclo o período lectivo dado (eficiencia interna anual) o al tránsito entre dos ciclos (eficiencia interna interanual); por su parte, la eficiencia externa pone el énfasis en los **resultados** obtenidos por el sistema o en lo que el sistema educativo entrega a la colectividad (nivel de escolaridad de la población, por ejemplo).

A continuación se describe y discute cada uno de estos conjuntos de indicadores en un esfuerzo por esclarecer sus fortalezas y limitaciones a efectos de mejorar su uso.

Eficiencia interna anual

Para un año o ciclo escolar dado es posible contar con información para cada grado acerca de la matrícula final y de cómo ella está compuesta en términos de la certificación académica esperada; es decir, podemos saber cuántos aprobaron y cuántos desaprobaron o reprobaron.

Del mismo modo, es posible conocer el número total de alumnos que se retiraron a lo largo del año. Eso se puede calcular, sea por diferencia con la información sobre matrícula inicial (en cuyo caso se requiere también la información sobre traslados de entrada y salida)¹¹⁴, sea porque esta se genera junto con la información de matrícula final. Es decir, con una sola toma de datos en un solo período se puede contar con dicha información. Esto es relevante, ya que permite eliminar los problemas de consistencia que se derivan de usar dos (o más) fuentes (o tomas) de datos diferentes.

¹¹⁴ Es crucial entender la naturaleza de los traslados a efectos de no asumir que estos no son importantes. Los traslados incluyen tanto los movimientos entre escuelas, como los que se realizan entre jurisdicciones administrativas e incluso los traslados internacionales. Estimar el saldo neto de traslados permite identificar, siguiendo una analogía demográfica, a las unidades que «atraen» y a las que «expulsan» matrícula. Indagar sobre estas pautas de comportamiento puede resultar muy útil.

El cálculo de estas tasas es simple y directo:

Donde:

TA_g : tasa de aprobación en el grado g

Apr_g : aprobados del grado g

TR_g : tasa de reprobación en el grado g

Rpr_g : reprobados del grado g

TRt_g : tasa de retiro en el grado g

Rt_g : retirados en el grado g , que es igual a:

$$TA_g = \frac{Apr_g}{Mat_g} * x$$

$$RT_g = abandonos_g + (traslados I_g - traslados O_g)$$

$$TR_g = \frac{Rpr_g}{Mat_g} * x$$

Donde:

$abandonos_g$: aquellos que dejaron el programa educativo durante el ciclo lectivo pero que no lo hicieron para trasladarse a otro programa sino que dejaron el sistema de modo temporal o permanente

$traslados I_g$: traslados recibidos por la unidad de observación

$traslados O_g$: traslados cedidos por la unidad de observación

$$TRt_g = \frac{Rt_g}{Mat_g} * x$$

Mat_g : matrícula total en el grado g , que es igual a:

$$Mat_g = Apr_g + Rpr_g + Rt_g$$

x : factor que puede ser igual a 1, 100 o 10.000 (fracción, porcentaje, cada 10.000 habitantes).

Se puede calcular para cualquier nivel de desagregación.

Del mismo modo, la lectura e interpretación de esta información también es sencilla: para una unidad de observación dada (país, jurisdicción, escuela, grado) en el período t , el x por ciento de los matriculados (en el grado g o en una agregación de estos) aprobó, reprobó o abandonó el grado g .

En teoría, los aprobados deberían ser aquellos que lograron los aprendizajes previstos para el período; mientras que los reprobados serían aquellos que no los lograron. Desde este punto de vista, estos indicadores podrían ser valiosas piezas de información sobre resultados de aprendizaje. Sin embargo, estas identidades no parecen sostenibles según la evidencia generada mediante las pruebas estandarizadas de logro académico. Esto plantea problemas relativos a las normas (formales o no formales) con las que realmente operan los docentes para certificar que un estudiante «aprobó» (o no) un grado y , por lo mismo, plantea preguntas centrales acerca de lo que los docentes efectivamente hacen en el aula.

El análisis concurrente de estas tasas con los resultados de pruebas estandarizadas podría brindar pistas a explorar en la política educativa.

Dada esta falta de identidad, cabe preguntarse para qué, entonces, sirven estas tasas. La respuesta es que ellas brindan información no sesgada sobre trayectorias escolares y , en ese sentido, sirven para aproximarse a fenómenos sustantivos. El hecho de permitir

cualquier nivel de desagregación, por su parte, presenta una ventaja que sugiere que se haga un uso más intensivo de tales tasas.

Ahora bien, esto nos lleva a un problema mayor ya mencionado: es de capital importancia que se explote la información que se genera a través de los sistemas de información. Los análisis de heterogeneidad hoy posibles gracias al levantamiento de información desagregada no se vienen desarrollando y eso cuestiona el sentido de hacer grandes inversiones en los operativos de acopio para obtener resultados agregados que se podían lograr en base al lápiz y el papel en décadas pasadas. Veamos en la tabla 5 un ejemplo de uso de las tasas de eficiencia interna anual con algunas desagregaciones.

Tabla 5. Matrícula oportuna en el último grado de educación primaria, Perú, 1999

Tipo de gestión de los centros educativos	Matrícula total			Matrícula urbana			Matrícula rural		
	Número de CE	Media (%)	Desviación típica	Número de CE	Media (%)	Desviación típica	Número de CE	Media (%)	Desviación típica
Estatal	21.953	23,1	22,8	4.015	39,4	17,8	17.938	19,4	22,2
No estatal	2.854	62,7	27,2	2.743	63,2	27,0	111	50,4	29,9
Total	24.807	27,6	26,5	6.758	49,1	24,9	18.049	19,6	22,3

Notas

- Solo se considera información para los centros educativos (CE) que reportaron datos al censo escolar 1999 y que cuentan con matrícula en el sexto grado. Información base descargada de <<http://escale.minedu.gob.pe>> (agosto de 2001).
- «Estatal» incluye solo a las instituciones gestionadas por el sector educación (es decir, se excluye otras formas de gestión estatal) y «No estatal» solo a aquellas de gestión «particular» (es decir, se excluye otras formas de gestión no estatal).

La tabla 5 muestra claramente: (i) un comportamiento diferenciado entre las instituciones estatales y no estatales, donde estas últimas tienen un nivel de matrícula oportuna en el grado final de la primaria mucho más alto que el observado en las primeras; (ii) un comportamiento diferenciado entre áreas urbanas y rurales en general y también según el tipo de gestión; y (iii) una variabilidad (medida como coeficiente de variabilidad, que es igual a desviación típica dividida por la media) entre las instituciones de gestión no estatal menor que entre las estatales.

Esta información podría calcularse a nivel de cada centro educativo (de hecho este ha sido el procedimiento a efectos de poder calcular la media y la desviación típica presentadas) y posteriormente agregarse de diversas formas (por circunscripciones administrativas, por ejemplo) o tratarse de modo desagregado a efectos de construir tipologías (por ejemplo, identificando centros educativos que compartan algunas características) o relaciones entre variables. Por ejemplo, se puede preguntar si los niveles de matrícula oportuna en el grado final se relacionan con el tamaño del programa (total de la matrícula) o con otras variables.

Eficiencia interna interanual

La aproximación más usual a la eficiencia está dada por las tasas de eficiencia interna interanual (principalmente promoción, repetición y deserción). Estas tasas se encuentran basadas en un modelo que busca representar los flujos interanuales de la matrícula.

El flujo de la matrícula puede ser representado tomando a una cohorte de la misma y haciendo un seguimiento «real» o, alternativamente, mediante un seguimiento «reconstruido». De ahí que los modelos puedan referirse tanto a cohortes reales como reconstruidas. La construcción de un modelo de cohorte real supone hacer un seguimiento individual de los miembros de la cohorte a través de un período dado y, a partir de ello, representar su flujo en el tiempo. Por su parte, las cohortes reconstruidas toman información agregada y asumen determinados comportamientos para definir trayectorias esperadas.

Es evidente que el flujo de una cohorte real plantea desafíos mayores en términos de la disponibilidad de información a nivel individual que permita su seguimiento. De ahí que los modelos reconstruidos hayan tenido una gran acogida, ya que con un volumen muy limitado de información permiten establecer trayectorias. Cuán fidedignas son esas trayectorias depende directamente de: (i) la capacidad del modelo mismo para representar la realidad; (ii) la validez de los supuestos usados para definir las trayectorias; y (iii) la calidad de la información usada. Discutamos estos elementos.

En general, los modelos de cohortes representan las trayectorias educativas en la forma que se presenta en el gráfico 11.

A partir de este modelamiento, se procede a reconstruir el flujo de las cohortes mediante el procedimiento que se presenta a continuación:

Se asume un volumen de partida (usualmente un número como 1.000 estudiantes, pero podría ser cualquiera) como matrícula en el grado inicial (g) en el período inicial (t). A partir de ahí, usando información interanual, se determina cuántos estudiantes se encontrarán en el período siguiente en condición de: (i) promovidos al grado siguiente ($g + 1$)¹¹⁵; (ii) repetidores en el mismo grado (g)¹¹⁶; o (iii) fuera del sistema. La tabla 6 ilustra el procedimiento.

¹¹⁵ Téngase en cuenta que «promovidos en el grado $g + 1$ » significa efectivamente matriculados en dicho grado y no solo que culminaron exitosamente el grado precedente. En algunos países la condición de «promovido» es atribuida al final de un grado, de modo que quiere decir: «culminó satisfactoriamente el grado g y se encuentra promovido al siguiente», definición que no verifica que la promoción se haya efectivamente dado como matrícula en el grado sucesivo en el siguiente período lectivo, sino que la persona está en condiciones de proseguir sus estudios (si lo hace o no, ese es otro tema).

¹¹⁶ Como en cualquier caso, es preciso prestar atención a la consistencia entre las definiciones del modelo y las usadas para generar la información que lo alimenta. Por ejemplo, hay países que consideran como repetidores a los alumnos que se matriculan en el mismo grado en dos períodos académicos consecutivos siempre y cuando lo hagan por haber desaprobado o reprobado en la primera ocasión. Otros países no requieren esta última condición, con lo que también incluyen como repetidores a aquellos que vuelven a matricularse habiéndose retirado el año anterior.

Gráfico 11. Representación de los flujos de matrícula en los modelos de cohortes

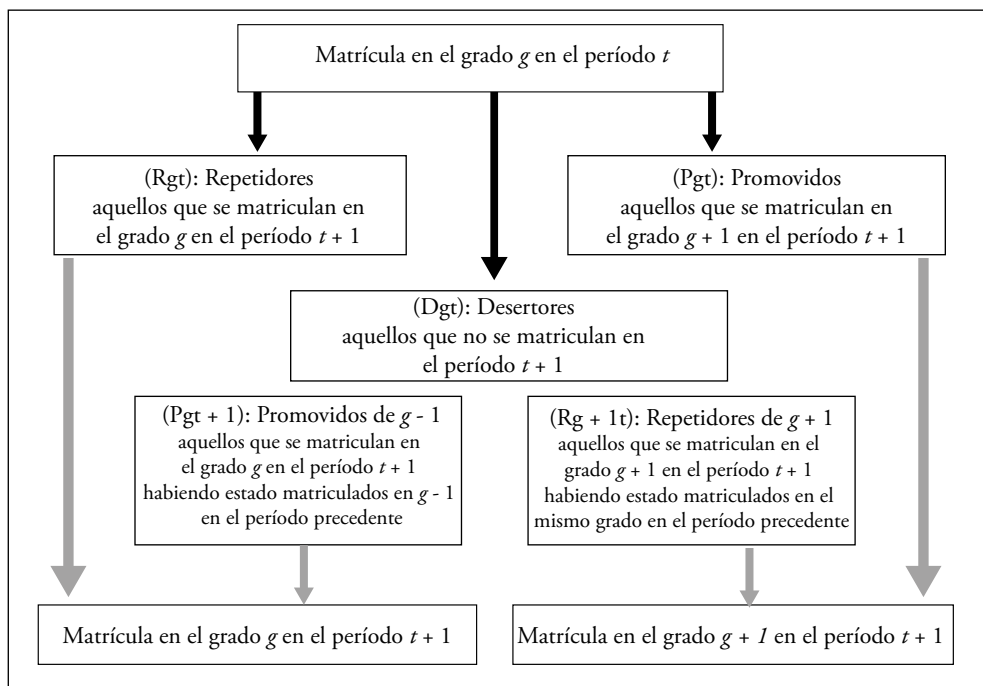


Tabla 6. Ilustración del modelo de cohortes (dos grados y un período interanual)

	Grado 1	Grado 2
Matrícula (t)	1.000	
Promovidos	815	
Repetidores	180	
Desertores	5	
Matrícula (t + 1)	180	815

Nota

Valores hipotéticos presentados con fines de ilustración.

A partir de la estructura mostrada en la tabla 6, se procede con un conjunto de iteraciones hasta lograr reconstruir todo el flujo; es decir, hasta el momento en que el «stock» inicial de estudiantes se agota en una de las dos categorías de destino: (i) se promovió hasta terminar el grado final; o (ii) desertó¹¹⁷. Nótese que la promoción y la repetición son categorías intermedias, ya que un promovido y un repetidor siguen en el sistema hasta que terminen o abandonen.

¹¹⁷ En ocasiones, se pone un límite predefinido a las iteraciones, de modo que se asume que tras un número dado de repeticiones el estudiante no repite más y deserta.

A efectos de poder aplicar un modelo de cohortes reconstruidas, se requiere la siguiente información mínima: (i) matrícula por grado; y (ii) número de repetidores por grado. Esta información se presenta en la tabla 7 para una educación primaria de seis grados¹¹⁸.

Tabla 7. Matrícula en educación primaria en dos períodos consecutivos y número de repetidores, por grados, País B, 2012-2011

	Grado 1	Grado 2	Grado 3	Grado 4	Grado 5	Grado 6	Grado 7
Matrícula en t (2011)	1.000	980	950	940	920	900	
Matrícula en $t + 1$ (2012)	1.000	955	940	900	870	850	800
Repetidores en $t + 1$ (2012)	180	140	120	90	80	60	50

Nota

Valores hipotéticos presentados con fines de ilustración.

Con esta información, se estima las siguientes tasas, que son las que alimentan el modelo:

Donde:

$TP_{g+1/g}$: tasa de promoción del grado g al grado $g + 1$

P_g : promovidos de g en t a $g + 1$ en $t + 1$, estimado como:

$$TP_{g+1/g} = \frac{P_g}{Mat_g} * x$$

TR_g : tasa de repetición en el grado g

R_g : repetidores de g $t + 1$ (dato observado)

$$TR_g = \frac{R_g}{Mat_g} * x$$

$TD_{g+1/g}$: tasa de deserción del grado g para el período $t + 1 / t$

D_g : desertores de g en el período interanual $t + 1 / t$ estimado como:

$$TD_{g+1/g} = \frac{D_g}{Mat_g} * x$$

$$Mat_g - P_g - R_g$$

Mat_g : matrícula en el grado g en el período t (dato observado)

x : factor que puede ser igual a 1, 100 o 10.000 (fracción, porcentaje, cada 10.000 habitantes).

En el ejemplo, las tasas serían las que se muestran en la tabla 8.

Tabla 8. Indicadores de eficiencia interna interanual (I), por grados, País B, 2012-2011

	Grado 1	Grado 2	Grado 3	Grado 4	Grado 5	Grado 6
Tasa de promoción	0,815	0,837	0,853	0,840	0,859	0,833
Tasa de repetición	0,180	0,143	0,126	0,096	0,087	0,067
Tasa de deserción	0,005	0,020	0,021	0,064	0,054	0,100

Nota

Valores hipotéticos presentados con fines de ilustración.

¹¹⁸ Nótese que se ha introducido información parcial de un grado adicional (grado 7) a efectos de poder calcular las tasas del grado final. Si se contara con información real de término del grado (por ejemplo aprobación y/o graduación), se utilizaría dicha información para el mismo propósito.

Una vez que se cuenta con estas tasas, se procede a calcular el modelo tal y como se muestra en la tabla 9.

Tabla 9. Modelo de cohortes reconstruidas, País B, 2012-2011

	Grado 1	Grado 2	Grado 3	Grado 4	Grado 5	Grado 6
Matrícula inicial (t)	1.000					
Promovidos	815					
Repetidores	180					
Desertores	5					
Matrícula ($t + 1$)	180	815				
Promovidos	147	682				
Repetidores	32	116				
Desertores	1	17				
Matrícula ($t + 2$)	32	263	682			
Promovidos	26	220	581			
Repetidores	6	38	86			
Desertores	0	5	14			
Matrícula ($t + 3$)	6	64	306	581		
Promovidos	5	54	261	489		
Repetidores	1	9	39	56		
Desertores	0	1	6	37		
Matrícula ($t + 4$)	1	14	92	317	489	
Promovidos	1	12	79	266	420	
Repetidores	0	2	12	30	42	
Desertores	0	0	2	20	27	
Matrícula ($t + 5$)		3	23	109	309	420
Promovidos		2	20	92	265	350
Repetidores		0	3	10	27	28
Desertores		0	0	7	17	42
Matrícula ($t + 6$)			5	30	118	293
Promovidos			5	25	102	244
Repetidores			1	3	10	20
Desertores			0	2	6	29
Matrícula ($t + 7$)			1	7	36	121
Promovidos			1	6	31	101
Repetidores			0	1	3	8
Desertores			0	0	2	12
Matrícula ($t + 8$)				1	9	39
Promovidos				1	8	32
Repetidores				0	1	3
Desertores				0	1	4
Matrícula ($t + 9$)					2	11
Promovidos					2	9
Repetidores					0	1
Desertores					0	1
Matrícula ($t + 10$)						2
Promovidos						2
Repetidores						0
Desertores						0

Nota

Valores hipotéticos presentados con fines de ilustración.

A partir del modelo, se calcula un conjunto de indicadores:

- **Tasa de supervivencia al grado g .** Esta tasa estima qué proporción de la cohorte logrará alguna vez matricularse en un grado dado. Por ejemplo, si se asume que cuatro grados de educación primaria deberían ser suficientes para asegurar que los estudiantes logren competencias básicas y sostenibles de lectura, escritura y uso de números, un indicador relevante sería el número de estudiantes que logran concluir dichos cuatro grados. Esto, tradicionalmente se ha aproximado con la tasa de supervivencia calculada al quinto grado (que subestima lo que se quiere medir, pues algunos de los que concluyeron exitosamente el cuarto grado pueden desertar inmediatamente)¹¹⁹.

Esta tasa se calcula de la siguiente manera:

Donde:

TS_g : tasa de supervivencia al grado g

$Prom_{g-1}$: promovidos desde el grado inmediato anterior ($g - 1$) en cada una de las iteraciones del modelo; es decir, el número de estudiantes que alcanza el grado g en cualquiera de los años para los que el modelo genera un número de promovidos a dicho grado

Coh : cohorte

x : factor que puede ser igual a 1, 100 o 10.000 (fracción, porcentaje, cada 10.000 habitantes).

$$TS_g = \frac{\sum Prom_{g-1}}{Coh} * x$$

- **Tasa de supervivencia sin atraso al grado g .** Esta tasa estima qué proporción de la cohorte logrará matricularse en un grado dado, habiendo transcurrido por el sistema sin repetir ninguno; es decir, qué proporción de la cohorte mostrará el comportamiento óptimo esperado. Nótese que supervivencia sin atraso no es igual a oportuna, ya que el flujo no toma en cuenta la estructura de edades de la matrícula y podría ser el caso de que el ingreso al primer grado se observe de modo mayoritario en personas de una edad diferente a la esperada, oficial o teórica.

Esta tasa se calcula de la siguiente manera:

¹¹⁹ En realidad este no es un ejemplo accidental sino la razón por la que la tasa de supervivencia al quinto grado ha sido considerada relevante por muchos años. El hecho de que la consecución de un determinado número de grados o años de educación no garantice que las personas logren determinados aprendizajes hace que este indicador pierda relevancia para este propósito más allá de sus fortalezas o defectos como esfuerzo de medición.

Donde:

$$TS_{g-a} = \frac{Mat_{g^1} * x}{Coh}$$

TS_{g-a} : tasa de supervivencia sin atraso al grado g
 Mat_{g^1} : matriculados en el grado g , solo en la primera
 ocurrencia
 Coh : cohorte
 x : factor que puede ser igual a 1, 100 o 10.000 (fracción,
 porcentaje, cada 10.000 habitantes).

- **Tasa de deserción acumulada al grado g .** Es el complemento de la supervivencia (total). Se asume que aquellos que en el modelo no superviven, desertan. Esta tasa se calcula de la siguiente manera:

Donde:

$$TDac_g = x - TS_g$$

$TDac_g$: tasa de deserción acumulada al grado g
 TS_g : tasa de supervivencia al grado g
 x : factor que puede ser igual a 1, 100 o 10.000 (fracción,
 porcentaje, cada 10.000 habitantes), teniendo que ser
 idéntico al usado para la tasa de supervivencia.

- **Años-alumno invertidos en el grado g .** Es el número total de matrículas registradas en un grado dado, de modo que considera tanto nueva matrícula como aquella originada en la repetición. Si se asume un costo por estudiante, este indicador permite estimar el volumen de inversión necesario para que una cohorte curse un grado dado. Por ejemplo, un escenario óptimo en el que 1.000 estudiantes cursen el primer grado debería suponer una inversión equivalente a 1.000 años-alumno (todos y cada alumno, una sola vez); sin embargo, la repetición hace que esto no sea así, aumentando los costos, y la deserción también impacta, reduciendo los mismos. Aquí se revela un problema conceptual en el modelo: a efectos de evitar el «desperdicio» de recursos y tener más eficiencia, es preferible que una persona deserte a que repita. Este supuesto no se condice con los objetivos de política y las finalidades de la educación en general¹²⁰. Este indicador se calcula de la siguiente manera:

$$SY_g = \sum Mat_g$$

Donde:
 SY_g : años-alumno invertidos en el grado g
 Mat_g : cada una de las ocurrencias de matrícula para el
 grado g en el modelo.

¹²⁰ Una observación equivalente a esta fue planteada a propósito del coeficiente de eficiencia en Argentina-Ministerio de Educación, Ciencia y Tecnología (s. f.: 79).

- **Años-alumno invertidos de forma acumulada hasta el grado g .** Este indicador permite representar la inversión acumulada hasta un grado dado para una cohorte de estudiantes. Por ejemplo, un escenario óptimo supondría que cada grado recibe a toda la cohorte una sola vez; así cada grado aportaría una inversión de 1.000 años-alumno y, por lo tanto, un nivel de seis grados implicaría una inversión de 6.000 años-alumno en total. Sin embargo la repetición hace que esto no sea así, aumentando los costos y, por su parte, la deserción impacta reduciéndolos. Aquí nuevamente aparece el mismo problema conceptual ya anotado: para el modelo, la deserción es «mejor» que la repetición. Sin embargo, en este caso el problema se hace aún más complejo, ya que volúmenes similares de deserción acumulada tendrán impactos diferenciados en el «desperdicio», dependiendo del momento en el que se presente. Una deserción en el grado inicial en el ciclo lectivo representa una «pérdida» menor que una deserción en el último grado, ya que en el primer caso el estudiante solo cursó un grado (solo se gastó una vez en él), mientras que en el segundo se invirtió en al menos seis oportunidades en el mismo estudiante (una por grado, más posibles repeticiones). De esta forma, al acumular los valores, un sistema resulta más eficiente mientras más temprano expulse a aquellos que no alcanzarán a culminar un nivel dado. Este indicador se calcula de la siguiente manera:

Donde:

$$SYac_g = \sum SY_g$$

$SYac_g$: años-alumno acumulados hasta el grado g
 SY_g : los valores de años-alumno para el rango de grados desde el primero hasta g inclusive.

- **Coefficiente de eficiencia.** Este coeficiente mide la eficiencia agregada del sistema a un grado dado. Para hacerlo compara los montos acumulados de inversión (en años-alumno) y el número de personas que culmina el grado, corrigiendo el cálculo al tener en cuenta el número de grados que entran en la operación. Así, un sistema óptimo tendría al final del nivel, si este comprende seis grados y considerando que ningún estudiante repite o deserta, una inversión acumulada de 6.000 años-alumno para lograr mil graduados. Cada graduado, por su parte, acumula seis años de educación, de donde el coeficiente es igual a 1 (valor óptimo). Como en los casos anteriores, la repetición tiene un efecto negativo sobre el coeficiente, ya que la elevación de costos que esta implica (porque un estudiante cursa más de una vez un grado dado) se manifiesta como reducción de eficiencia. De modo análogo, la deserción reduce los costos. Como en el caso del indicador anterior, este efecto de la deserción se hace más intenso mientras más tarde se produce. El modelo trata como «desperdicio» toda inversión en estudiantes que no lleguen a alcanzar un umbral dado (como por ejemplo el final del nivel).

Este coeficiente se calcula de la siguiente manera:

Donde:

$$CE_g = \frac{Prom_{g-1}}{SYac_g / N_g}$$

CE_g : coeficiente de eficiencia al grado g
 $Prom_{g-1}$: cada una de las magnitudes de promovidos desde el grado inmediato anterior ($g - 1$) en el modelo; es decir, el número de supervivientes al grado g
 $SYac_g$: años-alumno acumulados hasta el grado g
 N_g : número de grados acumulados hasta el grado g inclusive.

- **Tasa de término de los estudios.** Siendo posible estimar para una cohorte dada toda la trayectoria educativa, resulta lógico esperar que también se pueda calcular cuántos logran terminar el nivel; pero los modelos de cohortes fueron desarrollados para estimar eficiencia y, por lo mismo, no se planteó que contuvieran un indicador de este tipo. Sin embargo, eso no quiere decir que no sea posible considerarlo. Como se señaló en la nota 118, el cálculo de los indicadores de flujo para el grado final requiere información sobre el número de estudiantes que culminan el último grado y se gradúan del nivel o, a fin de estimar este valor, la matrícula en el grado siguiente (por eso se incluyó un grado 7 anteriormente).

Los temas vinculados a la culminación del último grado y la graduación del nivel serán discutidos más adelante, por lo que no cabe desarrollar esa discusión en este punto.

Sin embargo, sí es importante considerar que estimar los valores que deben aplicarse a la conclusión del último grado a partir de la matrícula en el siguiente revela un problema sustantivo: equiparar la conclusión de un grado con la promoción al siguiente no es una operación exenta de dificultades. Por ejemplo, en un país con una presencia extendida de la educación primaria pero no de la educación secundaria, terminar la primaria no se traduce fácilmente en la prosecución de estudios en el siguiente nivel, ya que esto puede implicar costos de transacción que la población podría no estar en condiciones de sufragar. Por ejemplo, graduarse en una escuela primaria en zonas rurales puede verse acompañado por el hecho de que continuar los estudios secundarios implique migrar a alguna ciudad cercana. Esa falta de identidad entre aprobación del grado y supervivencia en el siguiente es un problema presente en todos los casos, pero que puede hacerse más agudo en el cambio de nivel, dado que esto puede ser más demandante para el estudiante y su familia.

Esta tasa de término se calcula de la siguiente manera:

Donde:

TT_n : tasa de término del nivel n

$Prom_g$: cada una de las magnitudes de promovidos desde el grado final en el modelo; es decir, el número de supervivientes que va al siguiente nivel

Coh : cohorte

x : factor que puede ser igual a 1, 100 o 10.000 (fracción, porcentaje, cada 10.000 habitantes).

$$TT_n = \frac{\sum Prom_g}{Coh} * x$$

En el ejemplo, estos indicadores tendrían los valores señalados en la tabla 10.

Tabla 10. Indicadores de eficiencia interna interanual (II), por grados, País B, 2012-2011

	Grado 1	Grado 2	Grado 3	Grado 4	Grado 5	Grado 6
Supervivencia		994	970	942	854	725
Deserción acumulada		6	30	58	146	275
Años por alumno	1.219	1.159	1.110	1.046	963	886
Años por alumno acumulados		2.378	3.488	4.534	5.497	6.383
Coefficiente de eficiencia	0,815	0,815	0,810	0,753	0,660	0,464
Término						494

Notas

- Todos los valores están expresados por mil estudiantes, excepto el coeficiente de eficiencia, que es una fracción de uno.
- Valores hipotéticos presentados con fines de ilustración.

Hasta aquí se ha presentado el modelo de eficiencia interna interanual tal y como opera en los modelos de cohortes reconstruidas. Más allá de tal presentación, se ha llamado la atención sobre dos elementos potencialmente críticos. Antes de continuar con esta revisión del alcance del modelo y sus supuestos, es preciso subrayar algo que está en la base de la significativa presencia que este modelo ha tenido en las últimas décadas: permite decir mucho (calcular una batería amplia de indicadores) a partir de un conjunto muy pequeño de datos (solo los 20 puntos de datos contenidos en la tabla 7 [página 152]; esta tabla ha sido usada para calcular los indicadores contenidos en la tabla 10).

Ciertamente, este modelo parte de un pequeñísimo conjunto de datos de fácil acopio. El esfuerzo intelectual imbricado en hacer decir muchas cosas de modo coherente a un conjunto tan reducido de información debe destacarse como una gran ilustración de perspicacia e ingenio en el diseño de este modelo.

Sin embargo, la validez de tal modelo puede ser seriamente discutida considerando diversos elementos. En la presentación anterior se han anticipado dos de ellos:

- La preocupación por la eficiencia (que es completamente legítima y necesaria para cualquier reflexión sobre el estado de la educación) se ha desarrollado a expensas de criterios sustantivos sobre el fenómeno educativo. No resulta adecuado **considerar que desertar es más deseable que repetir** o, peor aún, que desertar temprano es mejor que hacerlo luego (habiendo tenido ciertas oportunidades educativas), si es que no se va a concluir el nivel. Este implícito del modelo es inaceptable para cualquiera que vea a la educación como un espacio para ampliar las oportunidades de las personas.
- De igual forma, la identidad implícita en el modelo entre completar un grado y promoverse al siguiente es potencialmente muy riesgosa, ya que oculta **posibles problemas vinculados a costos de transacción del acceso a los servicios educativos**. Tales problemas son más notorios en el paso de un nivel a otro, pero pueden ser igualmente complejos dentro de un mismo nivel (por ejemplo para personas que estudian en pequeñas escuelas rurales donde solo se ofrece un número limitado de grados).

En adición a estos temas, existe otro conjunto de elementos tanto o más complejo que compromete la validez de los modelos de cohortes:

- El modelo asume que el sistema educativo (o la unidad que se esté observando) constituye un sistema cerrado. Es decir, **no existe espacio en el modelo para transacciones dinámicas con otros componentes del sistema**. Por ejemplo, traslados de un subconjunto a otro (entre áreas urbanas y rurales, entre programas educativos de gestión estatal y no estatal, etc.) distorsionan el modelo de una manera incontrolable. Si, por ejemplo, existen escuelas rurales incompletas y los estudiantes tienen que continuar sus estudios (digamos en cuarto grado) en una localidad urbana, esto aparecerá como deserción para las zonas rurales y como reducción de la deserción para las zonas urbanas en el modelo. Así, el modelo solo debería ser usado para dar cuenta de conjuntos respecto de los cuales exista certeza de que las transacciones son de una magnitud tan pequeña que pueden ser obviadas. Este puede ser el caso de los agregados nacionales totales y por sexo, pero el modelo difícilmente sirve en otros casos. Así, si las migraciones internacionales son importantes (y lo son crecientemente en el mundo contemporáneo), el modelo deja de funcionar adecuadamente incluso para los propios agregados nacionales.

- Dado que el modelo se inspiró en las llamadas «tablas de vida» usadas en demografía, hay un tratamiento problemático de la deserción. **Toda deserción es tratada como definitiva**, es decir, como si se tratase de una defunción. Los desertores, para el modelo, nunca vuelven al sistema y dado que en la realidad sí lo hacen, el modelo termina viéndose afectado de una manera no determinable. Como se ha dicho anteriormente, la estadística no busca tener valores «exactos», sino dimensionar la inexactitud para poder manejarla. En este caso, tenemos la presencia de un error no identificable que cuestiona la validez del modelo por la imposibilidad de manejarlo.
- **Las tasas usadas para el modelo son estables en el tiempo e iguales para todos los sujetos.** Lo primero puede tener cierta solución usando información real de años consecutivos¹²¹, mientras que lo segundo es imposible de sostener. Existe suficiente argumento conceptual y evidencia empírica para saber que, por ejemplo, la probabilidad de deserción varía con la edad de las personas y el número de repeticiones, de donde aplicar las mismas tasas en cada iteración resulta problemático, nuevamente, en una magnitud desconocida. Corregir este problema puede suponer un esfuerzo enorme que podría no valer la pena si es que se desarrollan métodos alternativos al modelo de cohortes.
- **El modelo solo presta atención a la matrícula** cuando el derecho a la educación es algo que toca a la población y no solo al subconjunto de esta que se encuentra matriculado.

Como se señaló anteriormente, la confiabilidad de las estimaciones hechas con modelos depende directamente de: (i) la capacidad del modelo mismo para representar la realidad; (ii) la validez de los supuestos usados; y (iii) la calidad de la información usada.

Aquí hemos podido identificar problemas serios relativos a los dos primeros elementos. El tercero, por su parte, también agrega limitaciones que es necesario ponderar.

Por ejemplo, si los sesgos en el reporte de información afectan de modo diferenciado los diversos componentes de la información de partida, todo el esfuerzo termina completamente distorsionado. Así, ¿qué pasa cuando la agregación de la información sobre repetición y promoción lleva a que las tasas de deserción sean menores a cero?¹²², ¿las llamaremos «tasas de inserción»?¹²³

¹²¹ El valor de esta solución es dependiente de problemas de consistencia de datos que mencionaremos más adelante al referirnos a los problemas originados en su calidad.

¹²² Las tasas de deserción se calculan por diferencia. Por lo tanto, si la matrícula reportada como repetición de un grado (g en $t + 1$) más la estimada (o reportada si fuera el caso) como promoción al siguiente ($g + 1$ en $t + 1$) da como resultado una suma mayor a la matrícula original en el grado de referencia (g en t), la deserción resultará arrojando valores negativos.

¹²³ Esta es la solución que alguna vez escuché a un consultor muy original, que prefería esto a tener que renunciar a un modelo de estimación de inversiones que reposaba íntegramente en proyecciones de matrícula basadas en tasas de eficiencia interna interanual construidas de la manera aquí descrita. Por cierto,

Del mismo modo, si no se controla adecuadamente las tasas de respuesta para cada período, variaciones aparentes en los volúmenes de matrícula terminan afectando todo el modelo, ya que este se basa en un conjunto muy pequeño de datos y, si hay problemas ahí, estos se amplifican en la operación general del modelo.

Asimismo, es preciso considerar que estos modelos fueron construidos en otro contexto (con mucha menor disponibilidad de datos) y con otra agenda política (la preocupación mayúscula por la inversión educativa en un contexto de matrícula en expansión).

Por supuesto que todo lo hasta aquí discutido o, mejor dicho, los reparos puestos a los modelos de flujos y a los indicadores que de ellos se derivan son de naturaleza conceptual o lógica. Así, cabe dar una mirada a los datos reales y ver si lo hasta aquí descrito se traduce en problemas empíricos (tabla 11).

Tabla 11. Tasa de supervivencia al grado final de la educación primaria, países seleccionados, 1999-2007

País	1999	2000	2001	2002	2003	2004	2005	2006	2007
Bolivia	79,8	74,4	82,3	78,8	82,4			80,2	83,7
Brasil		79,9	84,4	83,5	80,5	75,6			
Chile	99,5			98,7	98,4			94,9	
Costa Rica	87,7	90,5	88,4	89,1	89,5	83,6	90,5	83,8	94,3
Ecuador	74,6	75,7	76,2	72,1	72,6	75,9	75,6	80,6	81,4
El Salvador	62,5	64,1	65,3	71,3	69,6	66,4	67,3	68,9	75,7
Guatemala	51,7	50,3	60,4	58,9	65,2	62,6	63,3	62,4	64,7
Honduras						54,0	80,9		76,2
México	87,2	87,3	89,0	91,3	90,4	91,6	92,1	92,3	91,5
Nicaragua	45,9	51,8	63,9	51,7	54,8	50,8	50,2	43,8	48,4
Panamá	90,1	86,2	88,0	79,5	81,0	82,4	85,2	88,2	85,2
Paraguay	73,3	72,6	63,6	76,5	76,5	84,0	76,7	79,1	79,2
Perú	83,0	80,9	78,1		84,1	85,5	84,9	89,7	83,0
República Dominicana	70,6	72,1	67,0	61,5	53,9	82,9	61,1		
Uruguay		86,6	92,1	87,0	89,6	93,8	92,4	93,7	94,3

Fuente: Unesco UIS (<<http://www.uis.unesco.org>>); elaboración propia.

Notas

- La educación primaria es definida de acuerdo a la adaptación de cada sistema nacional a la CINE.
- Los datos en cursiva son estimaciones del UIS. Las celdas en blanco corresponden a valores no disponibles.

Si un x por ciento de los ingresantes alcanza el grado final, solo estos se encuentran en condiciones de concluir el nivel; así, si asumimos que todos los que alcanzan el grado

el archivo electrónico en el que el consultor hacía todos los cálculos había sido creado y posteriormente mejorado por funcionarios de los ministerios de diversos países en los que el consultor había estado presente.

final lo concluyen y que toda la población ingresa, entonces los valores de la tasa de supervivencia al grado final de la educación primaria establecen el límite superior posible a la proporción de personas que culminan dicho nivel.

De esta forma, si los modelos de cohortes reconstruidas funcionan bien, estos valores de supervivencia deberían ser iguales o mayores a los niveles de conclusión que se observan en los países.

Esta comparación es posible, sin embargo, no está exenta de dificultades. La proporción de personas que ha logrado al menos la educación primaria completa es una medida directa e inesgada de la conclusión, pero es solo observable en edades superiores a la edad mínima posible de concluir, lo que plantea un desajuste temporal entre las tasas de supervivencia **actuales** y la conclusión de aquellos que transitaron por el sistema educativo en el pasado. Este desajuste es menor en la medida en que se considere grupos poblacionales cuyo tránsito por el sistema ha sido reciente, así como si se tiene en cuenta que los niveles de conclusión han tendido a incrementarse en el tiempo (así, los grupos poblacionales de edades mayores muestran niveles de conclusión menores que los más jóvenes).

Para los países considerados en la tabla 11, ha sido posible consultar información sobre niveles de conclusión para la población de quince a veinticuatro años de edad, con la que se ha elaborado la tabla 12¹²⁴.

Como se puede observar, en prácticamente todos los países (excepto Chile y Costa Rica) el último valor disponible en la serie de conclusión es sistemáticamente mayor que lo observable respecto de la supervivencia en el cuadro anterior. Así, cabe preguntarse, por ejemplo, ¿cómo es posible que el 95,1% de la población de quince a veinticuatro años en 2009 en Panamá o el 91,2% en 2009 en Paraguay tengan al menos primaria completa cuando las tasas de supervivencia que suponen haber experimentado durante su tránsito por la primaria (de 1999 a 2007) solo promedian valores de 85,1 y 75,7% respectivamente?

Estos valores empíricos no son consistentes entre sí y algo similar se observa en los otros casos (excepto Costa Rica y Chile, como ya se anotó). Esta inconsistencia solo es explicable si: (i) la tasa de supervivencia está subestimada, sea por problemas de datos o por falencias del modelo; (ii) la información sobre conclusión está sobrestimada; o (iii) ambas cosas a la vez.

Una primera indicación sobre la calidad de la información de las encuestas de hogares está dada por su consistencia temporal. En efecto, si se observa: (i) las tasas en la serie; (ii) el comportamiento de las tasas para distintos grupos etáreos en una encuesta singular¹²⁵;

¹²⁴ Dado que esta información suele generarse mediante encuestas de hogares, es necesario considerar franjas etáreas y no edades simples, ya que estas encuestas no están diseñadas para reportar información por edades simples. Nótese que aquellos que tienen en 2009 entre quince y veinticuatro años típicamente cursaron su educación primaria entre los años 1999 y 2007.

¹²⁵ Un ejercicio de este tipo se encuentra presente en Unesco Santiago (2004).

y (iii) el comportamiento de (i) y (ii) a través de los distintos países; entonces es posible encontrar un comportamiento consistente y acorde a lo esperable (los niveles de conclusión han aumentado a lo largo del tiempo) que sugiere que estos valores son confiables.

Tabla 12. Personas de quince a veinticuatro años de edad con al menos educación primaria completa, países seleccionados, 1990-2009 (en porcentajes)

	1990	1991	1994	1996	1998	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Bolivia						83,9		83,7			87,8	82,9	87,1		
Brasil							88,7			92,0		90,8	91,5	92,0	92,1
Chile	92,5		94,8	95,6	95,8	96,6			97,6			98,4			98,5
Costa Rica		87,8				89,5		90,1			92,9		92,0	92,8	93,6
Ecuador			94,9				90,6					92,5	96,3	94,0	94,5
El Salvador					68,7	72,5	73,9	74,7	74,5	75,2		78,4		81,4	
Guatemala							49,1					60,8			
Honduras	56,2						70,7			70,1		75,2	75,5		
México				87,1	88,0	90,0		91,3		92,1		93,7		94,5	
Nicaragua					57,8		63,8				71,4				
Panamá				92,7						94,4		93,6	94,0	94,7	95,1
Paraguay						80,6						88,4	90,2	89,2	91,2
Perú													93,6		
República Dominicana													85,0	84,2	
Uruguay							97,0	96,7	96,6	96,6	96,8	96,6	96,5	96,5	

Fuente: Sistema de Información de Tendencias Educativas de América Latina (Siteal) de Unesco y OEI (<<http://www.siteal.iipe-oei.org>>; recuperados en julio de 2012); elaboración propia.

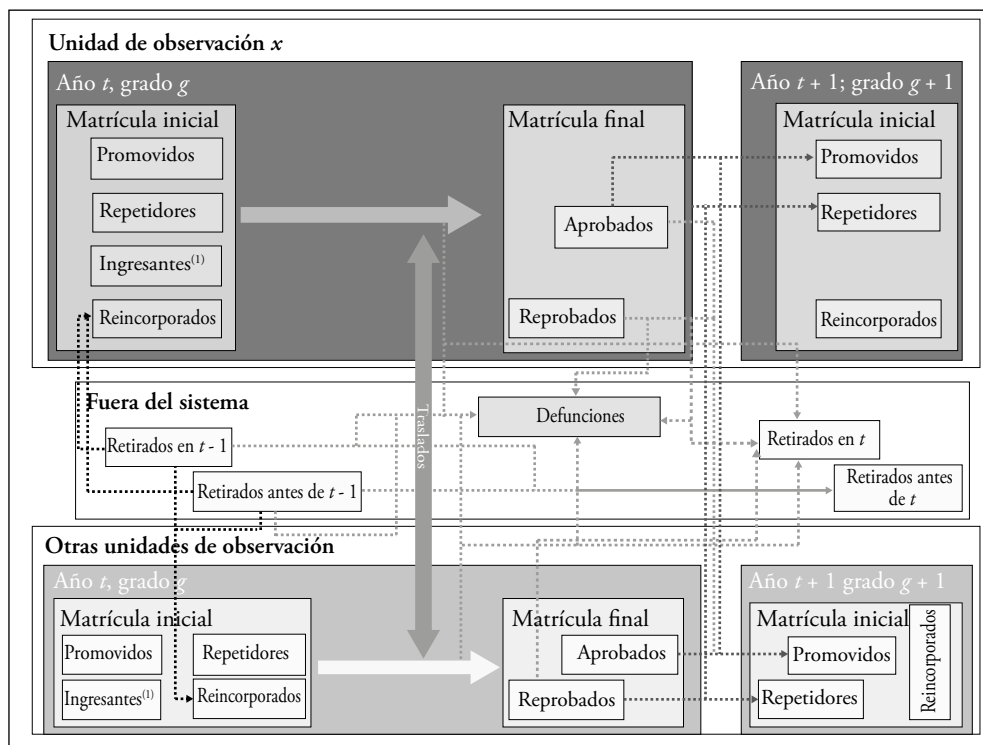
Notas

- Estos datos provienen de encuestas de hogares procesadas por el Siteal. El indicador publicado por Siteal corresponde al porcentaje de personas que tiene primaria incompleta o menos, aquí se publica el complemento de dicha tasa. No se cuenta con información acerca de cómo se ha definido «primaria» en cada caso (confiamos en que, al tratarse de un programa de la Unesco, se haya usado la CINE, pero esto no está documentado), ni acerca de los márgenes de error muestral de las estimaciones.
- Las celdas en blanco corresponden a valores no disponibles.
- No se ha proporcionado los errores muestrales por no encontrarse en la fuente tomada.

Por su parte, las observaciones anteriormente señaladas respecto de los modelos de cohortes reconstruidas sugieren que estos podrían no estar generando información lo suficientemente robusta.

El problema de fondo es que los flujos reales no están bien representados ni por el modelo ni por sus supuestos. Los flujos reales se parecen más al esquema del gráfico 12 (y no al que sustenta los modelos de cohortes del gráfico 11 en la página 151):

Gráfico 12. Representación de los flujos de matrícula



Fuente: adaptado de Guadalupe *et al.* (2002: 19)

Notas

⁽¹⁾ Ingresantes: solo si se trata del primer grado. (Cómo definir el «primer» grado también es algo que merece cierta atención. Una primera opción es «primer grado del nivel», con esta definición se pierde parte de la dinámica entre niveles. Una segunda opción es «primer grado en el sistema» y esto plantea la pregunta de si se trata solo de la fracción obligatoria de este o si se incluye niveles no obligatorios, como es el caso usual del preescolar).

El gráfico 12 cumple un doble propósito: por una parte, recuerda la complejidad de las dinámicas y trayectorias escolares y, por otra, es suficiente para ilustrar que los supuestos de los modelos de cohortes reconstruidas sobresimplifican la realidad que quieren representar.

Además, el gráfico 12 busca recuperar para la discusión algunos elementos clave que son obviados por los supuestos de los modelos de cohortes reconstruidas. En primer lugar, la necesidad de no considerar al sistema escolar como uno cerrado, donde los abandonos son una especie de *wormhole*¹²⁶ que simplemente arroja a los estudiantes fuera del universo observable. En segundo lugar, abrir el sistema deriva en una dinámica

¹²⁶ Se trata de una característica hipotética del espacio-tiempo que aparece como un puente Einstein-Rosen entre dos puntos del universo (o entre universos, en el tipo Schwarzschild) que de otra forma serían muy distantes.

compleja de transacciones (donde solo las defunciones reales son definitivas) con procesos temporales que también lo son.

Evidentemente, no es necesariamente obligatorio que los sistemas de información den cuenta de toda esta complejidad y siempre se necesita una solución parsimoniosa. Pero la discusión aquí desarrollada muestra que los modelos de cohortes reconstruidas y los indicadores de él derivados no son dicha solución, ya que no son lo suficientemente buenos como para dar cuenta de los temas clave de una manera metodológicamente robusta y consistente con las necesidades de la política educativa.

Entonces quedan al menos dos opciones: (i) invertir en un sistema de seguimiento de cohortes reales; y/o (ii) aproximarnos a estos problemas de una manera completamente diferente, que no tenga las dificultades de los modelos de cohortes.

Sin duda alguna, la primera opción parece ser muy atractiva para quienes trabajan en sistemas de información, pero demanda un esfuerzo muy grande y puede tener implicancias muy serias en cuanto al respeto a la privacidad de las personas. En cualquier caso, subsanar los problemas de los indicadores aquí discutidos no es un argumento suficiente para embarcarse en una empresa de esa magnitud y carácter. Dicha empresa solo tiene sentido en la medida en que pueda hacer una contribución a que las personas ejerzan su derecho a la educación y eso es algo que habría que argumentar muy sólidamente.

Por lo tanto, «no tenemos más opción» que ser creativos y buscar alternativas que parsimoniosamente nos permitan abordar algunos problemas clave. Para esto, es imprescindible volver a la pregunta de partida: ¿Qué se necesita saber con relación a la eficiencia del sistema?, y no estancarse en la pregunta: ¿Cómo arreglar los modelos de cohortes?¹²⁷

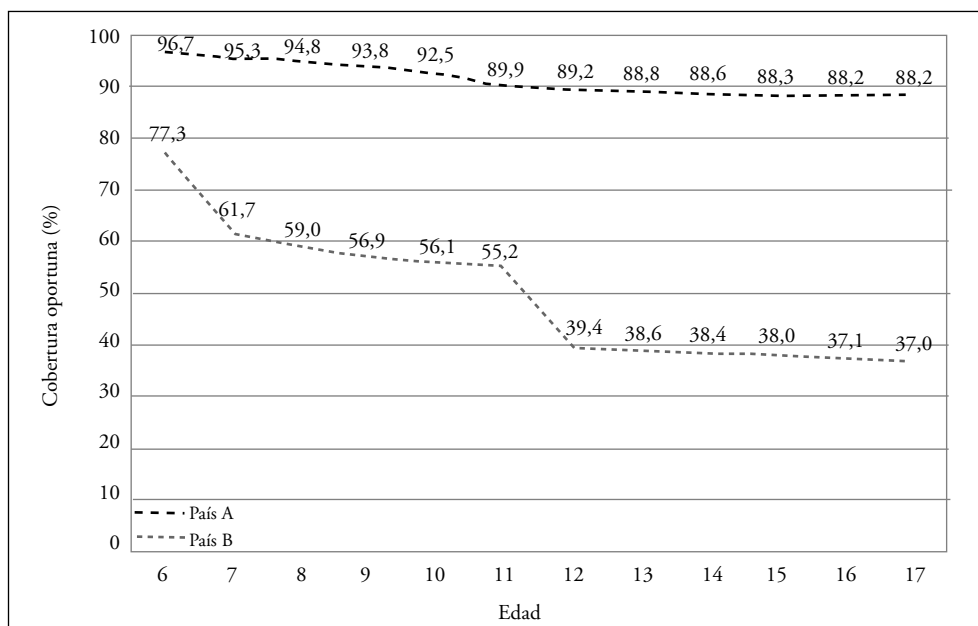
Observar la dinámica del sistema (aquello para lo que fueron desarrollados los modelos de cohortes) se podría lograr dando cuenta de modo sistemático de la estructura de la matrícula por edades para un período dado de tiempo¹²⁸. Es decir, si asumimos (como lo hacen los modelos de cohortes) que la dinámica del sistema no varía dramáticamente en un período limitado de tiempo, incluso la información de matrícula por edades de un solo año o ciclo escolar nos podría dar una imagen inicial de la dinámica del sistema. Esta imagen puede mejorarse usando una serie de tiempo, siempre y cuando se tomen en cuenta las posibles inconsistencias en dicha serie.

¹²⁷ Schiefelbein y Klein han hecho importantes esfuerzos en estos terrenos —véase Wolff *et al.* (2002) para una explicación de los procedimientos desarrollados por estos autores, así como Schiefelbein y Grossi (1984)—. Sin embargo, los modelos siguen siendo similares, aunque los controles de la información con la que se les alimenta sean mejores.

¹²⁸ Recuérdese que los modelos de cohortes reconstruidas usan datos de solo dos años consecutivos, aunque dada su mecánica de iteraciones parecieran estar referidos a un período temporal mayor.

En la sección destinada a la cobertura, se mencionó un indicador que mide la **cobertura oportuna**. Ahora bien, a efectos de ver la dinámica del sistema, es posible percibir los cambios en dicho indicador a través de las distintas edades simples (en una sola toma de datos o en una serie, en este último caso, observándose varias cohortes poblacionales simultánea pero diferenciadamente). El gráfico 13 ilustra este punto.

Gráfico 13. Cobertura oportuna, seis a diecisiete años, Países A y B, 2012 (en porcentajes)



Nota
Valores hipotéticos presentados con fines de ilustración.

En efecto, como se puede observar, el País A ha logrado niveles de acceso oportuno a su primer grado prácticamente universales (96,7%), aunque se aprecia una pérdida de dicha oportunidad a lo largo de los primeros seis años (el indicador cae a 89,9% para la edad de once años, es decir experimenta una caída de $[96,7 - 89,9] / 96,7 = 7,1\%$) y luego se mantiene muy estable. Así, para la cohorte de diecisiete años, el País A ha sido capaz de conservar el $88,2 / 96,7 = 91,2\%$ de la cobertura oportuna de la edad inicial.

Por su parte, el País B muestra una menor capacidad para lograr acceso oportuno (77,3%) y, sobretodo, para conservar dicho nivel, especialmente luego de un primer año de estudios (una caída de 20,2% en el valor del indicador) y entre las edades de once y doce años (una caída de 28,7% en el valor del indicador). Así, para la edad final considerada, el País B solo es capaz de conservar el 47,8% de la cobertura oportuna que había logrado en la edad de acceso al primer grado de primaria.

Este deterioro en los niveles de cobertura oportuna, que hemos medido como el cociente entre el indicador para la edad final versus la edad inicial, al que podemos llamar **coeficiente de conservación de la cobertura oportuna**¹²⁹, es un indicador directo de los grados de eficiencia del sistema educativo respecto de su capacidad para conservar su situación de partida a un grado dado respecto de una variable que mide de modo directo la calidad del flujo de las trayectorias escolares. Una trayectoria escolar eficiente es la que se da sin generar atrasos ni deserciones y es exactamente esto lo que mide el coeficiente aquí reseñado.

Formalmente, el grado de deterioro es medido por el coeficiente que compara el valor de la cobertura en edad para cualquier grado con el de la cobertura en edad en el grado inicial (se divide el primero entre el segundo), de manera tal que un valor 0 indica una pérdida total y un valor 1 una conservación total del punto de partida. En términos formales se calcula de la siguiente manera:

$$CCCO_g = \frac{CO_g}{CO_{gi}}$$

Donde:
 $CCCO_g$: coeficiente de conservación de la cobertura oportuna en el grado g
 CO_g : cobertura oportuna en el grado g
 CO_{gi} : cobertura oportuna en el grado inicial gi .

Como se aprecia en el gráfico 13, en el País B en el segundo grado el sistema solo había sido capaz de conservar 0,798 de la cobertura oportuna inicial. Es decir, solo el 79,8% de los que se encontraban matriculados en condición oportuna en el grado inicial logra mantenerse en esa situación en el siguiente ciclo lectivo. Del mismo modo, al tercer grado solo se conserva la cobertura oportuna a 0,763 del valor inicial y al décimo grado esta cae hasta solo 0,491 del valor en el grado inicial.

Asumiendo que estos valores de un solo año son indicativos de la dinámica, tal coeficiente muestra de modo directo el nivel total de eficiencia o ineficiencia de un sistema dado su punto de partida. Por cierto, esto debería verse complementado con una revisión de cuán eficaz es el sistema en lograr un punto de partida adecuado¹³⁰ (en el caso del País A dicha capacidad es clara, en el caso del País B es menor).

Adicionalmente, es posible calcular a qué edades se producen los mayores niveles de caída del coeficiente y explicar esa caída de acuerdo a sus componentes (la deserción se deriva de la cobertura total para la edad).

De modo resumido, es posible calcular la cobertura en edad para las edades inicial y final, sea de un nivel o del rango de escolarización obligatoria. Esto daría, por una parte,

¹²⁹ Este indicador fue usado en Unesco Santiago (2008: 121 y ss.) para medir el deterioro entre los grados inicial y final de la educación primaria.

¹³⁰ Sobre este punto, véase el acápite 2.4.4 sobre ingreso al sistema.

una imagen clara del ingreso oportuno, así como, por otra, del deterioro de la cobertura en edad, que puede ser representada por el coeficiente propuesto.

Por último, contar con información de matrícula por edades y grados para dos o más años consecutivos puede permitir una mejor reconstrucción de la dinámica real del sistema. Teniendo esta información, en la observación de una serie de datos se apreciaría la dinámica real de las cohortes de población (y no de matrícula). Así, por ejemplo, un cambio en la cobertura total en dos años consecutivos para edades consecutivas (digamos ocho años de edad en el año t y nueve años en $t + 1$) mostraría desertores de esa edad, si el cambio neto es negativo, o reingresos¹³¹ (cosa imposible de observarse con los modelos de cohortes) de esa edad, si el cambio neto es positivo.

Como se puede advertir, sistemas de información robustos que cuentan con un buen registro de las edades de los estudiantes pueden perfectamente prescindir de los modelos de cohortes para el análisis de su dinámica temporal.

Ahora bien, el enfoque propuesto ayuda con relación a la dinámica del sistema, pero no resuelve el problema de dar cuenta del «desperdicio» en términos de inversión. Para esto, sin embargo, los modelos de cohortes tampoco son necesarios, debido a que dicho «desperdicio» se hace real en el hecho de que las personas deban matricularse más de una vez en un mismo grado, es decir, matricularse en condición de repetidores. Una buena aproximación a esto está dada por el volumen relativo de la población matriculada en condición de repetidora, indicador computable con la información generalmente disponible. Este cálculo tiene la siguiente forma:

Donde:

FR_g : proporción de repetidores en el grado g o en un rango de grados (nivel)

$$FR_g = \frac{Rp_g}{Mat_g} * x$$

Rp_g : número de repetidores en el grado g o en un rango de grados (nivel)

Mat_g : matrícula en el grado g o en un rango de grados (nivel)

x : factor que puede ser igual a 1, 100 o 10.000 (fracción, porcentaje, cada 10.000 habitantes).

Nótese que la fracción de la matrícula de estudiantes en condición de repetidores no es idéntica a la tasa de repetición, ya que el numerador es la matrícula del mismo grado en el mismo período y no la del período precedente. Esto permite que el cálculo pueda hacerse para cualquier unidad de observación o desagregación que se requiera (cosa imposible con los indicadores interanuales, ya que no toman en consideración el efecto de los traslados entre dos períodos lectivos).

¹³¹ Si es que es posible asumir que las migraciones son inexistentes o tienen un efecto insignificante.

Este indicador permite estimar que, para un año dado, la proporción de la inversión necesaria como efecto de la repetición (es decir, desperdiciada por la incapacidad del sistema para asegurar un flujo óptimo de los estudiantes) es igual a una fracción calculada de la inversión total en un grado o nivel dado.

Otra forma de abordar el tema de la eficiencia tiene que ver también con la dinámica de la matrícula, pero, nuevamente, sin recurrir al modelo de cohortes. Es posible estimar, dada la estructura de la matrícula, un indicador postulado por la Unesco hace mucho tiempo, llamado **esperanza de vida escolar** (este indicador calcula el número de ciclos o años que se espera que una persona esté matriculada en el sistema—incluyendo las repeticiones—) y comparar dicho valor con el **grado que se espera que las personas aprueben** en un determinado momento¹³². Mientras más próximos ambos valores, menores las repeticiones y, por lo mismo, mayor la eficiencia interna.

Veamos cada indicador por separado y luego su uso conjunto.

La **esperanza de vida escolar** es un indicador que da cuenta de la cantidad de años que en promedio se espera que una cohorte poblacional se encuentre matriculada en el sistema educativo. Así, en un sistema con comportamiento «óptimo» y al que se ingresa a los seis años de edad al primer grado, el indicador mostraría que la población de siete años tendrá una vida escolar de un año, la de ocho dos y así sucesivamente, ya que en dicho sistema toda la población ingresa oportunamente y se mantiene así a lo largo de toda su trayectoria. En cambio, en un sistema con deserciones, cada año de edad implicará una magnitud de años de estudio adicionales menor que un año en proporción a dicha deserción. Del mismo modo, si hay ingresos tardíos, la población de siete años logrará, en promedio, encontrarse en el sistema menos de un año, donde la diferencia con el año deseable será proporcional al ingreso tardío.

La repetición no afecta a este indicador, pues solo contabiliza la cantidad de años cursados y no los grados aprobados.

El comportamiento general del sistema en un momento dado es leído como «**esperanza**» de vida escolar, ya que el valor se entiende de la siguiente manera: para aquellos en edad de ingresar (digamos seis años, para seguir con el ejemplo), se espera que tengan una vida escolar igual a x años (hasta que alcancen la edad e), la misma que se deriva de la observación transversal de los actualmente matriculados.

A efectos de calcular la esperanza de vida escolar, simplemente se estima las tasas de cobertura total por edades simples (expresadas como fracciones de uno, no como porcentajes) para el rango de edades considerado (a partir de la edad de ingreso hasta el punto de corte que se establezca, que puede ser cualquier edad) y se suma dichos valores. En términos formalizados se tiene:

¹³² Este nuevo indicador y su uso concurrente con la esperanza de vida escolar se documentan en Guadalupe y Taccari (2004).

$$SLE_c = \sum TC_{c+1-e}$$

Donde:
 SLE_c : esperanza de vida escolar para la población de la edad c
 TC_{c+1-e} : tasa de cobertura total para las edades $c + 1$ a e .

El ejemplo de la tabla 13 permite ilustrar el cálculo.

Tabla 13. Cobertura total y cálculo de la esperanza de vida escolar, por edades, País A, 2011

Edad	Tasa de cobertura total	Esperanza de vida escolar	
		Contribución de la edad e	Esperanza de vida escolar
6	0,967	0,967	0,967
7	0,991	0,991	1,958
8	0,995	0,995	2,953
9	0,988	0,988	3,941
10	0,982	0,982	4,923
...
65	0,001	0,001	...

Nota

Valores hipotéticos presentados con fines de ilustración.

Como se puede apreciar, en los casos en que la cobertura total es cercana a 1, cada año de edad representa un año más de educación y, por lo mismo, el valor de la esperanza de vida escolar se incrementa en un valor próximo a uno; mientras tanto, para las edades donde la cobertura total se distancia de 1, cada año de edad contribuye a elevar la esperanza de vida escolar en una fracción menor.

Ahora bien, como ya se anotó, la esperanza de vida escolar solo contabiliza los períodos lectivos anuales cursados (que se espera que se cursen) y no los grados aprobados. Una comparación entre estos dos valores daría una muestra de la eficiencia del sistema: si los grados aprobados fueran iguales a los años cursados, se tendría un sistema eficiente; por su parte, si los grados aprobados fueran menores a los años cursados, entonces se tendría una medida de la ineficiencia (la distancia entre lo esperado y lo logrado).

El problema entonces se convierte en: ¿Cómo generar información sobre los grados aprobados de las personas? Una primera opción sería recurrir a los censos o encuestas de hogares, pero esta tiene varias limitaciones. En adición a los temas mencionados anteriormente acerca de los problemas en los períodos de referencia para el registro de las edades (que es clave en un tema como este), la información sobre grados aprobados que se puede obtener de estas fuentes refiere a lo que las personas han obtenido en diversos momentos en el tiempo y no al desempeño actual del sistema. De esta forma, dicha información no resulta conceptualmente la más apropiada para la operación aquí

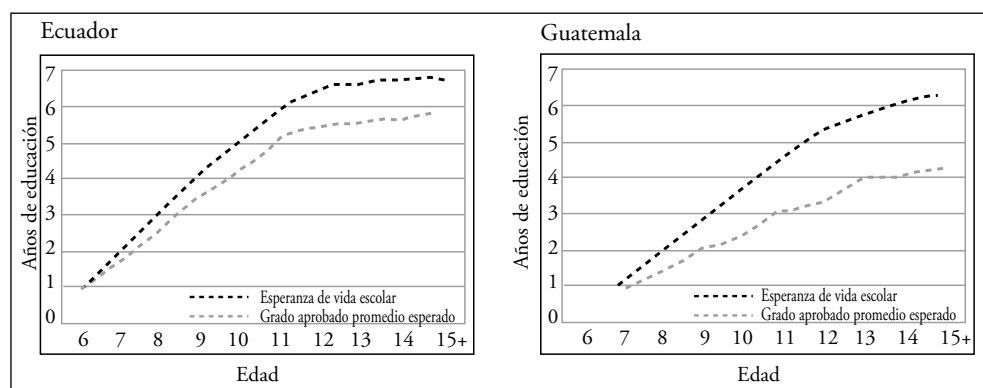
planteada, pues está destinada a medir la eficiencia actual del sistema, por lo que sería claramente inconsistente con la información de la esperanza de vida escolar que se utilizaría.

Por lo tanto, el desafío es construir un indicador sobre el número de grados que en promedio se espera que las personas aprueben dada su edad a partir de información de la estructura actual de la matrícula por edades.

En efecto, este procedimiento fue desarrollado por la Unesco Santiago –véase Guadalupe y Taccari (2004)–, denominándose a este nuevo indicador: **grado aprobado promedio esperado (GAPE)**.

El cálculo de este indicador se basa en el uso de una matriz de matrícula por edades simples y en el cómputo para cada edad singular del promedio de años aprobados. Este promedio se calcula considerando que las personas matriculadas en el segundo grado ya aprobaron un grado, las matriculadas en el tercero dos grados y así sucesivamente. Es decir, se calcula el número promedio de grados en los que la cohorte poblacional de cada edad singular se encuentra matriculada y se le resta uno. Los ejemplos del gráfico 14 ilustran el procedimiento.

Gráfico 14. Esperanza de vida escolar y grado aprobado promedio esperado, Ecuador y Guatemala, 2000



Fuente: Guadalupe y Taccari (2004: 12).

Nota

Nótese que la edad oficial de inicio de la educación primaria en Guatemala (en 2000) es siete años, por lo que no se ha registrado valores para los seis años de edad.

El gráfico 14 muestra tanto las diferencias que se originan en la cobertura (a fin de cuentas la esperanza de vida escolar es solo la agregación de las tasas de cobertura total por edad) como en la eficiencia. Así, cuando luego de tres años en la escuela los estudiantes ecuatorianos lograban en promedio aprobar cerca de tres grados, en Guatemala este valor solo era próximo a dos. La observación conjunta de ambos valores muestra una brecha y cómo esta evoluciona en el tiempo.

Tal brecha entre ambos indicadores señala claramente problemas de eficiencia originados en la repetición. La deserción no afecta la brecha, ya que esta siempre toma en cuenta que el mismo volumen de matrícula se usa en ambos indicadores.

Esta brecha de eficiencia se puede cuantificar de la siguiente manera:

Donde:

$$BE_e = \left(1 - \frac{GAPE_e}{SLE_e} \right) * x$$

BE_e : brecha de eficiencia en la edad e
 $GAPE_e$: grado aprobado promedio esperado a la edad e
 SLE_e : esperanza de vida escolar a la edad e
 x : factor que puede ser igual a 1 o 100 (fracción, porcentaje).

Mientras que su evolución se puede representar mediante la tasa a la que se amplía la brecha entre dos edades (consecutivas o no), la misma que se calcula de la siguiente manera:

Donde:

$$EBE_{e-f} = \left(1 - \frac{BE_f}{BE_e} \right) * x$$

EBE_{e-f} : evolución de la brecha de eficiencia entre las edades e y f siendo $f > e$
 BE_f : brecha de eficiencia en la edad f
 BE_e : brecha de eficiencia en la edad e
 x : factor que puede ser igual a 1 o 100 (fracción, porcentaje).

Como se puede apreciar, existe una batería de indicadores actualmente disponibles que permite dar cuenta de la eficiencia interna **sin necesidad de recurrir a los modelos de flujos interanuales**, los cuales, si bien hicieron una contribución de primera importancia durante más de tres décadas, hoy resultan demasiado problemáticos e innecesarios, dados los niveles actuales de disponibilidad de información.

Eficiencia externa

Un terreno adicional y completamente diferente del anterior, refiere al análisis de lo que muchas veces es llamado eficiencia «externa» del sistema educativo. En este caso se analiza los impactos que el sistema educativo tiene más allá de su esfera y, en ese sentido, con frecuencia aparece identificado como análisis de los impactos y no como análisis de eficiencia (ya que, en realidad, no es un tema de eficiencia sino de eficacia).

En efecto, la educación tiene impactos de corto, mediano y largo plazo en la vida de las personas, en sus familias y en las colectividades (locales, nacionales y global) a las

que pertenecen. Así, el análisis de los impactos de la educación cubre un universo muy amplio de temas, que incluyen¹³³:

- Aspectos demográficos, como:
 - Fecundidad y natalidad
 - Salud y mortalidad infantil
 - Nutrición infantil
- Aspectos sociales, entre ellos:
 - Movilidad educativa intergeneracional
 - El propio impacto de la escolaridad de los padres en los logros educativos de sus hijos
 - Identificación de variables que condicionan el impacto de la escolaridad
- Aspectos vinculados a la vida institucional democrática
- Aspectos vinculados a factores económicos, como:
 - Crecimiento económico
 - Rendimientos económicos de las inversiones en educación (tasas de retorno)
 - Ocupaciones de las personas según su experiencia educativa
 - Empleabilidad, subempleo y desempleo
 - Ajuste entre la experiencia educativa y la ocupación
 - Umbrales educativos para evitar la condición de pobreza
 - Modernización económica

Nótese que, dada la forma como se dan estos impactos en el tiempo, dichos análisis no dan cuenta del desempeño **presente** del sistema educativo sino más bien de sus impactos observables en un momento en el tiempo, de acuerdo al desempeño pasado del mismo. Por esta razón, este tipo de análisis no se suele incluir cuando el foco de la preocupación es el desempeño presente, que es lo que mayormente ocupa a los agentes de la política educativa.

No vamos a detenernos en una presentación detallada de los indicadores de eficiencia externa. Para el lector interesado, se sugiere que consulte el ya mencionado trabajo de Muñoz-Izquierdo *et al.* (2004). Baste aquí con señalar que estos análisis se centran en explorar asociaciones entre variables educativas (usualmente el número de años de

¹³³ Basado en Muñoz Izquierdo *et al.* (2004). Este volumen incluye una abundante revisión bibliográfica sobre el tema así como una propuesta de indicadores que pueden ser usados para abordarlo. Cabe destacar también la reflexión puramente económica de Walter McMahon, presente en McMahon y Boediono (1992) y en McMahon (1997, 1998).

escolaridad, la certificación obtenida o, de contarse con información, el nivel de destrezas adquirido)¹³⁴ con variables de otros dominios, como los mencionados.

Por ejemplo, si se quiere estudiar el impacto de la educación en la salud de las personas, normalmente se busca verificar la asociación entre variables educativas de las personas y/o sus hogares con variables de salud. Al encontrarse dichas asociaciones, se postula que, por ejemplo, la evidencia respalda la afirmación de que un mayor nivel de escolaridad de las personas contribuye a una reducción de la mortalidad infantil.

Ahora bien, en análisis como estos usualmente se considera a grupos poblacionales de diversas edades que han vivido sus experiencias educativas en diferentes momentos en el tiempo, por lo que estos impactos tienen que ver más con el desempeño de largo plazo de los sistemas educativos.

Lo anterior no implica que este tipo de análisis deba desdeñarse, sino que debe situarse en su espacio específico como parte de análisis más globales y de un horizonte temporal mayor. Por esta razón, suelen ser investigaciones académicas las que abordan este tipo de aspectos, aunque de ello no debe colegirse que no tienen impactos importantes en la política educativa. Por ejemplo, la política del Banco Mundial que estuvo a la base de muchos esfuerzos de reforma educativa en los años 1990 se basó justamente en un análisis de los retornos económicos de la inversión en educación, el cual llevó a concluir que había que centrarse prácticamente de modo exclusivo en la educación primaria, por ser este nivel el que mostraba mayores tasas de retorno¹³⁵.

¹³⁴ Esta es la principal motivación de estudios como el IALS, conducido en algunos países de la OECD en los años 1990, y sus secuelas, el Adult Literacy and Lifeskills Survey (ALL), conducido a inicios de los 2000, y el actual PIAAC. Estos estudios generaron información sobre las habilidades de las personas adultas en base a pruebas estandarizadas y luego vincularon sus resultados a otras variables sociales. Véase, entre otros documentos: OECD y Canadá-Statistics Canada (2000, 2005); OECD (2013b). Para una muy relevante crítica del enfoque seguido por la OECD en estos estudios (lo que también se aplica a PISA), véase, entre otros, Darville (1999); Hamilton (2001, 2012); Hamilton y Barton (2000); St. Clair (2012); Guadalupe (2015).

¹³⁵ El principal estudio sobre tasas de retorno que sustentó la definición de la política del Banco Mundial es Psacharopoulos y Woodhall (1985). Véase World Bank e Independent Evaluation Group (2006) para una evaluación de la política del banco, que ahora muestra una preocupación mayor por los logros de aprendizaje y no solo por la escolaridad por sí misma, ya que lo que explicaría los retornos son las habilidades de las personas y no las certificaciones. En cualquier caso, el cambio de énfasis en la preocupación no modifica el razonamiento subyacente. El resumen ejecutivo del informe se inicia con la siguiente frase, que ilustra los resultados de la evaluación y el ajuste de la política: «Los conocimientos y destrezas básicos –no el nivel educativo alcanzado– son fundamentales para reducir la pobreza. Expandir la matrícula y los niveles de conclusión de la educación primaria es necesario, pero no suficiente, para garantizar un nivel elemental de competencias de alfabetismo y uso de números» (World Bank e Independent Evaluation Group 2006: xiii). La traducción es propia, ya que la disponible en la versión en español del resumen ejecutivo no es satisfactoria (El texto original dice: «Basic knowledge and skills –not educational attainment– are key to reducing poverty. Raising enrollments and completing primary schooling are necessary –but not sufficient– to ensure basic literacy and numeracy»).

En muchas ocasiones, uno de los usos más importantes de estos análisis es contribuir con evidencia a la causa a favor de una mayor atención a la educación, por ejemplo mediante mayores niveles de inversión. Así, se señala que invertir en educación es rentable por los beneficios que esta trae a las personas, sus familias y colectividades. En ese sentido es un análisis de los retornos a dicha inversión.

Si bien este argumento es razonable y muchas veces logra un importante nivel de predicamento entre las personas, no se debe olvidar que la educación es un derecho humano fundamental y, por lo mismo, la inversión en garantizarlo se justifica por sí misma. Llevando al extremo el argumento a efectos de ilustrarlo cabalmente, se podría decir: **si la educación no tuviese ningún impacto en el crecimiento económico de un país, igual hay que invertir en ella, pues las personas tienen derecho a recibirla.**

Adicionalmente un análisis de la rentabilidad debería tomar en cuenta también los costos implicados en no invertir en educación. Entre los beneficios de hacerlo, habría que contabilizar el evitar los perjuicios de no hacerlo. O, para decirlo con la famosa frase atribuida a Derek Bok¹³⁶: «Si usted cree que la educación es cara, pruebe con la ignorancia».

Finalmente, dado que las variables usadas para medir el elemento educación en el discurso de la eficiencia externa corresponden al volumen de educación alcanzada (máxima certificación lograda) o de habilidades logradas (por ejemplo, la condición de alfabetismo), estos dos temas tienden a ser tratados bajo el rubro «impactos». Sin embargo, si bien pueden ser tomados en cuenta para el análisis de la «eficiencia externa», son fenómenos educativos por sí mismos y, por lo tanto, son tratados como tales en los siguientes acápite.

2.4.3 NIVEL DE ESCOLARIDAD DE LA POBLACIÓN

La medición de los niveles de escolaridad logrados por la población es una práctica extendida que se basa en registrar, a partir de censos poblacionales o encuestas de hogares, la información declarada por las personas con relación a su experiencia educativa.

Esta información suele ser generada a partir del registro de la respuesta a una pregunta del tipo: «¿Cuál es el último grado educativo que usted ha aprobado?». La respuesta a esta pregunta es usualmente registrada como el último grado aprobado y luego convertida en categorías que representan: (i) el número de años de educación acumulados hasta dicho grado; (ii) el último nivel educativo cursado diferenciando entre niveles culminados y no culminados.

Tal información es posteriormente tratada bajo dos modalidades: (i) como una variable continua, número de años de educación logrados, respecto de la cual es posible

¹³⁶ Educador y abogado norteamericano, fue rector de la Universidad de Harvard entre 1971 y 1990.

computar promedios, desviaciones estándar, etc.; y/o (ii) como una variable ordinal que permite dar cuenta de la distribución de la población según categorías de experiencia educativa lograda (por niveles o número de años). Adicionalmente, es posible utilizar esta información no solo para ver comportamientos típicos (como promedios), sino también para observar la dispersión que esta variable presenta¹³⁷.

Dichos valores son usualmente calculados para distintos grupos etéreos (edades simples o grupos de edad), así como para otras desagregaciones posibles, dadas las características de la fuente de datos.

Esta información es de mucha importancia para conocer diferencias entre grupos de población, dinámicas temporales (esto se detallará más adelante en el acápite sobre la conclusión de estudios), brechas a ser atendidas por el sistema (adultos que, por ejemplo, no han completado la educación primaria o los niveles que se consideran obligatorios), así como para análisis que vinculan esta variable a otros factores, tal y como se describió en el acápite anterior en las referencias sobre los impactos de mediano y largo plazo en otras esferas de la vida social («eficiencia externa»).

2.4.4 INGRESO

Un tema no siempre tratado con el suficiente detenimiento se refiere a los niveles de «ingreso» al sistema educativo. Usualmente, se ha medido el ingreso como la entrada al primer grado de primaria a partir de asumir que la educación preprimaria es una suerte de preescuela o preeducación, lo que deriva de una forma anticuada que desdeña la importancia que los procesos educativos tienen justamente en las edades tempranas¹³⁸.

Así, y como ha resultado usual en varias de las discusiones planteadas en este texto, es preciso definir con claridad y diferenciar algunos fenómenos que se suelen confundir en la discusión, como los siguientes.

(i) Acceso vs. ingreso

En este texto se viene utilizando de modo diferenciado estos dos términos, de modo que el primero hace referencia a las posibilidades de las personas de acceder a los servicios educativos que les corresponden (en cualquier momento en el tiempo), mientras que el segundo refiere al momento en que la persona llega «por primera vez» al sistema educativo.

¹³⁷ Véase el acápite 2.1 del anexo 1 para el detalle de las medidas de tendencia central y dispersión.

¹³⁸ Por cierto, hoy existe un bagaje muy grande de investigación que muestra que justamente los momentos más críticos para el desarrollo de las personas se concentran en los primeros años, motivo por el cual los esfuerzos educativos sistemáticos en las edades tempranas son de capital importancia. Para una discusión amplia sobre el tema, véase Unesco (2006); por su parte, la perspectiva de la OECD, que curiosamente destaca la importancia de la educación infantil debido a la incorporación de la mujer al mercado laboral, se encuentra disponible en OECD (2001, 2006b).

De esta forma, la cobertura trata específicamente de los temas de acceso (el texto asume una equivalencia entre las siguientes dos preguntas: ¿Qué proporción de personas accede al sistema? ¿Qué proporción de personas es atendida o servida por el sistema?), mientras que el ingreso busca registrar el momento y la proporción de personas que son captadas «por primera vez» por el sistema.

(ii) **Ingreso al sistema vs. ingreso al nivel**

A efectos de capturar el ingreso, es posible registrar en qué momentos en el tiempo y en qué puntos del sistema educativo se capta por vez primera a las personas. Por ejemplo, para una cohorte de población dada se puede tener lo mostrado en la tabla 14.

Tabla 14. Acceso al sistema educativo de los nacidos en el año 2000, 2001-2008

Año	Preescolar (por edades)						Primaria (grado 1)
	0	1	2	3	4	5	
2000	α						
2001		β					
2002			γ				
2003				δ			
2004					ϵ		
2005						Ω	ϕ
2006							λ
2007							μ
2008							π

Nota

En este caso la tabla termina en 2008, pues luego de esa fecha ningún nacido en 2000 entra al sistema en el ejemplo.

A partir de una tabla como esta, sería posible (y contando con la información poblacional correspondiente) calcular qué porcentaje de la población nacida en 2000 ha logrado acceder al sistema educativo.

Este cálculo tiene la siguiente forma:

Donde:

TI_n : tasa de ingreso de la población nacida en el año n

Ing_{n+i}^{ei} : proporción de personas de la edad ei que ingresa al sistema en el año $n + i$

Pob_{n+i}^{ei} : población de la edad ei en el año $n + i$

x : factor que puede ser igual a 1, 100 o 10.000 (fracción, porcentaje, cada 10.000 habitantes).

$$TI_n = \left(\sum \frac{Ing_{n+i}^{ei}}{Pob_{n+i}^{ei}} \right) * x$$

Este procedimiento equivale a calcular lo que representa cada celda con información de la tabla 14 como proporción de la población correspondiente como fracción de la población de 0 años en 2000, 1 en 2001, 2 en 2002, 3 en 2003, 4 en 2004, 5 en 2005, 6 en 2006, 7 en 2007 y 8 en 2008: celda α de entrantes de edad 0 en 2000... celda Ω de entrantes de edad 5 en 2005, celda ϕ de entrantes al primer grado de edad 5 en 2005... y celda π de entrantes al primer grado de edad 8 en 2008, como fracción de la población de 0 años en 2000, ... y 8 en 2008; y luego agregar cada una de esas proporciones de modo que se tenga el valor para la cohorte en total considerando el efecto de posibles defunciones. Este procedimiento da cuenta del ingreso al sistema independientemente del nivel educativo en el que este se registra (en el ejemplo, se captura el ingreso tanto en el preescolar como en el primer grado de primaria)¹³⁹.

Si, por otra parte, solo se requiere registrar el ingreso a un nivel educativo dado – por ejemplo educación primaria en este caso–, únicamente habría que considerar en la tabla 14 las celdas ϕ , λ , μ , π . Nótese también que estas celdas darían cuenta del ingreso oportuno (λ)¹⁴⁰, del ingreso tardío (μ , π) y del ingreso precoz (ϕ) al nivel. Así, tendríamos:

Donde:

TI_l^n : tasa de ingreso al nivel l de la población nacida en el año n

$$TI_l^n = \left(\sum \frac{Ing_{n+i}^{ei}}{Pob_{n+i}^{ei}} \right) * x$$

Ing_{n+i}^{ei} : proporción de personas de la edad ei que ingresa al nivel l en el año $n + i$

Pob_{n+i}^{ei} : población de la edad ei en el año $n + i$

$$TIO_l^n = \left(\frac{Ing_{eo}}{Pob_{eo}} \right) * x$$

TIO_l^n : tasa de ingreso oportuno al nivel l de la población nacida en el año n

Ing_{eo} : ingresantes de la edad oficial de ingreso eo

Pob_{eo} : población de la edad oficial de ingreso eo

$$TIP_l^n = \left(\frac{Ing_{n+i-}^{eo}}{Pob_{eo}} \right) * x$$

TIP_l^n : tasa de ingreso precoz al nivel l de la población nacida en el año n

Ing_{n+i-}^{eo} : población de la edad oficial de ingreso eo , que ingresó en un año anterior al esperado

¹³⁹ Es importante tener presente que el ingreso «por vez primera» se puede dar en cualquier momento en el tiempo, ya que población migrante puede ingresar por primera vez a la unidad bajo observación en cualquier punto del ciclo educativo. A efectos de no complejizar innecesariamente el ejemplo, se presenta solo situaciones de ingreso en el preescolar y en el primer grado de educación primaria.

¹⁴⁰ Asumiendo que la edad de ingreso normada al primer grado es seis años.

$$TIL_l^n = \left(\frac{Ing_{n+i+}^{eo}}{Pob_{eo}} \right) * x$$

TIL_l^n : tasa de ingreso tardío al nivel l de la población nacida en el año n

Ing_{n+i+}^{eo} : población de la edad oficial de ingreso eo , que ingresó en un año posterior al esperado

x : factor que puede ser igual a 1, 100 o 10.000 (fracción, porcentaje, cada 10.000 habitantes).

Como se puede apreciar, la tasa de ingreso oportuno aquí descrita es equivalente a la llamada **tasa neta de ingreso**, usualmente calculada solo para el primer grado de primaria. En este caso, sin embargo, se plantea un uso para cualquier nivel, utilizando un nombre que es más claro en términos de su significado y presentando el indicador en un marco específico de tratamiento del fenómeno del ingreso donde su alcance es más claro.

Esto, por su parte, también muestra que la actual **tasa neta de ingreso** es un indicador consistente y robusto, aunque insuficiente para dar cuenta del fenómeno del acceso.

(iii) Volúmenes vs. proporciones

Finalmente, en el tema del ingreso también es posible encontrar indicadores de volumen e indicadores de proporciones, como se discutió en el caso de la cobertura.

El indicador de volumen usualmente calculado es la **ratio bruta o aparente de ingreso**. Esta ratio, como sus contrapartes, las ratios brutas de matrícula, da cuenta del volumen relativo de población que ingresa al sistema educativo (en realidad, al primer grado de educación primaria, según la forma usual de cálculo) y, por lo mismo, sus alcances y límites son idénticos a los ya discutidos con relación a las ratios brutas de matrícula.

Reste aquí añadir que la ratio bruta de ingreso no es equivalente a la tasa de ingreso anteriormente planteada (ni tampoco a la suma de las tasas de ingreso oportuno, precoz y tardío), ya que la primera toma como punto de referencia un año dado (independientemente de la población), mientras que la segunda tiene como referencia una población dada, lo que no hace sino reiterar las preguntas acerca de cuál es el foco del análisis que se desea hacer: volúmenes de matrícula a gestionar en un momento dado o el grado en el que una cohorte poblacional ejerce su derecho a acceder al sistema educativo.

2.4.5 CONCLUSIÓN DE UN NIVEL EDUCATIVO

Un tema cada vez más importante se vincula a la medición de la conclusión de los estudios. La importancia del tema se explica por la creciente conciencia acerca de que

el acceso a un nivel educativo no se traduce necesariamente en concluir los estudios, de modo que se plantean objetivos específicos vinculados a, por ejemplo, la conclusión universal de la educación primaria, la baja secundaria, etc.

Los sistemas de indicadores tradicionalmente usados no cuentan con indicadores relativos a este tema. La presencia de la conclusión universal en la agenda de política y en compromisos internacionales como el marco de acción de la EPT ha contribuido a poner sobre el tapete la necesidad de desarrollar indicadores en este terreno.

¿Cómo medir la conclusión? Una medida adecuada supone, como en cualquier otro caso y en primer lugar, definir con precisión el problema. Así, la pregunta inicial implica determinar qué ha de entenderse por conclusión universal de, por ejemplo, la educación primaria. Esta interrogante parece trivial, pero se puede encontrar situaciones en las que se discute sobre los indicadores sin que haya un acuerdo mínimo sobre el significado de lo que se quiere medir.

En cierta oportunidad, en una reunión de trabajo con diversos funcionarios de un ministerio y de otras organizaciones en un país de América Latina me tocó discutir este tema. A dicho efecto, recuerdo que empecé planteando la pregunta acerca del significado de la meta sobre la universalización de la conclusión de la educación primaria y las respuestas que se dieron incluían lo siguiente:

- **Todas las personas**, independientemente de su edad, deben haber logrado concluir al menos la educación primaria.
- **Todos los estudiantes** de educación primaria deben terminar dichos estudios.
- **Todos los niños que hoy están en edad de ingresar** a la escuela deben hacerlo y terminar sus estudios primarios.
- **Todos los niños** (personas de hasta dieciocho años de edad) deben terminar sus estudios primarios.
- **Todas las personas en edades de escolarización propias de la educación primaria** deben terminar los estudios de dicho nivel.

Como se puede apreciar, cada una de estas respuestas implica algo diferente en términos del grupo poblacional al que refieren (se ha destacado en negrita cada caso). Así, la construcción de un indicador que mida la conclusión universal tendría la siguiente forma:

$$TC_n = \frac{PC_n}{Pob_r} * x$$

Donde:

TC_n : tasa de conclusión del nivel n

PC_n : el número de personas en la población de referencia que concluyen el nivel n

Pob_r : población de referencia

x : factor que puede ser igual a 1, 100 o 10.000 (fracción, porcentaje, cada 10.000 habitantes).

Pero, en cada caso, el denominador de la fórmula es diferente y, por consiguiente, el numerador también debería serlo para ser consistente con la población de referencia.

A esto hay que añadir que medir la conclusión efectiva de un nivel supone que las personas tengan una edad suficiente para haberlo logrado. Es decir, no tiene ningún sentido medir, por ejemplo, cuántos niños de cinco años han terminado la educación primaria, pues ninguno está en condiciones de haberlo hecho. De hecho, lo que cabría en este caso es estimar cuántos **terminarían** dados determinados supuestos. Es decir, la conclusión puede **medirse** de modo **retrospectivo** (cuántos han terminado, dado que ya tuvieron o debieron tener la oportunidad de cursar los estudios) o **estimarse** de modo **prospectivo** (cuántos habrían de terminar en el futuro).

Empecemos por la parte más fácil de este problema. Medir retrospectivamente es relativamente más fácil y puede hacerse de modo directo. Los censos de población y las encuestas de hogares suelen preguntar a las personas por su experiencia educativa y, como ya se ha anotado en un acápite previo (ver la página 43), al hacerlo generan información sobre el último grado o nivel aprobado o concluido.

Así, una primera medida directa de la conclusión está dada por **la proporción de personas que ha concluido el nivel objeto de estudio**. Esta tasa de conclusión se calcula de la manera descrita en la fórmula indicada líneas arriba. Esto, como es lógico suponer, solo debe calcularse para la población que tiene por lo menos una edad en la que se supone que debería haber concluido el nivel educativo motivo de análisis. Esta es la aproximación usada en Unesco Santiago (2004) y una de las usadas en Unesco Santiago (2008)¹⁴¹. La misma permite contar con una medida retrospectiva de la conclusión, siempre y cuando se acote a una población de referencia de al menos una determinada edad.

Una de las ventajas más importantes de este procedimiento radica en la fuente de datos que debe ser usada (dado que refiere a población no cubierta en los eventos estadísticos conducidos desde el propio sistema educativo). Al ser información generada regularmente mediante los censos de población o las encuestas de hogares, es posible calcular el indicador no solo para la población total de una edad x o mayor, sino también para desagregaciones por edades simples o rangos de edad (dependiendo de la capacidad de las muestras para permitir estimaciones de los valores poblacionales con un razonable margen de error), sexo, área de residencia, pertenencia étnica u otras desagregaciones posibles.

Ahora bien, existen dos temas claves que deben considerarse al querer usar este indicador: (i) definir el grupo de edades y (ii) su carácter retrospectivo.

¹⁴¹ Este procedimiento se encuentra completamente documentado y discutido en Guadalupe y Louzano (2003) y en Guadalupe y Taccari (2004). En el marco de los borradores sobre metas e indicadores de desarrollo sostenible, se ha introducido este indicador con el nombre «tasa de logro de la educación primaria» –«primary education attainment rate»– (Post-2015 Education Indicators Technical Advisory Group of the EFA Steering Committee 2014: 9).

¿Cuál es la edad mínima a considerar para definir la población a la que este indicador debe referir? La respuesta a esta pregunta implica considerar los siguientes elementos: en primer lugar, cuál es la edad mínima posible para haber culminado la educación primaria (o el nivel educativo que se esté considerando) en las condiciones propias del país; y, en segundo lugar, el hecho de que existe atraso escolar y, por lo mismo, resulta esperable que las personas concluyan el nivel no con la edad mínima esperada, sino en ese momento o después.

Por ejemplo, en un sistema educativo con una educación primaria de seis grados y edad de ingreso a la misma oficialmente estipulada como seis años, las personas deben tener al menos doce años para culminar la primaria. Así, caben las siguientes opciones:

- Medir la conclusión para aquellos que tienen doce años de edad solamente. Este sería un indicador de **culminación oportuna**. Sin duda es un indicador clave a utilizar. Sin embargo, antes de hacerlo es necesario preguntarse si la fuente de datos permite hacer dicha estimación para una edad simple. Si se trata de una encuesta por muestreo, el error muestral de una estimación para una edad singular puede ser de un tamaño tal que torne no recomendable hacerlo de esta manera, ya que las encuestas de hogares no suelen estar diseñadas para brindar información a ese nivel de desagregación (edad simple) con niveles de error manejables¹⁴². Si se trata de un censo poblacional esto es perfectamente posible, aunque aquí se depende de la frecuencia de realización de los censos. Igualmente, si los censos escolares producen información sobre graduación del nivel por edades simples, este indicador puede ser fácilmente calculado.
- Medir la conclusión para las personas de doce, trece, catorce, quince y más años de edad. Así se podría contar con información sobre las personas que terminan oportunamente (población de doce años); en edad y hasta con un año de atraso (trece); en edad y hasta con dos años de atraso (catorce); o en edad y hasta con tres años de atraso (quince). Nuevamente, la factibilidad de proceder de esta manera es dependiente de la fuente de datos: limitada en encuestas por muestreo, posible si se trabaja con censos.

Si se observa los patrones de conclusión, es posible encontrar que esta es creciente a partir de la edad mínima de egreso hasta que se estabiliza en edades un poco mayores. Así, si lo que interesa es medir cuántas personas concluyen, resulta conveniente elegir como punto de partida una edad a partir de la cual los incrementos de la misma no se

¹⁴² Por esta razón es muy importante que las estimaciones muestrales sean siempre publicadas con las correspondientes estimaciones de error que permitan manejar la información con rigurosidad.

traduzcan en incrementos en la conclusión (es decir, la edad en la que se maximiza la probabilidad de conclusión)¹⁴³, por ejemplo, a partir de los quince años.

El carácter retrospectivo de este indicador plantea un problema diferente. El indicador mide resultados del desempeño pasado del sistema educativo, cuando usualmente la gestión está más preocupada del desempeño presente. Entonces, ¿cómo estimar la conclusión de modo prospectivo?

Hay básicamente dos opciones para esto: proyectar a partir de la evolución histórica o estimar con comportamientos actuales.

La proporción de personas que ha culminado la primaria, según fue descrita en los párrafos precedentes, tiene una virtud importante: si se observa la información por edades simples o por rangos de edad (digamos grupos quinquenales), se tiene una buena aproximación a la evolución histórica de este indicador, ya que lo más probable es que aquellos que hoy tienen, digamos, 60 años de edad, culminaron, o no, su educación primaria hace por lo menos 45 años, mientras que poblaciones más jóvenes lo hicieron en momentos más recientes.

De esta manera, la observación de este indicador por edades (simples o grupos) permite, a pesar de basarse en una sola toma de datos, aproximarse a la dinámica temporal y brinda las bases para proyectar la tendencia observada, de modo que se pueden construir valores esperados para la población que cursa actualmente la educación primaria o que habría de cursar dichos estudios en el futuro cercano.

En la tercera sección se tratarán en detalle los temas vinculados a las proyecciones. Baste aquí con señalar dos cuestiones de capital importancia: (i) cualesquiera que sean los procedimientos que se usen para hacer una proyección, es preciso que estén claramente documentados, de manera que los usuarios puedan replicar los procedimientos y discutir los supuestos; y (ii) por muy buena que sea una proyección, siempre será dependiente de la experiencia pasada, de modo que políticas recientes que afecten de modo sustantivo el fenómeno observado (positiva o negativamente) no serán tomadas en cuenta y esto suele desmerecer el valor de las proyecciones, ya que, muchas veces, los agentes de política suelen considerar que «lo que se hace hoy es completamente distinto de lo visto en el pasado» y «nuestra gestión no puede juzgarse en base a lo que hicieron otros en el pasado». Por lo tanto, si bien hacer buenas proyecciones es importante, resulta necesario pensar en otros procedimientos complementarios que ayuden a tener una media prospectiva basada en información sobre el comportamiento **actual** del sistema educativo.

Existen en la actualidad varias opciones que tratan de atender esta preocupación. Revisémoslas en detalle.

¹⁴³ Esto se vincula a que es muy probable que a partir de determinada edad las personas que no han concluido simplemente abandonen (la probabilidad de abandono se incrementa con la edad y la propia dinámica temporal).

Los indicadores producidos por el UIS-Unesco¹⁴⁴

El UIS produjo cuatro indicadores vinculados al tema de la conclusión. Si bien todos ellos están referidos exclusivamente a la educación primaria (ya que fueron creados como respuesta a la necesidad de información vinculada a la segunda meta de la EPT), no es necesario que se limiten a dicho nivel. Nótese que actualmente el UIS solo publica los dos primeros.

Tales indicadores son los siguientes:

- **Ratio bruta de ingreso al grado final de primaria.** Da cuenta del volumen de personas que actualmente ingresan al último grado de la educación primaria expresado en términos relativos a la población de la edad de referencia para cursar dicho grado.

Esta ratio se calcula de la siguiente manera:

$$RBI_{gf} = \frac{Mat_{gf} - R_{gf}}{Pob_r} * 100$$

Donde:
RBI_{gf}: ratio bruta de ingreso al grado final
Mat_{gf}: matrícula en el grado final
R_{gf}: repetidores en el grado final
Pob_r: población de referencia (población que teóricamente debería estar matriculada en el grado final).

- **Ratio bruta de graduación de educación primaria.** Da cuenta del volumen de personas que actualmente culminan la educación primaria expresado en términos relativos a la población de la edad de referencia para dicha culminación.

Esta ratio se calcula de la siguiente manera:

$$RBG = \frac{Grad}{Pob_r} * 100$$

Donde:
RBG: ratio bruta de graduación
Grad: número total de graduados del nivel en un año dado
Pob_r: población de referencia (población que teóricamente debería estar graduándose en el año para el que se calcula el indicador).

- **Ratio bruta esperada de ingreso al grado final de primaria.** Proporciona una estimación del volumen de población que se espera ingrese al último grado dados los actuales volúmenes de ingreso y patrones de matrícula.

¹⁴⁴ Véase Unesco-UIS (2009a).

Esta ratio se calcula de la siguiente manera:

Donde:

$RBEI_{gf}$: ratio bruta esperada de ingreso al grado final

$$RBEI_{gf} = GIR * SUR^{t-1}$$

GIR : ratio bruta o aparente de ingreso al primer grado del nivel

SUR^{t-1} : tasa de supervivencia al grado final en el ciclo lectivo precedente.

Nótese que la ratio bruta o aparente de ingreso al primer grado de primaria es uno de los indicadores de volumen calculados tradicionalmente por el UIS y equivale al cociente entre el número total de ingresantes al primer grado en un año dado sobre la población en edad de ingreso a dicho grado. Para el detalle, consultar la guía técnica de indicadores del UIS (Unesco-UIS 2010b).

Por su parte, la tasa de supervivencia al grado final es la calculada mediante el modelo de cohortes reconstruidas ya discutido.

- **Ratio bruta esperada de graduación de educación primaria.** Proporciona una estimación del volumen de población que se espera culmine el último grado de la educación primaria dados los actuales volúmenes de ingreso y patrones de matrícula.

Esta ratio se calcula de la siguiente manera:

Donde:

$RBEG$: ratio bruta esperada de graduación

$$RBEG = RBEI * \frac{G}{E}$$

$RBEI$: ratio bruta esperada de ingreso al grado final

G : graduados del grado final

E : ingresantes al grado final

Estos cuatro indicadores comparten, como sus propios nombres denotan, el tratarse de medidas de volumen (ratios brutas), por lo que si bien pueden dar alguna información que toca a las necesidades de inversión no informan sobre el tema central que ocupa la agenda educativa que originó su creación: ¿Se está garantizando la conclusión universal de la educación primaria? Esta pregunta, como debe resultar evidente a partir de lo ya discutido en esta sección, más lo detallado anteriormente al referirnos a la cobertura (y los límites de las ratios brutas), no puede responderse con información sobre volúmenes, sino con información sobre proporciones. ¿Qué fracción de la población culmina la educación primaria? Si la respuesta es 100%, entonces se verifica la universalidad del fenómeno, de lo contrario, no.

Si bien los cuatro indicadores del UIS están bien definidos, eso no es suficiente para asegurar que brinden información suficiente y que lo hacen de una manera clara y transparente. Por ejemplo, tratar de analizar estos cuatro indicadores es una tarea compleja que puede muy fácilmente llevar a equívocos. Por eso es recomendable repensar todo el tema desde el principio enfocándose en qué es lo que se necesita.

Al mismo tiempo, vale la pena destacar el esfuerzo del UIS por precisar qué significa concluir un nivel educativo dado. En efecto, las normas nacionales sobre el tema son variables: en algunos casos concluir la primaria es equivalente a aprobar el grado final, en otros casos hay requerimientos adicionales como, por ejemplo, aprobar un examen de salida. Es de gran importancia incluir este tema de modo exhaustivo en la definición del indicador que se use, de modo que su significado sea nítido.

La tasa de conclusión del Banco Mundial

En adición a ser una institución financiera, el Banco Mundial ha desarrollado acciones muy intensas en diversas áreas vinculadas al desarrollo. Así, ha tenido una presencia muy importante en el terreno de las políticas educativas otorgando créditos y, en algunos casos, canalizando donaciones destinadas a promover determinadas políticas educativas.

Esta acción usualmente ha descansado en el establecimiento de acuerdos con los países que sancionan una «matriz de condicionalidades» a las que se sujeta toda la operación de crédito¹⁴⁵. Esto puede ser fácilmente criticado como un caso de injerencia en las políticas soberanas de los Estados completamente ajena a los principios de acción humanitaria y de promoción del desarrollo que se han venido sancionando en el pasado reciente (Good Humanitarian Donorship 2003).

Sin embargo, no es el propósito de este texto discutir esos elementos, pero son mencionados en este punto pues las «matrices de condicionalidades» suelen incluir indicadores contra los cuales se juzga el desempeño de los países¹⁴⁶. De hecho, el Banco Mundial tiene su propia unidad a cargo de los temas de estadística educativa (EdStats) y mantiene un sitio web con información educativa. Uno de estos indicadores popularizado en años recientes por el Banco Mundial es la llamada **tasa de conclusión de la educación primaria**¹⁴⁷.

El indicador de conclusión usado por el Banco Mundial es exactamente igual a la primera ratio adoptada por el UIS (ratio bruta de ingreso al grado final de primaria.).

¹⁴⁵ Por cierto, resulta curioso que exista una tendencia a afirmar que las acciones desarrolladas con créditos sean descritas como «esto ha sido hecho con recursos del Banco Mundial». Los créditos se pagan y, consiguientemente, son recursos del propio país facilitados mediante una operación de endeudamiento. Ese tipo de afirmaciones equivale a decir que cuando una persona compra una casa con un crédito hipotecario, en realidad la casa se la dio el banco.

¹⁴⁶ Más allá de estas matrices, el Banco Mundial ha venido desarrollando un ejercicio destinado a establecer «puntos de referencia» (*benchmarks*) en educación.

¹⁴⁷ Véase Bruns *et al.* (2003).

Actualmente, el banco menciona este nombre dado por el UIS como alterno del mismo indicador.

Es importante notar que, más allá de los límites de estas ratios para dar cuenta de lo que se supone que quieren dar cuenta, el nombre usado por el Banco Mundial solo sirve para introducir mayores equívocos en su uso, lo que es potencialmente nefasto tanto para los procesos de toma de decisión como para la credibilidad de los sistemas de información: llamar tasa de conclusión a una ratio bruta de ingreso sugiere que es una medida de lo que interesa, lo que no es el caso.

Es importante anotar que esta forma de usar los indicadores puede originarse tanto en la incompetencia de quienes los definen como en la falta de cuidado de quienes los usan o promueven y también en la tendencia a la sobresimplificación que ha caracterizado ciertas discusiones en el terreno de la política educativa. Por ejemplo, durante las últimas décadas, el Banco Mundial ha enfatizado: (i) la importancia de la educación primaria explicada exclusivamente por las tasas de retorno; (ii) la equivalencia de la «calidad» con medidas de logro según pruebas estandarizadas; (iii) la identificación de la fluidez lectora con un número mágico y éticamente reprobable de palabras decodificadas en un minuto; etc.

La tasa de conclusión actual de Orealc Unesco Santiago

En adición a los esfuerzos anteriormente mencionados en el terreno de la medición **retrospectiva** de la conclusión, la Unesco Santiago también ha desarrollado un esfuerzo con relación a contar con una medida **prospectiva**.

Este esfuerzo se expresó en un documento técnico (Guadalupe y Taccari 2004) y luego en el uso de una **tasa de conclusión actual** (Unesco Santiago 2008). La tasa de conclusión actual es una versión modificada del indicador propuesto en el primer documento mencionado; sin embargo, dicha modificación no ha sido completamente documentada hasta la fecha, lo que impide discutir en detalle sus alcances y limitaciones.

A pesar de ello, cabe hacer una reseña de las ideas centrales detrás de este indicador. Partiendo de la información actual sobre matrícula por edades simples y grados, es posible (como se mencionó en la sección sobre eficiencia interna¹⁴⁸) estimar el número de grados que una cohorte de población en promedio aprueba. Así, por ejemplo, en un sistema «perfecto» con seis años de edad de ingreso a la educación primaria y seis años de duración para esta, las personas de siete años tendrán un grado aprobado, las de ocho dos, y así sucesivamente hasta que las personas de doce tengan seis, lo que equivale a haber culminado la educación primaria. Si el promedio de grados aprobados es igual a la duración del nivel, entonces se tendría una conclusión universal. En sistemas no «perfectos» este cálculo se puede seguir haciendo para todo el rango de edades que tiene matrícula en primaria.

¹⁴⁸ Ver la parte sobre este tema en el acápite 2.4.2, página 147.

Esta idea inicial presenta algunas dificultades a tener en cuenta:

- (i) **Disponibilidad de información sobre aprobación por grado.** Ante la ausencia de esta, es posible estimar que los aprobados serán iguales a los matriculados menos los repetidores (asumiendo que en el año observado el número de repetidores se mantendrá igual al del año precedente –supuesto que es necesario verificar–).
- (ii) **Matrícula precoz.** El hecho de que una fracción de la población ha aprobado un número de grados mayor al esperado de acuerdo a su edad distorsiona el indicador haciendo posible valores superiores al 100%. Truncar el valor calculado es un procedimiento posible aunque discutible, ya que enmascara el problema.
- (iii) En situaciones de modificación (mejora o deterioro de la conclusión) **los patrones actuales de matrícula pueden resultar engañosos**, ya que son el resultado de dinámicas distintas experimentadas por cada cohorte poblacional. Claro está que esto es exactamente lo que siempre se ha hecho con mucho menos control de dichos cambios cuando se usaba los modelos de cohortes, pero el que estos contengan el mismo problema no lo resuelve para el indicador propuesto.
- (iv) Un procedimiento de **cálculo simplificado** –frente al que se detalla en Guadalupe y Taccari (2004)– podría partir por calcular la proporción de personas que aprueban el grado final (o que se matriculan en este menos la proporción de repetidores –con los límites mencionados en (i)–) por cada edad simple y luego agregar estos valores. Así, se podría tener lo que se presenta en la tabla 15.

Tabla 15. Cálculo simplificado de la conclusión actual de primaria, País A, 2011

Edad	Aprobados en el grado final ⁽¹⁾	Población	Proporción
12	192.948	216.840	0,890
13	1.770	216.749	0,050
14	2.328	216.555	0,011
15	1.076	216.297	0,005
16	619	215.919	0,003
Total			0,958

Nota

⁽¹⁾ Asumiendo que aprobar el grado final es equivalente a concluir el nivel. Si esta información no está disponible, se podría usar la matrícula menos los repetidores, si es que se cuenta con evidencia para defender este como un supuesto razonable.

La tabla 15 muestra que 0,89 de la población de doce años aprueba el grado final; es decir, 89,0% de dicha población concluye la primaria oportunamente. Asimismo, muestra que 0,05 de la población de trece años termina la primaria

(con un año de atraso) y así sucesivamente. Si estas proporciones se mantienen en el tiempo, es de esperar que la conclusión actual total sea igual a la suma de las mismas, es decir, 95,8%.

La principal dificultad en este procedimiento radica en el posible efecto del atraso escolar. Si los niveles de atraso han ido cambiando en el tiempo (digamos que la cobertura oportuna en el último grado ha ido mejorando), el indicador puede llegar a sobreestimarse, dado que las proporciones que se observarían en las edades más altas (asumiendo otros factores como constantes) serían mayores a las que se observarían en el grupo en edad de término, ya que este no contaría con similares proporciones de personas matriculadas en atraso. Esto, en algunos casos, se traduciría en contar con un valor del indicador superior a 100% (lo que no tiene sentido) o, peor aún, en una sobreestimación no observable cuando el valor es menor a 100%. Por esta razón, es importante verificar, antes de computar el indicador, cómo ha ido evolucionando la cobertura oportuna para la edad de cursar el último grado en los últimos años, de modo que se cuente con evidencia sobre esta posible sobreestimación¹⁴⁹.

Comentario final y opciones

Medir adecuadamente la conclusión supone entender que esto se puede hacer tanto retrospectiva como prospectivamente. Ambos enfoques son válidos y necesarios para dar cuenta de la situación educativa. Al mismo tiempo, es importante entender que las metas referidas a alcanzar determinadas cotas de conclusión por nivel educativo (sea o no conclusión universal) requieren medir proporciones y no volúmenes.

Finalmente, si bien una medida prospectiva basada en los patrones actuales de matrícula aún no está disponible, los esfuerzos de la Unesco Santiago muestran la dirección que hay que seguir y logran esto justamente porque parten por definir con precisión cuál es la pregunta de política que hay que responder y, por lo mismo, no se pierden en la discusión sobre «usar lo que tenemos» (aferrarse a razonamientos o procedimientos obsoletos –como los que subyacen a la existencia de ratios brutas de matrícula o a forzar indicadores derivados de los caducos modelos de cohortes reconstruidas–) o simplemente en buscar una «solución» incompetente y muy poco responsable.

¹⁴⁹ Contar con una sólida serie de datos sobre cobertura oportuna en el grado final puede permitir hacer correcciones al cálculo estimando las proporciones esperadas de personas que llegarían al grado final con uno, dos, tres o más años de atraso, de modo que solo se tome la proporción de graduados para la población con atraso y se aplique este valor a las poblaciones esperadas.

2.4.6 GASTO O INVERSIÓN Y ESFUERZO

Un tema clave de política es el que está definido por la inversión en educación. Al igual que otros temas ya tratados, este también demanda ciertas discusiones conceptuales y de política previas que permitan aclarar qué es lo que los sistemas de información deberían hacer en este terreno.

Un primer elemento se refiere a los términos que usamos para dar cuenta de este ámbito de la realidad educativa: ¿gasto o inversión? Se trata de una discusión relativamente sencilla vinculada con las connotaciones que ambos términos pueden tener con relación a la valoración que se da a la educación. Por un lado, «gasto» suele tener una connotación potencialmente negativa en el sentido de que lo que se destina a la educación no es algo «productivo», es decir, es una «carga» para el presupuesto y, como tal, puede ser objeto de cuestionamientos acerca de su conveniencia, especialmente si se considera que existen diversas cosas que deben ser atendidas con los recursos disponibles. En ese sentido, «inversión» tiene una connotación más favorable a la educación, ya que sugiere que los recursos que se destinen a ella lograrán retornos favorables. Esto, por su parte, sugiere que cabría preguntarse por las diferentes rentabilidades de diferentes destinos potenciales para los recursos y, eventualmente, la educación podría ser «menos rentable» que otras actividades.

Este texto parte de asumir que la educación es un derecho. Por lo mismo, las sociedades deben destinar los recursos necesarios para el ejercicio del derecho. Si a eso lo queremos llamar gasto o inversión, es irrelevante para el argumento de fondo. Tener derechos cuesta y esos costos hay que sufragarlos. Por ejemplo, organizar un proceso electoral mediante el cual los ciudadanos ejercen su derecho a elegir a sus gobernantes tiene un costo. Así, las sociedades «gastan» en tener elecciones o, si se prefiere, «invierten» en contar con este elemento de la vida democrática. En cualquier caso, ese es un gasto o inversión ineludible que debe hacerse porque la gente tiene derecho a ello, porque es uno de los elementos que hacen una sociedad democrática, y preguntarse por la «rentabilidad» o falta de rentabilidad de un proceso electoral es absurdo. Lo mismo se aplica a garantizar otros derechos civiles o políticos y, del mismo modo, habría que abordar la discusión sobre los derechos sociales, incluido el derecho a la educación.

Ahora bien, ¿quién ha de sufragar qué?

Una primera distinción en este terreno tiene que ver con los agentes que invierten en educación: el Estado o agentes privados, incluyendo a las familias y otros actores posibles.

Sin embargo, es importante tener en cuenta que el Estado no genera riqueza por sí mismo y que los recursos públicos son, en último término, recursos que la comunidad nacional pone en manos de sus órganos de gobierno (principalmente a través de los tributos) para hacer cosas que es mejor que se hagan de esa forma. Todas las correas siempre salen del mismo cuero: las familias y los diferentes agentes económicos.

Así, cuando se dice: «El Estado debe garantizar la educación invirtiendo x cantidad de dinero», en realidad se está diciendo que las sociedades deben garantizar la educación de sus miembros mediante una adecuada inversión en educación, una parte de la cual debe ser canalizada a través de la inversión pública. Los dineros públicos son dineros de las personas y eso no debe olvidarse.

Bajo todas las circunstancias posibles, la inversión en educación es siempre una combinación de recursos públicos y privados. Incluso en el caso más extremo de educación pública universal gratuita es posible encontrar que las familias contribuyen directamente (y no solo mediante los tributos) por lo menos asumiendo algunos costos de transacción implicados en el acceso a los servicios educativos (por ejemplo, el transporte). En este sentido, es muy importante que los sistemas de información sean capaces de dar cuenta de ambas fuentes de inversión en educación, no solo a efectos de obtener una buena descripción de la situación, sino y fundamentalmente por los efectos no deseados que puede tener la combinación de inversión pública y privada en el ejercicio del derecho a la educación.

Por ejemplo, es posible encontrar situaciones (que no son poco comunes) donde la inversión pública es muy pequeña y se distribuye de manera igualitaria entre toda la población. Dado que los recursos públicos son muy limitados, las familias empiezan a contribuir para sufragar algunos gastos que pueden ir desde cuestiones básicas (por ejemplo, pagar las cuentas de servicios como la energía eléctrica y el agua potable) hasta contar en forma particular, algunas familias, con profesores de materias especiales que la comunidad educativa valora (por ejemplo, una lengua extranjera, la introducción de nuevas tecnologías o talleres vocacionales). En este contexto, el servicio que los estudiantes reciben empieza a ser proporcional a la capacidad de gasto de las familias, de modo que justamente aquellos que tienen más recursos obtienen una mejor educación, lo que a fin de cuentas termina haciendo que el sistema educativo contribuya a reproducir las desigualdades sociales preexistentes.

Entonces, ¿qué información generar? Empecemos considerando alguna información usualmente disponible y su alcance.

Gasto como porcentaje del producto interno bruto

Un primer indicador usual está dado por los volúmenes de inversión anual disponibles en un país como porcentaje del total de la riqueza producida en ese año en dicho país (gasto en educación como porcentaje del producto interno bruto [PIB]). Este indicador da cuenta del nivel de **esfuerzo** que un país hace para dotar de recursos a su sistema educativo.

Una opción usual para abordar este tema es establecer un umbral deseable de niveles de inversión frente al cual se comparan los niveles observados. Por ejemplo, se considera que «un país debe gastar en educación al menos un equivalente al 6% de su PIB». Ahora

bien, este umbral tiene varios problemas que es preciso abordar: ¿Por qué 6% y no, digamos, 5,3%? La respuesta a esta pregunta puede tener dos formas: (i) cálculos hechos en el país han permitido estimar esa cifra como lo requerido para el período que va del año x al año y ; o (ii) la Unesco ha establecido esa recomendación.

La primera opción nos obligaría a mirar los estudios que la sustentan y es probable que estos tengan sentido, aunque, como se verá más adelante, indexar los niveles de gasto al PIB no es necesariamente un buen procedimiento.

La segunda opción es mucho más problemática. El recurso a la autoridad como argumento es siempre desdeñable por ser una forma de falacia (*ad verecundiam*) y muestra que quien a este recurre no tiene la menor idea de por qué plantea algo. Más allá de quién sea el que recomienda algo, es bueno entender el porqué de dicha recomendación. Este es un tema de responsabilidad y de seriedad profesional¹⁵⁰.

Más allá del origen de esta «recomendación», un análisis somero de la misma muestra que carece totalmente de sentido por las siguientes razones:

Primero, porque es imposible establecer una cifra común a todos los países (o incluso para diversas poblaciones particulares dentro de un solo país) y válida a lo largo del tiempo, dado que las necesidades de inversión pública en educación dependen, entre otras cosas, de las siguientes variables que tienen comportamientos diferentes para cada país y cada tiempo:

- Tamaño de la población a atender.
- Proporción de la matrícula estatal del sector público.
- Composición de la matrícula estatal por niveles educativos: el costo unitario de la educación terciaria tiende a ser mayor que el de la secundaria y el de esta, mayor que el de la primaria.

¹⁵⁰ En efecto, el documento de la Unesco llamado informe Delors (Unesco 1996: 187) plantea una recomendación de este tipo, pero si uno se detiene a revisar el texto no encuentra ni un solo argumento que lo explique. De hecho, si uno trata de buscar una explicación, lo más probable es que ubique el texto base de este capítulo del informe Delors (Peano 1993) y notará con sorpresa que el texto final se distancia del informe base justo en este punto. El documento de Peano (1993: 9) precisamente insiste en la imposibilidad de encontrar recetas únicas universales debido a la diversidad de situaciones existentes y, sin embargo, el informe Delors plantea esta recomendación sin ninguna explicación. Así, uno podría tratar de rastrear el origen de esta recomendación y encontrará que en el período 1964-1966 la Comisión Kothari en la India planteó este tipo de recomendación «basada en eventos internacionales anteriores», sin que estos se especifiquen. Este es el punto hasta el que he podido rastrear esta recomendación sin poder encontrar nada que se parezca a un argumento, en ningún caso. Con bastante generosidad, uno puede asumir que algo próximo a una línea argumentativa es la afirmación subyacente de que si los países de la OECD gastan, en promedio, 5% de su producto interno en educación teniendo una riqueza mayor y una proporción de la población en edad escolar menor, entonces, los demás deberíamos gastar más. Este razonamiento no permite especificar que «más» signifique eso: 6%, 7% o 20% del PIB.

- Patrones de asentamiento poblacional: los países con población dispersa (rural) requieren más recursos (para igual tecnología) para atender a la población.
- Tecnología a usarse: proporción aceptable de alumnos por docente (dadas las técnicas de enseñanza adoptadas); normas de repetición y promoción, estructuración de los grados (por ejemplo, existencia de ciclos de aceleración), etc.
- Características especiales de la población: multilingüismo, vulnerabilidad social, etc.

Segundo, porque la riqueza producida en cada país es diferente, de modo que así las necesidades de inversión fueran equivalentes, estas representarían necesariamente proporciones diferentes del PIB.

Tercero, porque la recomendación del «mínimo 6%» desconoce un hecho fundamental: la presión tributaria no es similar a través de los países, así cuando los países de la OECD –herederos del Estado del bienestar– recaudan por vía impuestos aproximadamente un 30% del PIB y destinan alrededor del 5% a educación, en, por ejemplo, América Latina la recaudación actualmente (dado que se incrementó en los años 1990) llega aproximadamente a 14% del PIB.

Cuarto, porque el indexar el gasto en educación al PIB hace que tal gasto sea sensible a los períodos de expansión o contracción de la producción nacional, lo que atenta contra la propia naturaleza del gasto social. En efecto, en los períodos de contracción económica son los sectores menos favorecidos de la población los que resultan más afectados, por lo que el gasto social en general (incluido el educativo) debería expandirse o, por lo menos, mantenerse, lo que no sería el caso si este se encuentra indexado al PIB.

En resumen, este tipo de «recomendación» es la que hace que los expertos de los ministerios de Economía o Hacienda miren con desdén a los expertos en educación, que vienen repitiendo este sinsentido por varias décadas, cuando dos minutos de análisis son suficientes para mostrar su profunda fragilidad y, sin embargo, los borradores actuales de Objetivos de Desarrollo Sostenible (SDG) incluyen una versión matizada de esta monserga¹⁵¹.

Entonces, el indicador gasto como porcentaje del PIB sirve para ver la magnitud del esfuerzo pero no dice nada acerca de los factores de los cuales depende ese nivel de esfuerzo, si es suficiente o cómo se distribuye.

¹⁵¹ A favor de la Unesco hay que decir que la Unesco Santiago ha presentado este tipo de consideraciones en el capítulo cuarto de Unesco Santiago (2007).

Cabe anotar que la fragilidad acá mostrada se aplica de igual forma a otras recomendaciones de la misma naturaleza, como que se debe «destinar x por ciento del gasto en educación a la educación de adultos», lo cual, siendo una táctica fácil para la movilización, resulta contraproducente por su falta de seriedad. De hecho, algunas personas que reconocen la debilidad de estas recomendaciones las consideran, sin embargo, útiles como mecanismos de movilización; sin embargo, esto parte por asumir que la gente no merece

Proporción entre gasto público en educación y gasto público total

Un segundo indicador usual está dado por la proporción que el gasto público en educación representa respecto del gasto público total. Este indicador, por su parte, muestra la importancia relativa, o prioridad financiera, que se le asigna a la inversión pública en educación en la operación del Estado.

Inversión por alumno

Un tercer indicador, aunque menos usual, está dado por los niveles de inversión por alumno por nivel educativo. Este indicador tiene la ventaja de aproximarnos a contar con una idea de lo que el propio servicio educativo implica. En ese sentido, empieza a ser una medida de lo que las personas efectivamente reciben como servicio educativo (siempre que se estime sobre la base de la parte del gasto que se destina a operaciones que acaecen a nivel escuela –por ejemplo, excluyendo los gastos de la administración–), aunque su utilidad está supeditada a tres elementos clave: (i) la omisión o no de la información sobre el gasto de las familias; (ii) la necesidad de contar con medidas de la dispersión o heterogeneidad del gasto y no solo de los niveles promedio; y (iii) los niveles de eficiencia en el gasto (gastos iguales pueden traducirse en servicios de calidades muy diferentes dependiendo de cuán competente sea la gestión de los recursos)¹⁵².

Detengámonos un poco en lo que cabría considerar a efectos de contar con información robusta sobre la inversión en educación.

Uno puede estar interesado en medir los niveles de inversión desde dos puntos de vista. Por un lado, puede haber un interés en hacer estudios sobre la rentabilidad (retornos) de la inversión y, por otro lado, se puede estar interesado en conocer si la inversión tiene la magnitud y forma suficientes como para garantizar el derecho de las personas a la educación. Dada la abundancia de trabajos sobre retornos económicos, véase una amplia compilación en el ya citado trabajo de Muñoz Izquierdo *et al.* (2004), este texto se preocupa principalmente de lo segundo y las líneas que siguen están orientadas a abordar ese tema.

argumentos serios y todo lo que se pueda ganar en movilización se puede muy fácilmente perder si alguien simplemente pregunta algo así como: «¿y por qué 6% y no 6,5%?»

Por cierto, esta situación no cambia por afirmar que el 6% es un «mínimo». Plantearlo como «mínimo» o como «óptimo» es igualmente injustificable.

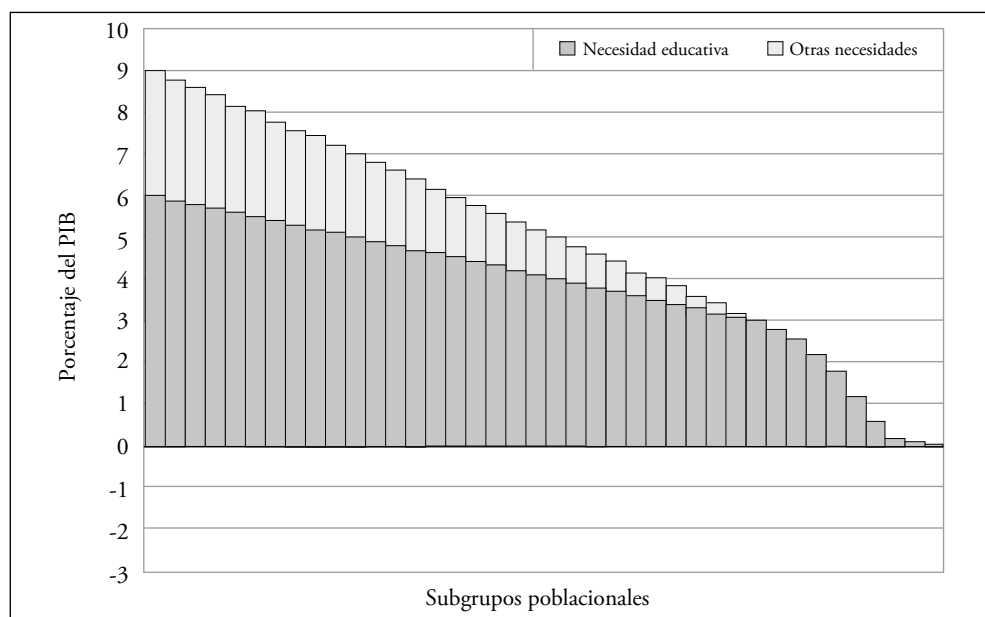
Luego de que en 2002 tuve la necesidad de formular por primera vez esta crítica a efectos de contribuir a las labores de la Unesco Santiago, este tema ha estado innumerables veces presente en mi trabajo. Recientemente, le he dedicado dos notas en mi blog personal, una de las cuales (<<http://saryagraha1602.blogspot.com/2014/02/mitos-y-miserias-de-la-educacion-un.html>>) se estructura alrededor de un gráfico irónico que muestra que en las regiones del Perú los resultados de aprendizaje se relacionan de modo **inverso** con el gasto como proporción del PIB.

¹⁵² A manera de ilustración, simplemente compárese lo que logran algunas instituciones benéficas o privadas con costos similares o menores a los de la provisión estatal.

¿Cuál es el nivel de inversión necesario para que las personas ejerzan su derecho a la educación? Esta pregunta no tiene una respuesta única, ya que las necesidades de las personas son diferentes en naturaleza y magnitud. Para efectos de ilustrar el argumento, vamos a presentar una situación hipotética de un país dividido en cuarenta segmentos de población (de igual tamaño, es decir, cada uno equivalente al 2,5% de la población total).

Cada uno de estos segmentos de población ha sido definido como homogéneo en su interior con relación a lo que las personas requieren para ejercer plenamente su derecho a la educación. Es decir, se asume que cada grupo requiere una canasta de bienes y servicios distinta. Así, habrá grupos con mayores necesidades educativas (por ejemplo, los hijos de padres de familia con menor nivel de escolaridad frente a los hijos de aquellos con educación terciaria completa) o con otras necesidades que satisfacer sin las cuales no se puede ejercer el derecho a la educación (necesidades de salud o nutrición, por ejemplo). Esto se presenta de modo simplificado en el gráfico 15.

Gráfico 15. Necesidades de inversión educativa por subgrupos. Información hipotética (en porcentajes del PIB)



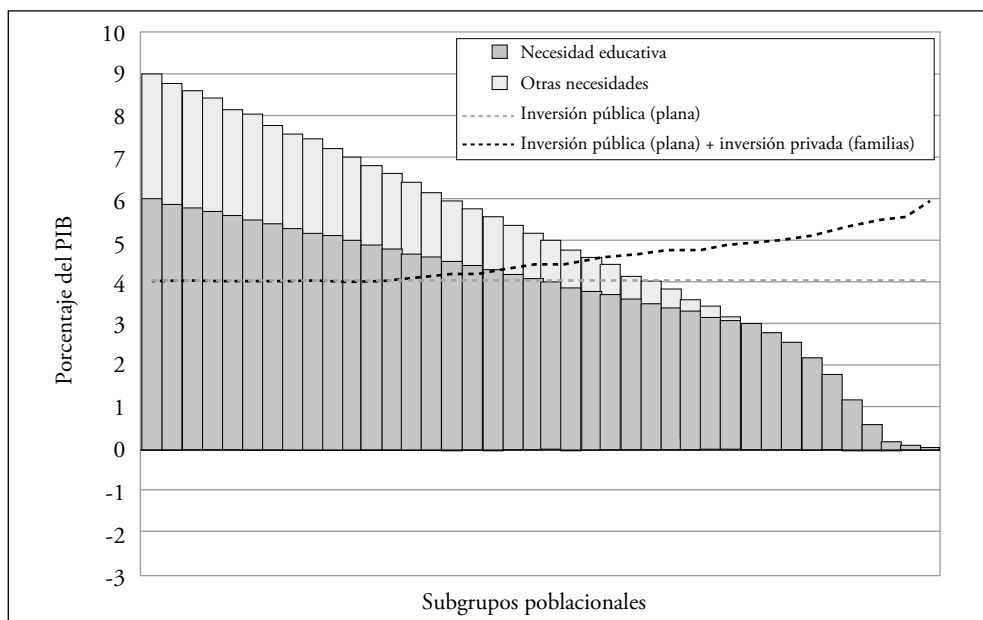
Nota

Los subgrupos han sido ordenados de menor a mayor necesidad de inversión educativa; el eje vertical expresa lo que la inversión representa como porcentaje hipotético proporcional del PIB. Es decir, los subgrupos son equivalentes a cuarenta países hipotéticos, donde en el más necesitado la inversión requerida sería equivalente a aproximadamente el 9% del PIB, mientras que el subgrupo más favorecido requeriría un nivel de inversión equivalente al 0,1% del PIB.

El gráfico 15 representa la diversidad en las necesidades de inversión. Ahora bien, frente a dichas necesidades habría que ver el comportamiento de la misma. Un esquema posible (y, en realidad, muy difundido) consiste en una asignación igualitaria de los recursos públicos bajo la forma de una educación gratuita para todos con recursos similares (similar número de alumnos por docente, docentes con calificaciones y salarios similares, etc.).

Ahora bien, junto a esta asignación pública, se tiene la inversión a cargo de las familias, que será diversa de un modo que se corresponde con las posibilidades financieras de cada grupo. Así, el efecto conjunto de la asignación plana (*flat*) de recursos públicos, más el gasto de las familias tendría la forma hipotética del gráfico 16.

Gráfico 16. Necesidades de inversión educativa pública y privada por subgrupos. Escenario uno: inversión pública (plana). Información hipotética (en porcentajes)



Nota

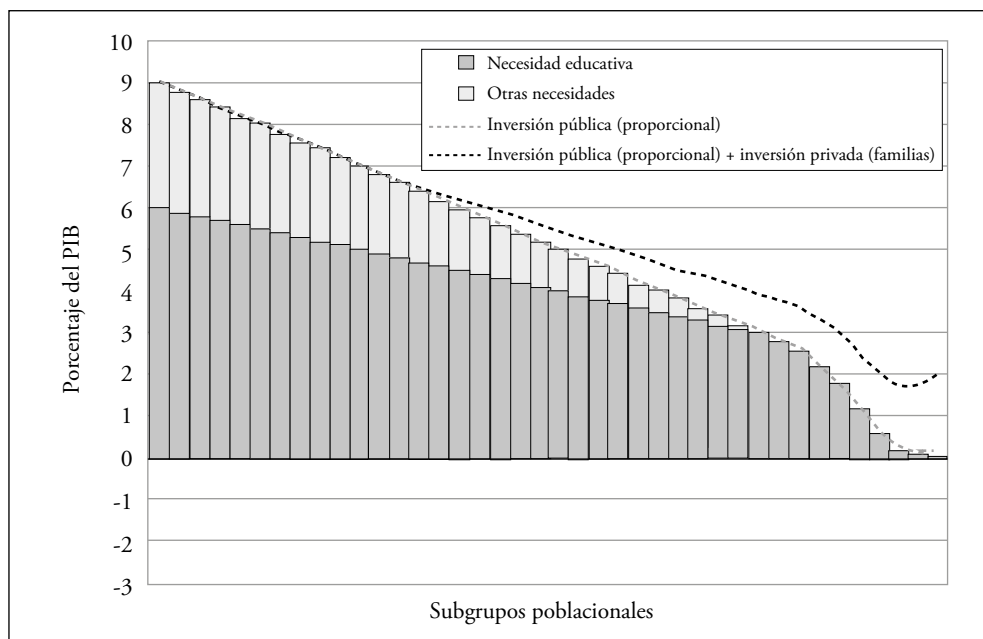
Véase la nota en el gráfico 15.

Como se puede apreciar, una asignación homogénea no hace sino reproducir las diferencias sociales dejando a un porcentaje de la población (en este caso hipotético, veintitrés de los cuarenta grupos) sin cubrir sus necesidades educativas y dando a los que más tienen más de lo que necesitan¹⁵³.

¹⁵³ El ejercicio asume que solo existe educación pública o que esta es reconocida como lo suficientemente buena como para que los sectores de mayores ingresos (que también contribuyen con mayor volumen de impuestos) opten por acceder a ella.

Entonces, ¿qué forma debería tener la inversión pública? Una primera opción es que esta sea directamente proporcional a las necesidades, es decir que tenga la forma del gráfico 17.

Gráfico 17. Necesidades de inversión educativa por subgrupos e inversión pública y privada. Escenario dos: inversión pública proporcional a la necesidad. Información hipotética (en porcentajes)



Nota

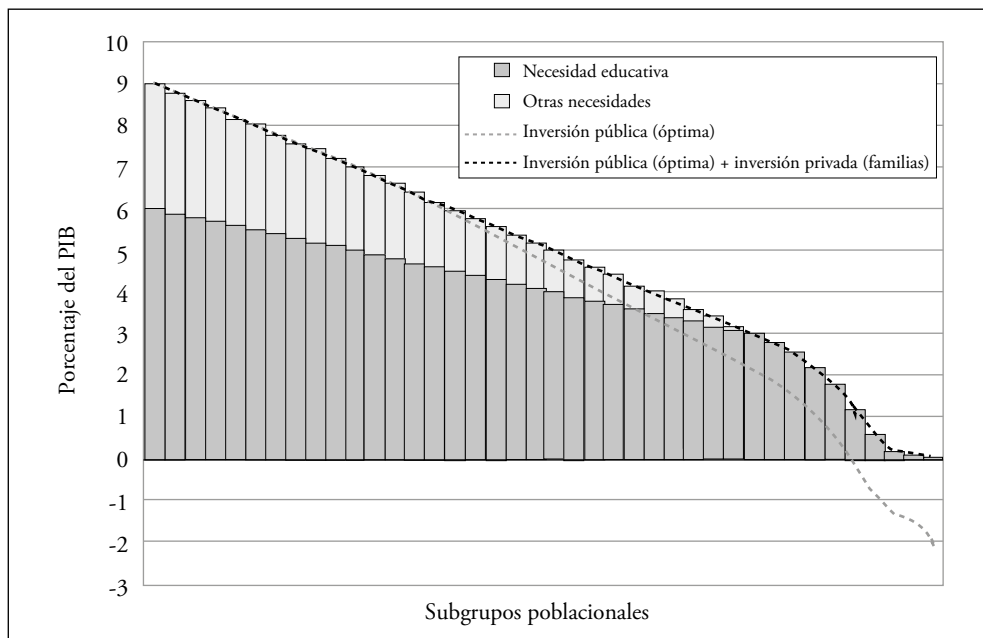
Véase la nota en el gráfico 15.

Sin embargo, las posibilidades de inversión de las familias seguirían teniendo un impacto negativo sobre la equidad, como también se aprecia en el gráfico 17. Entonces, una solución óptima sería la que señala el gráfico 18.

Como se puede observar, hay una fracción de la curva que representa el gasto público que tiene valores negativos (por eso desde el primer gráfico se incluyó esos valores negativos en el eje correspondiente). Es decir, una distribución «óptima» implicaría (en este caso hipotético) que las personas en los grupos más favorecidos de la población debieran recibir un subsidio negativo o, lo que es lo mismo, pagar por acceder a la educación pública.

Esta afirmación es consistente con el argumento en favor de una distribución equitativa de los recursos públicos; sin embargo, choca con un supuesto básico consagrado en la afirmación de que la educación es un derecho: nadie debería pagar por un derecho.

Gráfico 18. Necesidades de inversión educativa por subgrupos e inversión pública y privada. Escenario tres: inversión pública proporcional a la necesidad y corregida tomando en cuenta las posibilidades de las familias. Información hipotética (en porcentajes)



Nota
Véase la nota en el gráfico 15.

Sin embargo, si nadie debe pagar, entonces, ¿de dónde salen los recursos para garantizar los derechos? Dado que nada es gratis, siempre hay alguien que paga, ya que las cosas cuestan. Se puede pagar de modo directo o indirecto, pero siempre alguien paga, y lo importante es tener esto claro y hacerlo transparente.

Las personas tenemos derecho a tener un certificado de nacimiento en el que se registre nuestro nombre, fecha de nacimiento y filiación. Sin embargo, cuando solicitamos un certificado como ese usualmente tenemos que pagar el costo de su emisión (y de mantener los registros, etc.). En este caso, se paga un tributo que tiene una contraprestación directa (pagar el costo del certificado que se va a emitir). Esto es lo que se llama una tasa y es diferente de los impuestos en que estos últimos no suponen una contraprestación directa individualizada en el sujeto que paga: se paga un impuesto al valor agregado presente en el precio final de los productos que se compra o un impuesto a la renta, y este dinero financia un conjunto de servicios públicos de los que cada uno de los contribuyentes hace uso de manera diferente.

Es decir, el pago al que hace referencia el ejercicio hipotético puede tener una forma directa (como una tasa por estudiar en una escuela pública) o no (a través de un impuesto general), pero al final siempre es necesario (y muchas veces deseable) que las

personas paguen por el servicio. Que sea un pago directo hace que sea más visible, pero en el fondo siempre las personas pagamos justamente para garantizar nuestros derechos básicos (para eso existe el Estado y financiar su operación se traduce –o debe traducirse– en esa garantía de derechos).

* * *

Esto nos lleva a una discusión que va mucho más allá de los límites de este texto pero que es preciso tener en cuenta: la política fiscal es un elemento clave del financiamiento de los servicios educativos y del nivel en el cual se garantiza el derecho a estos. Las políticas de captación de ingresos (tributaria) y de distribución de los mismos como subsidios a la población (gasto público) son dos caras de una misma medalla (la política fiscal) y entender en qué medida el financiamiento público de la educación opera de acuerdo a principios democráticos y a la búsqueda de la equidad supone entender ambas partes del fenómeno simultáneamente.

Este ejercicio hipotético sirve para ilustrar una serie de discusiones centrales sobre el financiamiento de la educación que se verían profundamente enriquecidas si los sistemas de información abordasen estos temas de modo sistemático. Pero la información actualmente disponible de modo regular está muy lejos de ser suficiente para este fin. Sin embargo, hay algunas cosas que se pueden hacer que tienen que ver con el análisis de la heterogeneidad o falta de heterogeneidad del gasto público.

Si las personas tienen necesidades diferentes, las asignaciones presupuestales iguales no son la mejor forma de favorecer la equidad. Menos aun las que terminan beneficiando a los sectores que menores necesidades tienen.

Un tema clave en este terreno son los salarios docentes y la dotación de docentes en general. Si el gasto en la planilla de docentes es una fracción importante de la dotación de recursos existente, entonces la forma como se distribuye este gasto es de primera importancia para las políticas sobre financiamiento educativo. Si las personas tienen necesidades diferentes, cabría esperar que se requiera docentes con habilidades distintas y, consiguientemente, que los salarios docentes sean diferenciados. Salarios docentes planos o que solo obedezcan a consideraciones ajenas a las necesidades de los estudiantes (por ejemplo, solo el tiempo de servicios como criterio clave para obtener un ingreso mayor) revelan una asignación de recursos ajena a la preocupación por la equidad y a la garantía del derecho a la educación. De modo análogo, salarios docentes que son proporcionales al tamaño de las escuelas (no de las clases) suelen significar que se asigna mayores recursos a las escuelas urbanas que a las que atienden a población rural dispersa, lo que muy probablemente es completamente atentatorio contra la equidad.

De igual forma, ¿cómo se distribuyen las ratios de alumnos por docente según las necesidades?, ¿son estas ratios menores donde mayor necesidad hay?, ¿a qué obedecen las

ratios actualmente existentes?, ¿a una política expresa por la equidad?, ¿a una negociación sindical?, ¿a políticas de distribución de prebendas?

Sin embargo, las principales limitaciones en este terreno no se refieren a la disponibilidad de información sino a los prejuicios, ideologías e intereses que impiden explorar estos temas con mayor detalle y que ocultan que una gran parte de la educación estatal ha dejado de servir el interés público y, más bien, sirve intereses particulares (del gremio magisterial, por ejemplo)¹⁵⁴. Entre tales prejuicios e ideologías se encuentran: la creencia cuasi religiosa en una gratuidad que en realidad no permite ver que existe un problema de redistribución del gasto que hay que afrontar; el interés corporativo docente impide que se postule la necesidad de introducir diferenciaciones salariales en el magisterio; políticas de distribución de prebendas privatizan el uso de los recursos públicos; y así sucesivamente. Entonces, si nadie tiene interés en introducir diferencias para lidiar con realidades diferentes, ¿quién puede estar interesado en explorar la diversidad?

Existe un problema adicional que no es de menor importancia. La mayor parte del gasto público en educación corresponde a salarios docentes y los regímenes laborales de los docentes tienden a ser muy rígidos. Esto conlleva que los responsables de la política educativa perciban que tienen muy poco margen de maniobra con relación a los recursos que manejan y, por lo mismo, que la política educativa se concentre en lo que se puede hacer con el resto de los recursos.

Esta situación explica la importancia desmedida que parecen tener las iniciativas desarrolladas con endeudamiento externo: es dinero disponible para actuar; si bien representa una fracción minúscula de la inversión total en educación, es una que se puede manejar para hacer política educativa.

El problema que de esto resulta es la gran dificultad para tratar de hacer con el 3% –por poner una cifra como ejemplo– del presupuesto aquello que el 90% no logra. La pregunta clave entonces es cómo hacer para que el grueso de los recursos públicos destinados a la educación puedan ser palancas de políticas educativas que busquen asegurar el derecho de las personas a la educación. Esto, ineludiblemente, lleva al tema de los docentes, que se tratará más adelante.

¹⁵⁴ Los gremios tienen (o deberían tener donde este no sea el caso) todo el derecho de defender los intereses particulares de sus asociados y lograr beneficios para ellos. Lo que se quiere afirmar en este párrafo no es una negación de ese derecho, sino la necesidad de asegurar que la acción estatal se oriente, en primer lugar, a salvaguardar el interés público (y no un interés particular, cualquiera que este sea), es decir, en este caso, a proteger el derecho a la educación. Lamentablemente, la educación estatal, en muchas partes, ha dejado de salvaguardar dicho interés público, lo que no hace sino mostrar que estatal no es sinónimo de público. Cuando la legislación laboral docente, por ejemplo, asegura derechos a no asistir a clases por motivos sindicales sin que exista provisión alguna para asegurar que dichas clases no se pierdan, se tiene una legislación que hace que los recursos y la acción estatal se supediten a un interés particular (es decir, se privaticen) por encima del interés público. Los estudiantes no tienen un gremio, es función de la administración pública poner primero los intereses de los estudiantes y la ciudadanía.

Adicionalmente, es preciso considerar las posibles fuentes de información sobre el financiamiento, las mismas que incluyen:

- Registros administrativos sobre presupuestos públicos y su ejecución.
- Registros administrativos que brinden luces sobre los atributos del gasto público (en qué medida este es sensible a la diversidad).
- Información estadística (usualmente generada mediante encuestas) sobre los niveles de gasto de las familias y otros agentes privados.
- Idealmente, encuestas que permitan generar información sobre la magnitud de las necesidades de las personas en diferentes contextos.

Del mismo modo, es importante que la generación de información tome en cuenta algunos aspectos centrales que incluyen:

- Definir con precisión qué debe ser considerado como inversión en educación tanto a nivel de la acción pública como de las familias y/o agentes privados. Por ejemplo, ¿deben los aportes estatales a los fondos de pensiones de maestros ser considerados gasto en educación?
- Definir con precisión los procedimientos para asignar costos por niveles educativos cuando estos deben ser estimados, ya que usualmente es imposible contar con una diferenciación nítida. Por ejemplo, la depreciación de activos fijos que son compartidos por varios programas educativos.
- Diferenciar con precisión los gastos corrientes de los gastos de capital. En el caso de las familias, identificar también de modo diferenciado los costos de transacción.
- Diferenciar claramente la proporción del gasto que efectivamente se traduce en prestación del servicio de los elementos que hacen a la gestión general y que «no llegan a la escuela», sea porque no corresponde que lleguen o por otras razones.
- Prestar particular atención a la armonización de las fuentes de datos con relación a unidades monetarias¹⁵⁵, períodos de referencia (años calendario, años escolares,

¹⁵⁵ Las unidades monetarias pueden calcularse de diversas maneras. Por ejemplo, si la moneda local es el dracma transilvano (la moneda del Count Duckula © Thames Television), los valores pueden ser expresados en dicha moneda a valores corrientes (tal y como están registrados en un momento dado) o constantes, es decir, expresados en términos del valor de la moneda en un momento específico (por ejemplo: dracmas transilvanos de 1990). Igualmente, en el caso de que interese hacer comparaciones internacionales, es posible convertir la moneda local en una moneda de curso internacional (dólar norteamericano, euro, libra esterlina, yen, rupia, yuan, etc.) o en una canasta de monedas que permite corregir las fluctuaciones en los valores relativos de cada una, o también convertir los valores observados en otros países a la moneda local. Ahora bien, estas conversiones deben, a fin de lograr una mejor medida relativa, corregirse por la llamada PPP. Esta corrección busca tomar en cuenta que los costos de vida no son idénticos en todos los países, de modo que los bienes que pueden ser adquiridos con, por ejemplo, un dracma transilvano en un lugar, no son equivalentes a los que se pueden adquirir con dicha cantidad en otro país.

años fiscales, período de referencia de encuestas) y esquemas de clasificación y contabilización de la información.

2.4.7 EQUIDAD

El acápite anterior ha brindado alguna atención al tema de la equidad. Ciertamente, si la educación es un derecho humano fundamental, diferencias en el ejercicio del derecho tocan un tema básico de justicia que debe enfrentarse.

Pero, ¿qué es la equidad en educación? Siguiendo lo planteado en el capítulo segundo de Unesco Santiago (2007), la reflexión sobre la equidad implica encontrar un equilibrio entre los principios de igualdad (lo común) y diferencia (lo diverso). Esto refiere a que la reflexión sobre la equidad parte por identificar los aspectos en los que se busca la igualdad¹⁵⁶ entre las personas (por ejemplo, en su carácter como sujetos de derechos) reconociendo la diversidad de situaciones en las que estas viven. Así, es obligación de los sistemas educativos asegurar la equidad en el acceso, en las experiencias educativas y en los logros que de ellas derivan. Para ello, tales sistemas educativos deben tratar de forma diferenciada situaciones que son desiguales para llegar a resultados de aprendizaje equiparables y no reproducir las inequidades presentes en la sociedad.

Avanzar hacia esta aspiración de justicia demanda desarrollar instituciones educativas inclusivas capaces de atender la diversidad de necesidades de aprendizaje de cada persona.

Así, los sistemas de información deberían ser capaces de dar cuenta tanto de situaciones de paridad y/o disparidad con relación a fenómenos en los que se aspira a la igualdad (por ejemplo, niveles de acceso y de conclusión de estudios y logros de aprendizaje) como de la heterogeneidad en el tratamiento de la diversidad (por ejemplo, en qué medida el financiamiento de la educación es proporcional a las necesidades de las personas).

Lo primero, las situaciones de paridad y/o disparidad, respecto de fenómenos en los que se aspira a la igualdad, son relativamente de fácil manejo siempre y cuando se cuente con los indicadores apropiados para ello. En cualquier caso, esto plantea la necesidad de **promover un uso de indicadores con las mayores posibilidades de desagregación** por grupos de población relevantes en un país dado, de modo que se puedan comparar los valores observables para los indicadores entre distintos grupos, los cuales pueden incluir:

- Áreas de residencia: urbanas y rurales; regiones o provincias
- Pertenencia étnica y/o lengua materna
- Niveles de ingreso: deciles, cuartiles o quintiles

¹⁵⁶ Al decir de Sen (1995), la búsqueda de la igualdad implica definir (optar) por las dimensiones en las que esa igualdad se plantea como objetivo, lo que bien puede suponer desigualdad en otros ámbitos.

- Condición de pobreza: población pobre, población en pobreza extrema, población no pobre
- Grupos étnicos
- Sexo
- Combinaciones de los anteriores

La construcción de índices de paridad es una tarea relativamente sencilla, ya que solo supone la comparación directa entre los valores de un mismo indicador para dos grupos de población diferentes para los que se cuente con información censal. En el caso de que la información sea muestral, lo que corresponde es hacer una prueba de diferencia de medias o proporciones, según sea el caso o identificar la distancia entre los correspondientes intervalos de confianza. Por ejemplo: el porcentaje de la población de quince a diecinueve años que ha concluido la educación primaria según sexo.

Sin embargo, se debe tener cuidado en este procedimiento con relación a tres elementos:

- (i) Tener una presentación consistente de los índices, de modo que su lectura sea directa e inequívoca; por ejemplo, se puede optar por colocar siempre en el numerador al grupo de personas para las que se espera un valor menor del indicador, así el resultado del cálculo será de la forma: el valor del indicador x para el grupo de población α es z (una proporción) de lo observado en la población β .
- (ii) Hacer este cálculo solo en los casos en los que tenga sentido; es decir, es importante asegurarse de que, por ejemplo, un valor más alto del indicador es equivalente a una mejor situación, de modo que la lectura de la comparación tenga un significado claro (un índice menor a 1 quiere decir que la población en el numerador está peor que la considerada en el denominador). Este, por ejemplo, no es el caso de los índices de paridad de ratios brutas (de cualquiera de las ratios brutas existentes), ya que en estas más no quiere decir mejor (ni peor), ya que las diferencias se pueden explicar por fenómenos de diferente signo.
- (iii) Leer los resultados de modo consistente. Las proporciones pueden ser objeto de lecturas erróneas si es que no se toma en cuenta que los valores complementarios de fracciones de 1 no son intercambiables.

Un ejemplo interesante que ilustra esto está dado por la meta de la EPT referida al alfabetismo¹⁵⁷. Esta meta afirma, literalmente, que se debe buscar una mejora de los niveles de alfabetismo del 50%.

¹⁵⁷ El siguiente ejemplo se basa en un relato que me hiciera un colega que fue miembro de la comisión de redacción del marco de acción aprobado en la Conferencia Mundial de la Educación para Todos en Dakar (Unesco 2000).

Lo interesante de esta ilustración es que, durante la conferencia de la EPT de Dakar, la discusión sobre el tema del alfabetismo estuvo centrada en proponer que los países se comprometieran a reducir a la mitad las actuales tasas de analfabetismo, pero la siempre presente buena voluntad llevó a que alguien sugiriera que la meta debía ser redactada en términos positivos, así que «reducir a la mitad» fue transformado en «aumentar en 50%» el alfabetismo. ¿Es esto correcto?

La información con la que se trabajó en Dakar señalaba que, para entonces, la tasa de analfabetismo en el mundo era de aproximadamente 20%, es decir, la tasa de alfabetismo (su complemento) era de 80%. Reducir a la mitad el analfabetismo (es decir, hacerlo pasar de 20% a 10%) es equivalente a incrementar la tasa de alfabetismo de 80% a 90%, lo que es un incremento de 10 puntos en 80, es decir, del 13% y no del 50% como se terminó redactando. De hecho, incrementar la tasa de 80% en un 50% es hacerla subir en 40 puntos, es decir, hacerla llegar a 120%, lo que es, aritméticamente, un sinsentido.

Del mismo modo, si se está comparando los valores de un indicador para dos grupos de población cualesquiera, es preciso considerar lo siguiente: supongamos que el valor del indicador es 40 en un caso y 50 en el otro. La comparación (índice de paridad) $40 / 50$ dará un valor de 0,8; es decir, el primer valor es superior al segundo en 20% (la diferencia de 10 puntos porcentuales representa un 20% de 50). Sin embargo, la comparación inversa daría lo siguiente: $50 / 40 = 1,25$, lo que quiere decir que el segundo valor es superior al otro en 25% (10 puntos es el 25% de 40).

- (iv) Tener particular cuidado con las posibilidades reales de las fuentes de datos. Por ejemplo, si se combinan criterios y se está utilizando información muestral, puede que el nivel de error de las estimaciones sea demasiado grande como para calcular índices de paridad. En cualquier caso, lo más apropiado al usar información muestral es partir verificando la significancia estadística de las diferencias (sobre esto, véase la sección tercera del anexo 1).

En adición a los índices de paridad, es posible utilizar otras formas para observar diferencias y la heterogeneidad en las distribuciones de determinadas características. Dichas formas incluyen las diversas medidas de dispersión que dependen del nivel de medición de cada variable (razón de no modalidad, rango intercuartil, desviación estándar) y son presentadas en la segunda sección del anexo 1. Asimismo, hay gráficos que representan las distribuciones de frecuencias (simples o acumuladas) que ayudan a ver la forma de tales distribuciones y a hacer comparaciones entre ellas (estos gráficos también se describen en el anexo 1).

2.4.8 MEDICIÓN DE APRENDIZAJES (CON CONSIDERACIONES SOBRE EL CASO DEL ALFABETISMO ADULTO)

En este acápite se ha de abordar distintos aspectos vinculados a la medición de logros de aprendizaje. Se partirá por el tema del alfabetismo adulto por cuatro razones: (i) la primera tiene que ver con tratar de explicar los fundamentos de este tema, poniéndolo, momentáneamente, fuera del ámbito escolar para no perder de vista que el tema de los aprendizajes va más allá de un *curriculum* dado o de prácticas escolares; (ii) un motivo más bien biográfico, ya que mi ingreso con profundidad en este tema se dio a través de los problemas implicados en la medición internacionalmente comparable de competencias de alfabetismo entre adultos; (iii) porque permite establecer un puente entre el mundo de la medición de aprendizajes y la estadística educativa más tradicional, que aquí se busca reforzar al colocar el tema de la medición de logros en un marco general de información sobre los asuntos educativos; y (iv) porque una parte importante del mundo de las evaluaciones internacionales actualmente existentes (especialmente los programas conducidos desde la OECD) se inician conceptual y operacionalmente en este ámbito.

La forma más extendida de medir el alfabetismo de la población está dada por las llamadas **tasas de alfabetismo** (y su complemento, el **analfabetismo**) y por su contrapartida, la distribución de la población en dos categorías: alfabetizados y analfabetos.

Las tasas de alfabetismo suelen ser producidas mediante uno de los siguientes mecanismos:

- Las respuestas a una pregunta del tipo «¿Sabe usted leer y escribir?» o variantes de ella.
- La asunción de que las personas que han completado un determinado número mínimo de años de estudio (según la medida de nivel de escolaridad de la población presentada en el acápite anterior) han logrado ser alfabetizadas y, consiguientemente, las que no han alcanzado ese umbral han de ser consideradas analfabetas.

Estos procedimientos se basan en recomendaciones aprobadas por la Conferencia General de la Unesco en 1958 y posteriormente adoptadas por la UNSD (1997; § 2.145)¹⁵⁸.

Es necesario detenerse a revisar qué es lo que estos procedimientos efectivamente miden a efectos de evaluar su alcance y definir cuál es la mejor manera de usar esta información.

¹⁵⁸ La revisión 2008 de este documento (UNSD 2008: 147-148) brinda un tratamiento un poco distinto al tema, como veremos más adelante.

Cuando tal información es tomada como un indicador de lo que las personas son o no capaces de hacer a partir de su relación con materiales escritos, ambos procedimientos comparten una aproximación al fenómeno del alfabetismo que presenta dificultades muy importantes:

- En primer lugar, el fenómeno del alfabetismo es definido en términos **dicotómicos**, de tal suerte que las personas son consideradas como poseedoras o no de las capacidades de lectura, escritura y uso de números¹⁵⁹. Sin embargo, es perfectamente posible afirmar que las competencias no son algo que se posee o no, sino que tienen diferentes niveles de gradación correspondientes con la distinta complejidad de las operaciones a desarrollar. Las competencias de alfabetismo se presentan en un *continuum* (en realidad en tres *continua*: lectura, escritura y uso de números) que es impropriadamente representado por una dicotomía.

Así, la dicotomía alfabetizado/analfabeto es una representación inapropiada de la distribución de las capacidades de las personas y, lamentablemente, la revisión 2008 del documento ya mencionado de la UNSD (2008) persiste en dicha forma de abordar el tema, a pesar de anotar que la discusión actual tiende a ser planteada de una manera distinta. La experiencia de medición basada en usar *continua* de habilidades se ha traducido en declaraciones políticas, como las de la Conferencia Internacional de Educación de Adultos celebrada en 2009 (Unesco 2009), que insisten en la necesidad de abandonar la dicotomía anterior.

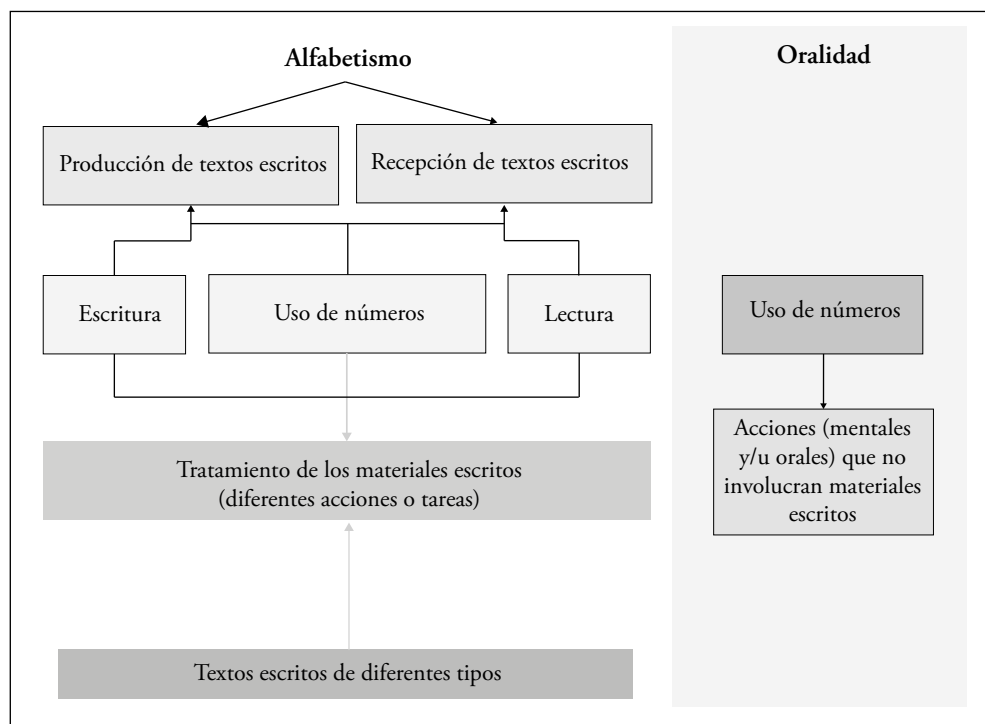
Cabe agregar, sin embargo, que es posible encontrar cierta utilidad en una representación dicotómica (por ejemplo, distribuir a la población entre aquellos que necesitan ser atendidos por un programa de alfabetización y aquellos que no), pero esta no tiene que ver con la naturaleza del fenómeno, sino con una operacionalización posterior que ayuda a definir las acciones de intervención y debería ser construida a partir de la definición sustantiva de un umbral relativo a lo que las personas pueden o no hacer, cosa que ninguno de los procedimientos previos permite.

- En segundo lugar, el fenómeno del alfabetismo es impropriadamente tratado como si fuera **unidimensional**; es decir, se asume que si una persona «no tiene la capacidad», entonces no la tiene de igual forma en ninguno de los dominios que componen el alfabetismo: no lee, no escribe y no puede hacer operaciones numéricas, y esto es claramente inapropiado. Leer, escribir y calcular son diferentes capacidades que se

¹⁵⁹ Usualmente no se pregunta por la capacidad de las personas para operar con números, pero las definiciones contemporáneas de alfabetismo incluyen la lectura, la escritura y la capacidad para operar con conceptos (como volúmenes, proporciones) y operaciones numéricas. Ejemplos de conceptos contemporáneos, así como una discusión de detalle de los temas tratados en este acápite, pueden consultarse en Unesco-UIS (2009b).

pueden desarrollar con cierto grado de independencia (por ejemplo, una persona puede calcular muy competently y, al mismo tiempo, puede no escribir)¹⁶⁰ y, además, están vinculadas con la naturaleza de los materiales que se ha de manejar. Estas limitaciones obedecen a que las competencias de alfabetismo son diversas y pueden representarse de la manera planteada en el gráfico 19.

Gráfico 19. Las competencias del alfabetismo



El gráfico 19 muestra que el alfabetismo involucra dos grandes operaciones: la producción y la captura o recepción de información escrita.

Estas grandes operaciones, por su parte, se corresponden con tres dominios de acción: la escritura, la lectura y el uso de números. Nótese que los dos primeros están siempre vinculados a la operación con textos escritos, mientras que el uso de números puede o no involucrar dicho tipo de mediación. Así, una parte del uso de números (el cálculo mental y su verbalización independientes de textos escritos) escapa al universo del alfabetismo, del mismo modo que la competencia comunicacional tiene una dimensión completamente oral.

¹⁶⁰ Si bien la experiencia escolar tiende a desarrollar la lectura, la escritura y el uso de números de modo conjunto, para las personas carentes de una experiencia escolar, este no es necesariamente el caso.

Los materiales escritos tienen diferente naturaleza como, por ejemplo, los contenidos en este texto: (i) textos continuos, organizados en frases y pasajes que se leen linealmente; (ii) textos no continuos, como los organizados en tablas y gráficos, que se leen de una forma no necesariamente lineal; (iii) textos que combinan ambas formas. De igual forma, los textos pueden ser de ficción o no ficción, así como escribirse o leerse para diferentes propósitos y de modo individual o colectivo, etc., de donde las prácticas del alfabetismo no se sustentan únicamente en habilidades individuales. Las habilidades requeridas para captar o para generar unos y otros tipos de textos son distintas. Asimismo, el nivel de dificultad de cada uno de estos se puede representar en un *continuum*.

La relación con los materiales escritos conlleva un conjunto de operaciones o tareas posibles que son el resultado de la conjunción entre el tipo mayor de dominio (leer, escribir, usar números) y el tipo de material escrito del que se trate. Por ejemplo, no es lo mismo localizar una pieza de información en un texto continuo que en una tabla matricial, ni es lo mismo calcular a partir de información provista en formato continuo que hacerlo a partir de una tabla.

De esta forma, una representación consistente de lo que las personas pueden hacer pasa por diferenciar:

- los grandes ámbitos a los que las operaciones refieren (lectura, escritura, uso de números);
- los dominios en los que las habilidades operan (lectura de textos continuos o discontinuos, producción de textos continuos o discontinuos, uso de números a partir de textos continuos o discontinuos);
- las diferentes operaciones que se realizan a propósito de los materiales escritos, como localizar, generar e integrar información y diversos tipos de operación numérica;
- los tipos de información (concreta, abstracta);
- los tipos de texto (ficción/no ficción); y
- los propósitos y contextos de la lectura¹⁶¹.

En resumen, medir lo que las personas pueden hacer es algo que las tasas de alfabetismo (computadas mediante cualquiera de las dos variantes antes descritas) no pueden hacer en ningún sentido.

A esto se debe añadir que el primer procedimiento de cálculo de las tasas (el basado en la pregunta), al requerir una autodeclaración sobre la condición de alfabetismo, introduce una característica más asociada a: (i) lo que cada entrevistado entiende como ser alfabetizado o analfabeto, que puede ser diferente; (ii) posibles incentivos para que

¹⁶¹ Sobre este último tema, véase Street (1998).

una persona se sienta impulsada a declarar de modo contrario a lo que considera que es su situación. Estos incentivos pueden obedecer a cuestiones de estigmatización o autoestima (que impulsan a declararse como alfabetizado) o a potenciales expectativas del entrevistado respecto de verse beneficiado por algún programa social destinado a los considerados analfabetos (que impulsan a declararse como tal).

Por su parte, el segundo procedimiento (el uso de un número de años de educación como umbral) tiene como problema que existe mucha evidencia que indica que lamentablemente la posesión de un número dado de años de educación no es un indicador directo de lo que las personas son capaces de hacer, tanto porque la capacidad de los sistemas educativos para asegurar los aprendizajes es variable (algunos sistemas educativos no logran para toda su población o para una fracción de esta lo que se supone que deben lograr en x años) como por el hecho de no tomar en cuenta que, dependiendo de las experiencias de vida, las habilidades logradas en la escuela se pueden desarrollar o perder.

Estos problemas no se subsanan cambiando, por ejemplo, las preguntas censales o las de las encuestas; o hacerlas más específicas y plantear: «¿Puede usted escribir su nombre y una pequeña carta?», en un intento bienintencionado pero inútil, dada la distancia entre lo que habría que conocer y lo que esta pregunta genera.

Entonces, si las tasas de alfabetismo no dicen nada sobre las competencias o habilidades de las personas, cabe preguntarse qué es lo que ellas miden, si es que miden algo.

La respuesta a esta pregunta pasa por considerar el procedimiento utilizado. El uso de los niveles de escolaridad logrados solo nos dice cuáles son los niveles de escolaridad logrados, nada más. Por lo tanto, ¿para qué adscribirle algún poder descriptivo (que no tiene) con relación a las competencias de alfabetismo?¹⁶² Esto no cambia la naturaleza de lo que es medido por el indicador (posesión de cierta certificación) y usarlo para un propósito para el que no se presta puede tener un efecto negativo sobre lo que sí mide. La pregunta sobre la autodescripción, por su parte, da cuenta de algo interesante: la manera cómo las personas se autoposicionan con relación a la distribución de las oportunidades educativas. Decir: «No sé leer ni escribir» es más que una evaluación de los propios saberes, es una afirmación de elementos de identidad social, ya que ello en determinados contextos puede ser entendido como: «Pertenezco a los grupos sociales más desfavorecidos, pues ni siquiera he podido lograr los rudimentos de una educación básica». En ese sentido, las tasas de alfabetismo informan sobre un hecho relevante (autopercpciones sobre el acceso a oportunidades) que vale la pena registrar

¹⁶² Por cierto, la propia idea de «analfabeto» fue históricamente construida en un proceso en el que la representación estadística jugó un papel muy curioso, ya que, lo que empezó como una cuantificación asociada a registros civiles, se transformó por una asunción fetichista de la estadística en un concepto educativo. Al respecto, véase Vincent (2009) y Hamilton (2012).

en la medida en que la información haya de ser interpretada en ese sentido. Así, estas tasas miden una percepción subjetiva que, siendo muy importante para dar cuenta de temas vinculados a la formación de identidades y su relación con la exclusión educativa, no informan acerca de lo que las personas son o no capaces de hacer cuando se ven confrontadas por una realidad mediada por la escritura¹⁶³. Este último aspecto, por otra parte, puede ser indagado de una forma no dicotómica y atendiendo a los diversos dominios (lectura, escritura, uso de números)¹⁶⁴ presentes en la idea contemporánea de alfabetismo.

En conclusión, las tasas de alfabetismo pueden brindar información valiosa, pero para ello es preciso no confundir su sentido.

Ahora bien, tan importante como conocer este elemento de formación de la identidad es saber lo que las personas son capaces de hacer con textos escritos. Por lo tanto, se necesita una medida efectiva de las competencias de las personas, ¿cómo construirla?

Se ha mostrado que las competencias de alfabetismo son un fenómeno complejo que implica diversas dimensiones no reductibles entre sí. Por lo mismo, es de esperar que la medición de las competencias no sea una tarea sencilla. De hecho, evaluar un fenómeno complejo requiere medidas lo suficientemente complejas que, si bien han de ser parsimoniosas, se correspondan con la naturaleza de lo que se quiere conocer.

El primer elemento de respuesta a la pregunta acerca de cómo evaluar competencias o habilidades es relativamente obvio: ellas se miden mediante una prueba o test¹⁶⁵. Este primer elemento de respuesta nos lleva, ineludiblemente, a la discusión sobre la evaluación de capacidades aprendidas y a todo un universo de aspectos y consideraciones conceptuales y metodológicas que es preciso tener en cuenta para hacer las cosas de una manera robusta y capaz de brindar información relevante. Es esa una razón por la que la discusión sobre la medición del alfabetismo nos deriva directamente al tema de la medición de aprendizajes. Una segunda razón es que la discusión actual sobre este tema está excesivamente centrada en aprendizajes vinculados a la lectura y al uso de números,

¹⁶³ Evidentemente, debe también considerarse que, en muchas ocasiones, las respuestas a la pregunta censal o a las encuestas es dada por un solo miembro del hogar, lo que introduce un factor que limita la posibilidad de considerar esta información como una «autopercepción».

¹⁶⁴ En este texto insisto en la necesidad de asociar la idea de alfabetismo a esos tres dominios no solo por lo ya expuesto, sino también por la tendencia actual a hablar de múltiples alfabetismos (digital, en salud, etc.), tendencia que convierte la idea de alfabetismo en sinónimo de posesión de habilidades básicas, lo que ciertamente tiene mucho sentido, aunque no resulta claro por qué se llama a eso «alfabetismo» y se hace perder la especificidad que este tiene, dada su relación con textos y números en formato escrito.

¹⁶⁵ Es interesante anotar que la Conferencia General de la Unesco en 1958, junto con proveer las definiciones tradicionales que han inspirado la medición del alfabetismo, también mencionó la conveniencia de desarrollar test muestrales de la población sobre este aspecto. Sin embargo, la investigación educativa relativa a la medición de aprendizajes mediante pruebas estandarizadas se encontraba, en ese momento, en sus fases más iniciales y es recién en los años 1980 que algunos países inician experiencias en este sentido, las que llevan a las primeras iniciativas internacionales en los años 1990.

aunque esto no tiene por qué ser así (como lo ilustra el tipo de estudios conducidos por la IEA que se reseñan en la cronología incluida en el anexo 2).

2.4.8.1 Temas fundamentales vinculados a la medición de aprendizajes

Tipos de evaluaciones y tipos de resultados. ¿Cuál es el propósito de la evaluación?

Las evaluaciones o mediciones de conocimientos, habilidades o actitudes pueden tener varios propósitos, los mismos que se traducen en elementos claves para definir el diseño de las pruebas e interpretar sus resultados.

Un primer tipo de evaluaciones son las pruebas que buscan **establecer una jerarquía en el desempeño de las personas (o *rankings*)**, de modo tal que se pueda determinar a los sujetos que se encuentran en posiciones relativas mejores o peores. Este tipo de prueba se usa, por ejemplo, para seleccionar a los «mejores» (de acuerdo a una determinada variable o conjunto de variables que se consideran importantes para la decisión). En este sentido, las pruebas son construidas de manera que maximicen las diferencias entre las personas a efectos de hacer la selección o el *ranking* más confiable.

Este tipo de evaluaciones es conocido como **pruebas referidas a normas**, de acuerdo al término acuñado por Glaser (1963). Ahora bien, las posiciones relativas en un *ranking* no necesariamente informan sobre qué es lo que las personas son capaces de hacer. Al priorizar la diferenciación, no prestan mayor atención acerca del espacio de la variable en la que esas diferencias se dan: «mejor» o «peor» no son necesariamente equivalentes a «bien/competente» o «mal/incompetente». Por esta razón, las lecturas de resultados de pruebas internacionales (como PISA) que se enfocan en el autoflagelamiento (muy común en América Latina), ya que ocupamos «los últimos lugares», son particularmente intrascendentes a pesar de que en algunos casos (como PISA) los *rankings* ocupan un lugar privilegiado en la agenda y en los esfuerzos de diseminación desarrollados por los responsables del estudio¹⁶⁶.

La necesidad de maximizar la diferenciación hace que estas pruebas tengan, además, una utilidad particular, ya que permiten **asociar los puntajes con otras variables que se pueden medir simultáneamente** y, entonces, identificar cuáles son los factores que muestran una mayor asociación con los resultados obtenidos. La promesa de este tipo de análisis está en poder identificar dichos factores a efectos de que la política pueda intervenir sobre ellos y, consiguientemente, alterar la situación encontrada.

Por ejemplo, si se encuentra una alta asociación positiva¹⁶⁷ entre los puntajes de escolares y la presencia de libros en el aula, se concluye que dado que las diferencias

¹⁶⁶ Esto obedece al énfasis de la OECD en los problemas de competitividad económica entre países, lo cual es una agenda política perfectamente legítima pero no precisamente una centrada en problemas educativos o en la necesidad de garantizar el derecho de las personas a la educación.

¹⁶⁷ Véase la segunda parte del anexo 1 con relación a lo que es una asociación (positiva/directa o negativa/inversa) entre dos variables.

en los resultados entre los estudiantes se asocia a dicha variable una acción de política deseable sería dotar a todas las aulas de libros.

La paradoja más importante de la lectura política de este procedimiento estadístico es que, si se dotase de libros a todas las aulas por igual, esta variable (la presencia de libros en el aula) dejaría de ser una (tendría un valor constante a través de toda la población) y, por lo tanto, no podría estar estadísticamente asociada a los resultados. Así, un nuevo estudio podría encontrar que las diferencias en los resultados no se explican por la dotación de libros, de donde **no** debería concluirse que los libros no importan para el aprendizaje, ya que la asociación estadística no es equivalente a la importancia de un factor dado. Lo mismo se aplica a otros elementos que podrían estar presentes de modo universal (y, por lo tanto, no presentar variaciones).

Una lectura simplista que equipara asociación con importancia es una lectura potencialmente peligrosa y debe evitarse a toda costa. Hay cosas que importan para los aprendizajes (y no solo para aquellos aprendizajes que son medidos mediante una prueba) o respecto del ejercicio del derecho de las personas a la educación que no necesariamente se representan adecuadamente por su asociación con la variación en los resultados. De hecho, incluso se puede encontrar asociaciones con hechos no deseables: estudiantes con docentes autoritarios podrían lograr mejores puntajes en, por ejemplo, matemáticas y de esto no se debe colegir que hay que impulsar el autoritarismo entre los docentes, pues este atenta contra otros aprendizajes (el desarrollo de la autonomía y de la capacidad de vivir en colectividad de una manera democrática), contra principios democráticos básicos inmersos en el carácter de la educación como derecho humano fundamental y contra la propia dignidad de las personas.

Por otra parte, es importante entender que la educación es un fenómeno humano entroncado en un tejido amplio de relaciones sociales y cultura, tal y como se expresan en contextos específicos. Esto quiere decir, como se señaló anteriormente (ver la página 91), que lo que «funciona» en un contexto dado, no necesariamente «funcionará» de igual forma en otro; por el contrario, el carácter generativo de las interacciones hará que medidas iguales en contextos distintos muy probablemente conduzcan a resultados diferentes (Pawson y Tilley 1997).

Por lo tanto, usar los resultados de los estudios de logros de aprendizaje –incluidos sus hallazgos en términos de asociaciones que explican la variabilidad– como recetarios de política es un despropósito mayúsculo.

Un segundo tipo de pruebas está diseñado a efectos de **describir lo que las personas saben hacer** sin poner el énfasis principal en maximizar las diferencias. Por ejemplo, si se desea medir quiénes lograron los aprendizajes necesarios y se encuentra que la respuesta es un valor como 99%, esto son buenas noticias que también son necesarias, pero no habría suficiente variabilidad en los resultados como para hacer análisis de asociación,

como los anteriormente indicados. Glaser (1963) llamó a estas **pruebas referidas a criterios**.

En el caso de las competencias de alfabetismo a las que nos venimos refiriendo, pruebas de este tipo resultan muy pertinentes ya que importa primariamente definir qué es lo que las personas saben hacer y no construir *rankings*.

Ambas aproximaciones brindan información útil pero de diferente naturaleza. Ahora bien, también es posible construir pruebas que combinen ambos enfoques.

Psicometría

Todo lo anterior puede sonar muy interesante, pero hacerlo implica desafíos técnicos de importancia. La psicometría se ha dedicado a indagar estos temas y a proveer herramientas que permiten construir pruebas más robustas. Elaborar adecuadamente una prueba, computar sus resultados y analizarlos con propiedad requiere un conocimiento especializado complejo.

Los aspectos involucrados en esta tarea incluyen los siguientes elementos básicos a tener en cuenta:

- Es necesario entender que los **rasgos** sobre los que se quiere obtener información (la habilidad lectora, por ejemplo) no son directamente observables a la manera de otras variables (por ejemplo, la edad de una persona extraída directamente de un registro en el que se consigna la fecha de nacimiento). Por lo mismo, se habla de variables **latentes** que son observadas mediante variables manifiestas (por ejemplo, la manera como una persona responde un conjunto de preguntas que suponen la comprensión de un texto escrito que se le presenta en una prueba)¹⁶⁸.
- El análisis de variables latentes, como otros ámbitos del análisis estadístico, tiene diversas formas que dependen de la naturaleza o nivel de medición tanto de las variables manifiestas u observables como de las propias variables latentes. Por ejemplo, si tenemos variables manifiestas continuas, que se entiende dan cuenta de una variable manifiesta también de tipo continuo, es posible conducir **análisis de componentes principales**; si, por otro lado, se cuenta con variables observables de nivel ordinal (respuestas correctas e incorrectas a un conjunto de preguntas) que han de dar cuenta de una variable latente continua, es posible conducir modelamientos basados en el **análisis de rasgos latentes**, que es, exactamente, lo que tiende a hacer la parte de la psicometría que es normalmente usada en las pruebas de logro de aprendizajes; de igual forma, si tenemos variables observadas

¹⁶⁸ Sobre el análisis de variables latentes y sus diversas formas, véase: Borsboom *et al.* (2003); Henry (1999); Lazarsfeld y Henry (1968); y Magidson y Vermunt (2001).

nominales que dan cuenta de una variable latente nominal, se conduce un **análisis de clases latentes**, etc.

- Más allá de lo anterior, es necesario, en primer lugar, contar con **constructos**¹⁶⁹ bien definidos; es decir, especificaciones claras y unívocas de aquello que se ha de medir mediante los instrumentos. Por ejemplo, se puede asumir que la capacidad lectora incluye (pero no se limita a) **la capacidad para localizar una pieza de información explícita en un texto discontinuo carente de información adicional**. Estos constructos usualmente se organizan en un marco general que acota la evaluación.
- A partir de los **constructos**, en segundo lugar, se debe definir **ítems** que los midan con propiedad, es decir, preguntas que clara y unívocamente supongan desarrollar las tareas implicadas en dichos **constructos**. Tomando el anteriormente definido constructo, para el gráfico 20 una pregunta posible sería: «¿Cuál es la velocidad máxima permitida para la circulación vehicular en una vía marcada con el símbolo aquí reproducido?»

Gráfico 20. Velocidad máxima permitida



Este caso muestra, además, que el ítem (la pregunta) está construido con relación a un **estímulo** o reactivo (la imagen en este caso; sin embargo, el estímulo puede estar inmerso en el propio ítem como, por ejemplo: «¿Cuántos vasos de medio litro de aceite pueden llenarse con un barril de 150 litros?»).

Por otra parte, es importante validar mediante análisis de la información empírica la validez y confiabilidad de un ítem para medir determinado atributo.

- Una adecuada representación de los resultados a través del cálculo de **puntajes** en una escala. Por ejemplo, una respuesta correcta al ítem anteriormente presentado contribuirá en una determinada magnitud al puntaje general obtenido en la escala de lectura de textos no continuos.

¹⁶⁹ Un constructo es cada una de las entidades conceptuales que se establecen de modo que se cuenta con una clara especificación de lo que se desea medir.

Este tema plantea una dificultad técnica adicional: ¿cómo determinar la contribución de cada ítem al puntaje final?

Tradicionalmente, los resultados de las pruebas cognitivas, como las que se aplican usualmente en los programas educativos desde la primaria hasta la educación superior, se computan como una fracción de respuestas correctas sobre el total de preguntas. Esta fracción, por su parte, se expresa en términos de una escala de 100 puntos (porcentaje de respuestas correctas) u otra que se haya adoptado como convención (sobre 7, 10 o 20 puntos, por ejemplo) y donde se define un umbral que se considera satisfactorio.

Ahora bien, debe notarse que dos personas que logran en una misma prueba un resultado equivalente, por ejemplo 50% de respuestas correctas, no necesariamente tienen un rendimiento igual, ya que las preguntas que cada una respondió correctamente pueden ser diferentes entre sí con relación a su dificultad. No es lo mismo obtener 50% habiendo respondido correctamente la mitad más fácil de la prueba que haberlo hecho respondiendo algunas de las preguntas fáciles y algunas de las difíciles.

Por lo tanto, la forma tradicional de computar resultados como porcentaje de respuestas correctas puede mejorarse mediante una «ponderación» de cada pregunta. La ponderación puede ser establecida de modo arbitrario o se puede hacer un análisis detenido del comportamiento empírico de los ítems, de forma tal que se conozca de modo sistemático sus propiedades básicas (como el nivel de dificultad observado –y no solo esperado–). Justamente, esta es la tarea primordial a la que se aboca la llamada **teoría de respuesta al ítem**¹⁷⁰.

En efecto, es posible hacer un análisis empírico del comportamiento que efectivamente tienen los ítems para determinar: (i) su nivel de dificultad; (ii) su capacidad discriminatoria; y (iii), en el caso de las preguntas de respuesta múltiple, el efecto posible de la adivinación en las respuestas. Calcular los parámetros de cada ítem permite una medición más robusta de las capacidades de las personas, pero supone un complejo análisis estadístico.

Lo anterior muestra de modo muy sumario la complejidad imbricada en la construcción de una prueba robusta y, del mismo modo, la fragilidad de enfoques que tienden a sobresimplificar la tarea de medición. Asimismo, estas reflexiones subrayan la

¹⁷⁰ La misma que, por cierto, no es una teoría sino más bien un conjunto de modelos matemáticos usados para determinar las propiedades de ítems (ordinales) y, a partir de ahí, analizar los resultados sobre una escala continua. Como su nombre lo indica, el foco del tratamiento de las pruebas es cada ítem singular y cómo varios ítems operan en conjunto, a diferencia de las aproximaciones «clásicas», que se centran en el desempeño agregado de cada test.

importancia de que los sistemas de información cuenten con las competencias técnicas requeridas para una medición que brinde información realmente valiosa.

Existen muchos ejemplos de niveles de sobresimplificación absurda a los que se puede llegar en esta materia. Por ejemplo, a propósito de la medición del alfabetismo, es claro que la administración de un número pequeño de ítems contruidos sin procesos de validación rigurosos no es la mejor manera de abordar tal tarea. Lamentablemente, esta es una práctica que se ha venido extendiendo mediante la «mejora» muy bien intencionada de los cuestionarios que se aplican en encuestas de hogares, aunque la buena intención, una vez más, no garantiza buenos resultados. Del mismo modo, pretender que un fenómeno complejo como el alfabetismo puede capturarse adecuadamente midiendo un solo aspecto, así sea de gran importancia, como la fluidez lectora, con una sola toma de datos (por ejemplo, mediante la identificación del número de palabras por minuto que la persona puede leer) es un ejercicio muy poco riguroso.

Este último ejemplo merece una mención adicional detenida. La fluidez en la lectura de un texto continuo (aproximada por el número de palabras que una persona decodifica en una unidad de tiempo) es un aspecto crucial en la capacidad lectora, pero no es el único y difícilmente se puede usar como predictor del comportamiento de las personas en los otros dominios al respecto.

En los últimos años se ha desarrollado significativos esfuerzos para medir las capacidades de lectura adquiridas por los estudiantes en sus primeros grados de estudio (de modo que cuando un problema se diagnostique no sea demasiado tarde como para poder actuar, riesgo implicado en una medición más tardía). Un esfuerzo muy importante está dado por el Early Grade Reading Assessment (EGRA)¹⁷¹, programa que busca dar cuenta de la adquisición de diversas habilidades implicadas en el desarrollo de la capacidad lectora¹⁷².

Sin embargo, algunas personas y organizaciones, han mostrado una tendencia a la sobresimplificación que busca reducir este tipo de esfuerzos a la simple medición del número de palabras decodificadas en una unidad de tiempo.

La fluidez lectora es de capital importancia, nadie lo niega. Medirla puede ser muy importante para tener una señal de alerta, pero tenerla es insuficiente para saber qué hay que hacer. Por eso el reduccionismo suele ser un mal consejero. De hecho, sobre enfatizar los aspectos de la mecánica de la lectura (la decodificación, en este caso) puede llevar a fortalecer prácticas pedagógicas fundamentalmente conductistas que no consideran que

¹⁷¹ Véase <<https://www.eddataglobal.org/documents/index.cfm?fuseaction=showdir&ruid=1&statusID=3>>.

¹⁷² El EGRA mide los siguientes **componentes de lectura**: reconocimiento de letras, pronunciación, lectura de palabras usuales, lectura de palabras no usuales, lectura y comprensión de pasajes, además de dictado (escritura) y comprensión auditiva –para verificar competencia en la lengua de la evaluación– (RTI International 2009).

la operación de leer es, fundamentalmente, una capacidad de manejar significados y acciones comunicativas.

En adición a lo anterior, y esto es mucho peor aún, en ocasiones se trata de introducir una suerte de estándar universal de, digamos, cuarenta palabras por minuto como objetivo deseable al final del segundo o tercer grado de primaria.

Las discusiones técnicas sobre estos temas suelen incluir «justificaciones» que afirman que «si bien esto no es completamente riguroso, es suficiente». Este tipo de afirmaciones incurren en dos problemas: (i) enunciar la suficiencia no es equivalente a demostrarla; y (ii) ¿por qué algo no muy riguroso es aceptable cuando se trata de la educación y no en otros casos? ¿Los investigadores, consultores y agentes de política que afirman esto, aceptarían, por ejemplo, que un médico les practique una intervención quirúrgica que «no va a ser muy rigurosa»? ¿La educación merece menos rigurosidad que la salud del consultor? ¿O solo la educación de «otros» merece esa menor rigurosidad? Si para conducir un automóvil, para no hablar de la cirugía, se necesitan varios indicadores, ¿por qué librar la suerte de los niños pobres a un cronómetro?

Del mismo modo, ¿por qué 40 palabras por minuto y no 37 o 53? Esa cantidad se basa en observaciones que dan cuenta de que la velocidad de lectura entre los niños de grados tempranos en diversas partes del mundo es muy limitada, así 40 palabras por minuto (o 50, 43, o algo así, definido siempre de modo arbitrario) parece ser un objetivo mínimo alcanzable. Sin embargo, esto plantea dos problemas mayúsculos:

- (i) Los idiomas y sistemas de escritura son distintos y la longitud típica de las palabras varía entre ellos. Por ejemplo, el concepto «visión del mundo» en español requiere tres palabras con catorce letras, en inglés (*worldview*) es una palabra con nueve letras, mientras que en alemán (*Weltanschauung*) es una con catorce letras. Nótese que este ejemplo solo está usando lenguas europeas; si lo ampliásemos de modo de considerar lenguas no europeas como, por ejemplo, el kiswahili o el vietnamita, los contrastes serían mucho mayores.
- (ii) Niños de sectores sociales favorecidos, a esa edad, muestran que pueden leer con comprensión (y no solo decodificar) unas 200 o 250 palabras por minuto (en inglés o español). Entonces, asumir un estándar tan bajo puede resultar regresivo y discriminatorio. Regresivo, por la tendencia a considerar que el indicador es idéntico a lo que «debe» lograrse, y discriminatorio, pues los que recomiendan dichos niveles de fluidez como «aceptables» para pobres jamás los considerarían efectivos para sus propios hijos.

La sobresimplificación, incluso cuando es bienintencionada, suele resultar nefasta y revelar profundos prejuicios y actitudes discriminatorias.

Estandarización de las pruebas y equivalencia de los resultados

La medición de las capacidades de las personas no es una tarea simple, menos cuando se trata de contar con información sistemática a nivel nacional o internacional.

Un docente puede administrar pruebas que él mismo construye (o debería construir) a estudiantes que él conoce (o debería conocer) prestando atención (o debiendo prestarla) al proceso educativo en el que los alumnos se encuentran inmersos. Así, los resultados de las pruebas son (o deberían ser) leídos desde un complejo conocimiento sobre el entorno específico de los alumnos, donde la información de las pruebas es (o debería ser) una pieza más que informa (o debería informar) sobre la manera como se gestionan los procesos de aprendizaje. Sin embargo, aun en el mejor de los casos, los resultados de este tipo de pruebas no pueden simplemente agregarse para contar con información nacional (o internacional que para el caso es lo mismo), ya que cada prueba es construida de una forma que se ajusta (o debería ajustarse) a un contexto específico y reacciona a él, de modo que su utilidad para el proceso educativo esté (o debería estar) asegurada.

Esto, por otra parte, ni siquiera es una posibilidad remota cuando se trata de conocer las habilidades de las personas adultas que no se encuentran necesariamente participando de algún programa educativo en un momento dado. Este es el caso de la medición del alfabetismo adulto.

Entonces, se plantea un desafío vinculado a la generación de información consistente y con significado más allá de un entorno educativo específico. La respuesta inmediata a este desafío sugiere construir pruebas estandarizadas, es decir, pruebas que midan lo mismo a través de diversos contextos (como estos son siempre locales, la agregación de información en el plano nacional plantea el mismo tipo de desafíos que la agregación a escala internacional, aunque la magnitud de algunos de esos desafíos pueda ser diferente).

Sin embargo, esta no es una tarea de poca envergadura, como sugieren algunos defensores de métodos «simples» y «baratos» que no «compliquen» las cosas en demasía¹⁷³. **La medición de conocimientos, habilidades, destrezas y actitudes de una forma estandarizada choca con el hecho de que estamos hablando de fenómenos humanos firmemente entroncados en la cultura y en contextos que siempre son específicos.** Esto, por cierto, ya había sido observado por Luria (1976) en los años 1930.

En el caso del alfabetismo, por ejemplo, estamos hablando de una forma de comunicación (la basada en el texto escrito) y uso del lenguaje y no hay nada más dependiente del contexto que esto.

De hecho, ni siquiera las operaciones matemáticas (algunas de las cuales hacen parte del mundo del alfabetismo cuando están mediadas por material escrito) pueden considerarse independientes del contexto, como lo demuestra un conjunto muy amplio de

¹⁷³ Si bien esta es una tendencia que se observa en muchos lugares, en el mundo académico es representada por Wagner (2003, 2011).

investigaciones que se suelen agrupar bajo el rótulo de «etnomatemáticas»¹⁷⁴. De hecho, por ejemplo, extraer la décima parte de una cantidad (dividir entre diez o calcular el 10%) es una operación diferente en culturas que usan sistemas de numeración no decimal (de base cinco o veinte, por ejemplo). Del mismo modo, diferentes culturas realizan las operaciones aritméticas básicas usando notaciones o procedimientos diferentes que implican operaciones mentales distintas¹⁷⁵.

Así, la construcción de pruebas estandarizadas es, a juicio de algunos, un imposible conceptual o un despropósito empírico, ya que justamente estandarizar implica remover el efecto del contexto y esto es un completo sinsentido en el plano de la lengua y la cultura¹⁷⁶. De hecho, algunos expertos en «evaluación educativa» buscan incesantemente utilizar instrumental «libre de cultura»¹⁷⁷ o «contextualizar» los instrumentos como un procedimiento que, en realidad, tiene como propósito «domesticar» los efectos del contexto, de modo que la medición estandarizada sea posible.

En efecto, el diseño de las pruebas estandarizadas ha tendido (al menos en algunos casos) a ser más sensible al contexto. Así, la mayor parte de este tipo de pruebas se construye incluyendo preguntas que no tienen un carácter algorítmico abstracto (aunque lo que se mide sea una capacidad abstracta) y se apela, por ejemplo, a ítems «reales» (con estímulos tomados de la vida real) o que simulan condiciones de la vida cotidiana de las personas. El problema es que la vida cotidiana nunca es estándar, por lo que estos ítems pueden, justamente, mostrar más dificultades¹⁷⁸. Por otra parte, al construir una «situación cotidiana estándar» se corre el riesgo de asumir que ciertas formas de vida cotidiana son universalizables, lo que tiene una potencial vena etnocentrista tremendamente peligrosa en términos éticos y que, además, liquida cualquier pretensión de validez.

Por ejemplo, podemos tomar un estímulo o reactivo usado en Adult Literacy and Lifeskills Survey (ALL; gráfico 21).

¹⁷⁴ Este término fue acuñado por el educador brasileño Ubiratan D'Ambrosio.

¹⁷⁵ Por ejemplo, los métodos tradicionales de multiplicación usados en Etiopía.

¹⁷⁶ Tanto la lingüística y la filosofía del lenguaje ordinario como las ciencias sociales han prestado suficiente atención a estos temas como para que quede bastante claro que el significado de las acciones comunicativas resulta de la interacción entre el texto, el contexto y la acción de las personas.

¹⁷⁷ No es broma, esto lo he visto por escrito en un documento muy serio que no es necesario identificar.

¹⁷⁸ Cooper y Dunne (1998, 2000) muestran, por un lado, sesgos clasistas en los resultados de pruebas estandarizadas en el Reino Unido, los que, justamente, se fundan en la «contextualización» de los ítems; y, por otro lado, cómo esto plantea problemas vinculados a diferencias en el desarrollo de capacidades de abstracción (lo que recuerda los hallazgos de Luria). Esto, en sí mismo, introduce un problema de validez de la prueba (ya que no termina midiendo lo que se propone).

Gráfico 21. Estímulo usado por la encuesta ALL como situación cotidiana estándar

PROPER FRAME FIT

RIDER MUST BE ABLE TO STRADDLE BICYCLE WITH AT LEAST 2 cm CLEARANCE ABOVE THE HORIZONTAL BAR WHENSTANDING.

NOT LESS THAN 2cm

NOT LESS THAN 2cm

NOTE: Measurement for a female should be determined using a men's model as a basis.

OWNER'S RESPONSIBILITY

PROPER SIZE OF BICYCLE	
FRAME SIZE	LEG LENGTH OF RIDER
430mm	660mm-760mm
460mm	690mm-790mm
480mm	710mm-790mm
530mm	760mm-840mm
560mm	790mm-860mm
580mm	810mm-890mm
635mm	860mm-940mm

1. **Bicycle Selection and Purchase:** Make sure this bicycle fits the intended rider. Bicycles come in a variety of sizes. Personal adjustment of seat and handlebars is necessary to assure maximum safety and comfort. Bicycles come with a wide variety of equipment and accessories . . . make sure the rider can operate them.
2. **Assembly:** Carefully follow all assembly instructions. Make sure that all nuts, bolts and screws are securely tightened.
3. **Fitting the Bicycle:** To ride safely and comfortably, the bicycle must fit the rider. Check the seat position, adjusting it up or down so that with the sole of rider's foot on the pedal in its lowest position the rider's knee is slightly bent.

Note: Specific charts illustrated at left detail the proper method of determining the correct frame size.

The manufacturer is not responsible for failure, injury, or damage caused by improper completion of assembly or improper maintenance after shipment.

Fuente: OECD y Canadá-Statistics Canada (2005: 285).

Nota

Traducción propia:

«Ajuste adecuado del marco.

El ciclista debe ser capaz de montar dejando al menos dos centímetros libres de la barra horizontal al sentarse.

Nota: la medida para una bicicleta femenina debe determinarse usando la masculina como referencia.

[La tabla específica a la izquierda es el método apropiado para determinar el tamaño correcto del marco y se refiere al tamaño adecuado de la bicicleta de acuerdo al del marco y a la longitud de la pierna del ciclista.]

Responsabilidades del propietario

1. Selección y compra de la bicicleta. Asegurarse de que esté ajustada al ciclista que la habrá de usar; las bicicletas vienen en tamaños diversos; el ajuste personal del asiento y el timón es necesario para asegurar el máximo de confort y seguridad; las bicicletas vienen con una amplia variedad de equipos y accesorios... asegúrese de que el ciclista los puede operar.
2. Armado. Siga cuidadosamente todas las instrucciones de armado, asegúrese de que todos los pernos, tornillos y tuercas estén ajustados firmemente.
3. Ajustando la bicicleta. Para manejar cómodamente, la bicicleta debe estar ajustada para el ciclista. Revise la altura del asiento ajustándola de modo que al poner la suela del zapato del ciclista sobre el pedal en su posición más baja, la rodilla esté ligeramente flexionada.

El fabricante no es responsable por fallas, lesiones o daños causados por un armado inadecuado o por un inapropiado mantenimiento posterior al envío.»

Al respecto surgen varias preguntas posibles:

- ¿En qué medida cabe asumir que las personas en diversos lugares del mundo compran las bicicletas desarmadas y las ensamblan siguiendo un conjunto de instrucciones impresas?
- ¿En qué medida cabe asumir que las personas efectivamente definen el tamaño de la bicicleta que han de comprar a partir de las medidas y procedimientos que aquí se muestran?
- Ahora bien, ¿cuáles son las preguntas que cabe sensatamente hacer usando un estímulo como ese? Un despropósito mayúsculo sería, por ejemplo, preguntar acerca de las razones para que las personas sigan las instrucciones de ensamblaje con cuidado (usualmente las personas tienen razones para ello independientemente de la lectura del texto: que la bicicleta funcione correctamente; y esto no aparece en el estímulo bajo ninguna forma) y asumir que la respuesta sería que esto debe hacerse así ya que el fabricante no asumirá ninguna responsabilidad por los daños que no hacerlo puede generar (párrafo final en el estímulo)¹⁷⁹.

Entonces, ¿es posible medir las competencias de alfabetismo de las personas (en realidad esta problematización se aplica también a cualquier otra forma de medición de logros educativos) de una forma sistemática y equivalente a través de diversos contextos? O, dicho de otra manera, ¿es posible contar con medidas nacionales o internacionales de logros de aprendizaje que sean equivalentes?

En realidad, cualquier persona involucrada en pruebas estandarizadas argumentará que sí, que se cuenta con mecanismos estadísticos que permiten verificar potenciales fuentes de sesgo asociadas al contexto (por ejemplo, el análisis del comportamiento diferencial de los ítems entre diferentes grupos de población), y eso es efectivamente así como resultado de un amplio trabajo desarrollado a lo largo de décadas (Hambleton 2005). Esto, si bien hace posible construir pruebas estandarizadas en la medida en que ello se funde en una técnica muy robusta (de ahí que los enfoques «simples» sean potencialmente muy peligrosos y nefastos desde el punto de vista de la desinformación que pueden generar), no lleva a desdeñar las consideraciones antes expuestas. No existe una prueba estandarizada «perfecta», no por defectos de la técnica (o de quienes las usan o

¹⁷⁹ De hecho esta es exactamente la tercera pregunta de la encuesta ALL: «Subraye la oración que explica por qué usted debe seguir cuidadosamente todas las instrucciones» y la respuesta esperada: «Subraya o encierra en un círculo la oración o parte de la oración que contiene “El fabricante no es responsable por fallas, lesiones o daños causados por un armado inadecuado”» (traducción propia), que se usaron en ALL de acuerdo a lo publicado en el sitio web del Centro Nacional de Estadísticas Educativas de los Estados Unidos (NCES) (<<http://nces.ed.gov/surveys/all/Items.asp?sub=yes&SectionID=2&CatID=3>>). Esto denota un nivel de etnocentrismo que difícilmente se podría esconder.

no) sino por la propia naturaleza de lo que se busca medir, que está entroncado en la cultura y los contextos.

Así, enfoques cuidadosos y robustos permiten contar con información estandarizada, es decir, información que se basa en medir lo mismo a través de distintos grupos humanos. Para ello, un primer requisito está dado por la capacidad de los operativos de medición para adaptarse al contexto. En efecto, un elemento que ilustra muy rápidamente si estamos o no ante un enfoque robusto lo da la actitud hacia la adaptación de los instrumentos de medición.

Un enfoque pobre insiste en la estandarización no solo de los **constructos** a evaluar sino de los propios instrumentos y sus detalles. Este enfoque olvida que procedimientos iguales en contextos diferentes llevan a resultados distintos. Así, a efectos de medir con propiedad los mismos **constructos**, lo que se requiere es un énfasis informado en la **equivalencia cognitiva** de las operaciones implicadas en las pruebas y no una rígida búsqueda de uniformidad. Claro está que en algunos casos no existe equivalencia cognitiva, dado que existen operaciones que son específicas de cada lengua, sistema de escritura o cultura¹⁸⁰.

Es claro que existe una tensión entre la búsqueda de información equivalente y la necesidad de adaptar los instrumentos para que sean significativos para las personas. Obviamente, no hay una regla general que resuelva estos predicamentos y en cada caso es preciso adaptar lo que sea necesario adaptar y conservar lo que se debe conservar, para, en ambos casos, lograr resultados equivalentes en los terrenos en los que esto sea posible. Por otra parte, se ha discutido estos temas con relación a pruebas de habilidades de la población adulta, lo que hace el tema aún más complejo debido a las grandes diferencias que pueden existir en trayectorias educativas entre distintas cohortes poblacionales. Sin embargo, mayor complejidad no quiere decir que los temas sean necesariamente distintos. Por ejemplo, en los sistemas educativos donde existe el hábito de rendir exámenes estandarizados, los estudiantes tendrán una relación con las pruebas diferente a la de los estudiantes de sistemas educativos en los que la evaluación tiene otra forma, lo cual podría sesgar los resultados comprometiendo la comparabilidad de los mismos.

Ahora bien, ¿para qué se quiere tener resultados equivalentes o comparables? La respuesta a esta pregunta tiene que ver con los propósitos de contar con información que represente en dimensiones comunes, en un solo país o internacionalmente, los fenómenos educativos. Así:

¹⁸⁰ Por ejemplo, conocer el uso de las mayúsculas y las minúsculas es necesario para ser un escritor competente en español y no, por ejemplo, en árabe, donde no existe esa diferencia. De modo análogo, la lectura del árabe requiere reconocer cada letra en sus diferentes formas (aislada, en la posición inicial, final o intermedia en una palabra e incluso cuando su escritura se omite), sin embargo, estas diferencias no existen en el español. De ahí que no hay forma de medir de modo equivalente este tipo de aspectos.

- Si la educación es un derecho universal, es preciso verificar si todas las personas ejercen dicho derecho. Usar medidas diferentes no lo permite y, por el contrario, implica el riesgo de establecer distintos raseros para distintas poblaciones, de modo que algunas tengan más derechos que otras (como el ejemplo de la lectura de cuarenta palabras por minuto cuando los hijos de aquellos que postulan esto leen cuatro o cinco veces más rápido). El informe Coleman (Coleman *et al.* 1966), justamente necesitaba este enfoque para ver problemas de desigualdades en las oportunidades educativas.
- Es necesario verificar si hay progresos, y por lo mismo se necesita información que sea comparable en el tiempo.
- El análisis que compara información para distintas unidades de observación (por sexo, grupos de edad, área de residencia, niveles de ingresos o países) permite identificar elementos que informan la discusión pública y la toma de decisiones. Ahora bien, ¿qué es comparar entre diferentes unidades de observación?

Existe una práctica muy difundida, especialmente en los estudios de medición de logros académicos de estudiantes, aunque también se observa en estudios sobre alfabetismo adulto, consistente en comparar mediante la construcción de *rankings*.

Este procedimiento en realidad es una de las formas más pobres del análisis comparado. El propósito de este no es definir un *ranking* sino observar cómo diferentes atributos de las distintas poblaciones se vinculan a los resultados, de modo que existan elementos para una rica reflexión acerca de lo que explica el mayor o menor desempeño.

Recordemos una frase sobre la que el Cockcroft Committee (Cockcroft 1982) llamó la atención: «Nadie se ha hecho más alto como resultado de haber sido medido». Es decir, la medición en sí misma es insuficiente, por eso es preciso que los operativos de medición de logros educativos (en estudiantes o en la población adulta) se vean acompañados por otros instrumentos que permitan enriquecer el análisis de los datos para que estos se puedan convertir, efectivamente, en un elemento valioso para la política educativa. Esto incluye la posibilidad de conjugar diversos métodos de investigación, de forma que se genere información de diversa naturaleza que enriquezca el análisis: por ejemplo, conducir observaciones etnográficas en paralelo a la evaluación¹⁸¹, realizar grupos focales con los estudiantes a efectos de discutir diversos temas vinculados a las áreas a evaluar o de conocer la forma como se condujeron durante la evaluación estandarizada, etc.

Sin embargo, los *rankings* parecen tener una presencia muy tenaz. Esto obedece a que, de acuerdo a algunas visiones, en sí mismos cumplen una función legítima.

¹⁸¹ El programa LAMP de la Unesco incluye entre su batería de instrumentos una guía para conducir observaciones etnográficas en paralelo a la evaluación. Este tipo de indagación fue usado durante la implementación de LAMP en Paraguay y en Mongolia (Maddox 2014, 2015).

Ilustremos esto de la siguiente manera: cuando uno participa en un evento deportivo como una carrera sabe que hay dos cosas que importan en los resultados: el tiempo y la posición obtenida. El tiempo nos permite medir el desempeño contra un rasero, digamos, absoluto y, por lo mismo, monitorear nuestro progreso frente a él; la posición, sin embargo, nos permite monitorear nuestro desempeño vis a vis el desempeño de otros.

Sin embargo, la educación no es una carrera entre escuelas, jurisdicciones o países, ¿o sí lo es? Este es exactamente el punto: para algunos sí lo es. De hecho, el propósito principal de los sistemas de información educativa a escala de la OECD es justamente mostrar la posición relativa de los países entre sí, dada la importancia de las variables educativas para la competitividad económica. Un enfoque centrado en la competitividad económica dota de sentido al *ranking*¹⁸².

Del mismo modo, si se considera que la educación debe ser manejada como cualquier otro negocio, es posible postular que la operación del mercado necesita, para funcionar bien, información. Así, un sistema de evaluación de logros puede convertirse (si cubre a todas las escuelas, es decir, si es censal) en un elemento de información a los «consumidores» que, entonces, tomarán decisiones informadas sobre dónde matricular a sus hijos. Este es el caso del Sistema de Medición de la Calidad de la Educación (Simce) chileno, creado por la dictadura militar en los años 1980 exactamente para cumplir ese propósito¹⁸³.

Uno puede o no estar de acuerdo con estos enfoques, pero son parte integrante del debate educativo contemporáneo. Es claro que la opción expuesta en este texto es diferente: la educación importa porque la gente tiene derecho a ella, entonces, no es subsidiaria a la competitividad económica y no puede ser tratada como un bien o servicio más en el mercado. De hecho, uno de los resultados más nefastos del Simce ha sido justamente el contribuir a crear una imagen de «mejores escuelas» a aquellas que logran mejores puntajes por la vía de seleccionar a los estudiantes (violentando sus derechos) con miras a, justamente, aparecer mejor en el *ranking*. Asimismo, la investigación educativa desde por lo menos los años 1960 ha mostrado con total claridad que los resultados de aprendizaje se explican no solo por variables dependientes de la acción de los actores educativos, sino también por otras, como el nivel socioeconómico de las familias. Así, no debe sorprender que las «mejores escuelas» sean justamente las que atienden a los sectores sociales más favorecidos.

¹⁸² En noviembre de 2010 el gobierno británico presentó ante el parlamento un documento que fundamenta su política educativa («White Paper»). Este literalmente parte de constatar el posicionamiento relativo del Reino Unido en pruebas internacionales estandarizadas para postular los objetivos de política: mejorar su posición frente a otras naciones «competidoras» (Reino Unido-Department of Education 2010).

¹⁸³ Aunque, debe anotarse, su evolución posterior ha permitido también abordar otras tareas más significativas, como la identificación de las escuelas en situación de mayor necesidad de intervención.

¿Cuán compleja o simple debe ser la medición de aprendizajes?

La discusión anterior es suficiente para sugerir que la medición de los aprendizajes de las personas no es una tarea sencilla que se pueda abordar sensatamente con instrumentos improvisados o mediante operativos sin mayores controles.

En este acápite, sin embargo, es necesario precisar algunos elementos conceptuales y metodológicos que deben ser suficientes para ilustrar las posibilidades de trabajo más pertinentes y dimensionar el tipo de desafíos implicados por ellas.

La primera tarea es definir con precisión qué se desea medir. El gráfico 19 (ver la página 207) brinda el punto de partida para esto en el caso del alfabetismo: ¿Se desea medir lectura, escritura y uso de números? ¿Sólo lo primero?, ¿lo primero y lo segundo?, ¿lo primero y lo tercero?, etc. Y si se trata del uso de números, ¿solo las habilidades mediadas por textos?

Evidentemente la respuesta depende de los propósitos de la medición así como de consideraciones conceptuales y operacionales. Por ejemplo, nadie negaría la importancia de medir la escritura, dado que esta implica habilidades activas de producción de textos y no solo de uso o recepción de los que vienen dados y, por tanto, son producidos por otros. Sin embargo, evaluar tareas de escritura en una forma comparable entre diferentes grupos poblacionales (dentro de un país o entre países) es una tarea de una inmensa complejidad, no imposible, pero demanda recursos y capacidades operacionales de una magnitud muy considerable.

Supongamos que se decide excluir la escritura del esfuerzo de medición (al menos en una primera etapa) y que este se enfoque exclusivamente en operaciones (incluso numéricas) que suponen la lectura de textos escritos. Esto puede perfectamente excluir de la evaluación las habilidades numéricas no mediadas por textos (que no hacen parte del alfabetismo pero que son de primera importancia para desarrollar las habilidades del uso de números mediado por textos).

Aquí, entonces, cabe preguntarse qué tipo de textos incluir o excluir (continuos, discontinuos, impresos, manuscritos, de ficción, de no ficción, etc.), lo mismo que el tipo de operaciones y los niveles de dificultad que deben tener. Esto consiste en desarrollar un marco de evaluación que especifique claramente el alcance de la medición y sus propiedades.

Tal marco, por su parte, no puede ser el fruto de un trabajo arduo y bienintencionado pero carente de sustento. La lingüística, la pedagogía y la ciencia cognitiva tienen mucho que decir para informar la construcción de un marco de evaluación y, en este sentido, es un trabajo de gran calibre y sobre el que reposa todo el esfuerzo de medición. Asimismo, se debe tener en cuenta en qué medida los marcos de evaluación deben o no corresponderse con los *curricula* nacionales o con los marcos nacionales de calificaciones que puedan existir.

A partir de estos marcos de evaluación es posible definir las especificaciones de lo que se ha de medir siguiendo una pauta como la de la tabla 16.

Tabla 16. Niveles de alfabetismo lector en PISA

	Obtención de información	Interpretación	Reflexión y valoración
Nivel 5	Localizar y posiblemente ordenar o combinar varios fragmentos de información que no resultan evidentes en absoluto, algunos de los cuales podrían encontrarse fuera del corpus principal del texto. Inferir qué información del texto es relevante para la tarea. Manejar información muy verosímil y/o abundante información en conflicto.	O interpretar el significado de un lenguaje lleno de matices o demostrar una comprensión completa del texto.	Valorar de manera crítica o formular hipótesis haciendo uso de conocimientos especializados. Manejar conceptos contrarios a las expectativas y hacer uso de una comprensión profunda de textos largos o complicados.
	Textos continuos: analizar textos cuya estructura no resulta obvia ni está marcada con claridad, para discernir la relación entre partes específicas del texto y el tema o la intención implícita en el mismo. Textos discontinuos: identificar las pautas existentes entre muchos fragmentos de información expuestos de manera extensa y detallada, a veces haciendo referencia a información externa a la exposición. Es posible que el lector tenga que percatarse independientemente de que para comprender por completo la sección del texto es necesario consultar otra parte distinta del mismo documento, como una nota al pie.		
Nivel 4	Localizar y posiblemente ordenar o combinar varios fragmentos de información que no resultan evidentes, que es posible que tengan que ajustarse a varios criterios, en un texto cuyo contexto o forma resulta habitual. Inferir qué información del texto es relevante para la tarea.	Utilizar un nivel elevado de inferencia basada en el texto para comprender y aplicar categorías en un contexto poco habitual e interpretar el significado de una sección del texto teniendo en cuenta el texto en su totalidad. Manejar ambigüedades, ideas contrarias a las expectativas e ideas expresadas de forma negativa.	Utilizar conocimientos públicos o formales para formular hipótesis o analizar de manera crítica un texto. Mostrar una comprensión precisa de textos largos y complicados.
	Textos continuos: seguir los vínculos lingüísticos o temáticos a lo largo de varios párrafos, a menudo sin nexos claros en el discurso, para localizar, interpretar o evaluar información que no resulta evidente o inferir significados psicológicos o metafísicos. Textos discontinuos: realizar una lectura rápida de un texto largo y detallado para encontrar información relevante, a menudo con muy poca o ninguna ayuda de elementos organizadores como marcadores o una maquetación especial, para localizar diversos fragmentos de información que deberán ser comparados o combinados.		
Nivel 3	Localizar y en algunos casos reconocer la relación entre distintos fragmentos de información que es posible que tengan que ajustarse a varios criterios. Manejar información importante en conflicto.	Integrar distintas partes de un texto para identificar una idea principal, comprender una relación o interpretar el significado de una palabra o frase. Comparar, contrastar o categorizar teniendo en cuenta muchos criterios. Manejar información en conflicto.	Realizar conexiones o comparaciones, dar explicaciones o valorar una característica del texto. Demostrar un conocimiento detallado del texto en relación con el conocimiento habitual y cotidiano o hacer uso de conocimientos menos habituales.

	<p>Textos continuos: utilizar convenciones de organización del texto, cuando las haya, y seguir vínculos lógicos, explícitos o implícitos, tales como causa y efecto a lo largo de frases o párrafos, para localizar, interpretar o valorar información.</p> <p>Textos discontinuos: tomar en consideración una exposición a la luz de otro documento o exposición distintos, que posiblemente tenga otro formato, o combinar varios fragmentos de información espacial, verbal o numérica en un gráfico o en un mapa para extraer conclusiones sobre la información representada.</p>		
Nivel 2	<p>Localizar uno o más fragmentos de información que es posible que tengan que ajustarse a varios criterios. Manejar información en conflicto.</p>	<p>Identificar la idea principal del texto, comprender relaciones, crear o aplicar categorías simples, o interpretar el significado con una parte limitada del texto cuando la información no es importante y se requieren inferencias sencillas.</p>	<p>Hacer una comparación o conectar el texto y el conocimiento externo, o explicar una característica del texto haciendo uso de experiencias y actitudes personales.</p>
	<p>Textos continuos: seguir conexiones lógicas y lingüísticas dentro de un párrafo para localizar o interpretar información; o sintetizar información a lo largo de textos o partes de textos para inferir la intención del autor.</p> <p>Textos discontinuos: demostrar que se ha captado la estructura subyacente de una exposición visual, como un diagrama de árbol, o combinar dos fragmentos de información de un gráfico o una tabla.</p>		
Nivel 1	<p>Localizar uno o más fragmentos independientes de información, generalmente ajustándose a un criterio, con muy poca o ninguna información en conflicto en el texto.</p>	<p>Reconocer el tema principal o la intención del autor de un texto sobre un tema habitual, cuando la información requerida es importante.</p>	<p>Realizar una conexión simple entre la información de un texto y el conocimiento habitual y cotidiano.</p>
	<p>Textos continuos: usar las redundancias, los encabezamientos de los párrafos y las convenciones de imprenta habituales para formarse una impresión de la idea principal del texto o para localizar información expuesta de manera explícita en un breve fragmento de texto.</p> <p>Textos discontinuos: centrarse en fragmentos de información separados, generalmente dentro de una única exposición, como un mapa sencillo, un gráfico lineal o de barras que tan solo presenta una pequeña cantidad de información de una manera sencilla y que en la mayoría de los textos verbales está limitada a un reducido número de palabras o frases.</p>		

Fuente: OECD (2006a: 64).

Una vez que se cuenta con especificaciones como las mencionadas, se procede a precisar el detalle de cada una de las celdas, como en la tabla 16, en términos de los niveles de dificultad esperados, el tipo de situaciones o contextos a los que debe referir la evaluación, etc. Con este nivel de detalle se puede proceder a construir ítems que midan cada una de las características que componen cada dimensión, representada en la tabla 16 por las columnas (la medición debe ser unidimensional, de lo contrario no se sabe qué es lo que se mide) a través de los distintos niveles de dificultad y contextos.

Un elemento a considerar es que una medición confiable necesita un número mínimo de ítems por atributo y nivel de dificultad (por celda en la tabla 16) y, por lo mismo, es necesario producir una batería grande de ítems, ya que probablemente algunos no funcionen bien y haya que descartarlos. Está claro que una «pruebita sencilla, no muy científica» se aparta de estas consideraciones de manera muy significativa.

Asimismo, es importante que los ítems, estén basados en ficción o en situaciones reales, se formulen pensando en el potencial impacto del conocimiento previo. Por ejemplo, veamos el siguiente caso:

Estímulo

Las ciudades pueden jugar diferentes roles en cada país. Por ejemplo, hay ciudades que son consideradas capitales, otras que son centros administrativos, otras que son centros financieros, etc.

En los Estados Unidos de Norteamérica, Washington DC es la capital y Nueva York, el centro financiero.

En el caso de Gulevandia, la capital es Wilferiburgo¹⁸⁴.

Ítem:

¿Cuál es la capital de Gulevandia?

Si a la hora de construir o adaptar este ítem a las condiciones de un determinado país, se decide reemplazar la información de Gulevandia por la propia, tendremos un problema: lo más probable es que las personas puedan responder a la pregunta de modo correcto **sin necesidad de leer el texto**. Este ítem, construido para evaluar la capacidad de la persona para recuperar información de un texto continuo donde dicha información está explícitamente contenida, y donde existe información competitiva, entonces, no serviría para nada. El conocimiento previo anularía completamente su capacidad para medir lo que se propone.

Adicionalmente, es necesario componer los instrumentos combinando los ítems de una manera que no introduzca sesgos (por ejemplo, no conviene colocar todos los ítems difíciles al final, ya que hay un efecto fatiga que afectaría de modo sesgado a los ítems de mayor dificultad por ir siempre en la posición final).

Hay diversas técnicas para componer cuadernillos de pruebas. Por ejemplo, se puede usar bloques y rotarlos, de modo que el mismo bloque vaya en diferentes posiciones en diferentes cuadernillos, o se puede producir un número de bloques mayor al que se usa en un cuadernillo cualquiera, de modo que los cuadernillos incluyan tanto ítems comunes como otros no comunes. Esto permite ampliar el universo de lo que se mide sin sobrecargar a cada respondiente singular, ya que, a fin de cuentas, no se busca generar estimaciones para cada individuo sino para poblaciones y grupos de población (salvo en el caso de las evaluaciones censales, donde, efectivamente, se busca computar un resultado para cada individuo).

Llegados a este punto es de fundamental importancia probar los ítems con una muestra (que sin ser necesariamente probabilística debe tratar de cubrir la diversidad de

¹⁸⁴ Gulevandia y su capital son entidades ficticias creadas por los músicos y humoristas argentinos Les Luthiers.

situaciones en las que se aplicará la prueba) de la población. Esta prueba piloto brindará información sobre el funcionamiento de los ítems y permitirá descartar algunos y ajustar otros. Asimismo, brindará información sobre la dificultad relativa de cada uno tal y como se puede observar de modo empírico y no solo como se espera dadas las consideraciones conceptuales con las que se les construyó.

Con esta información se procede a revisar los instrumentos (y también los procedimientos, de ser necesario) y recién ahí es posible pasar a conducir la evaluación.

Del mismo modo, es importante esclarecer cuál es el tipo de instrumentos que ha de aplicarse: ¿suponen estos brindar respuestas escritas? Si este es el caso (asumiendo que es una prueba de lectura y no de escritura), ¿cuál es el impacto de esta mediación de la escritura en el desempeño en la prueba?, ¿es mayor para los sectores con menores competencias lectoras?¹⁸⁵

Conducir la evaluación, por su parte, también supone otros elementos que tienen que ver con el diseño general. Así, se debe definir: si la evaluación será censal o muestral y, si es muestral, qué tipo de muestreo va a usarse; las formas de administración; la logística; el personal capacitado para la aplicación; la supervisión; el aseguramiento de la calidad y los controles de potenciales fuentes de sesgo; etc.

Y aún no hemos dicho nada de las técnicas implicadas en el análisis de los resultados: controles de calidad y depuración de datos; análisis psicométricos (¿qué modelos usar?, ¿solo considerar la dificultad de los ítems?, ¿también su capacidad de discriminación?; y, si hay preguntas de respuesta múltiple, ¿se buscará controlar el efecto de la posible adivinación?); el análisis que se hará para verificar si los ítems tienen un comportamiento diferenciado según distintos grupos poblacionales; también, dado que no interesa reportar información a nivel individual (como sí es el caso en pruebas de selección académica o en las evaluaciones censales, por ejemplo), qué procedimientos deben usarse para que las estimaciones den cuenta tanto de las pautas típicas como de la diversidad en la población¹⁸⁶, etc.¹⁸⁷

Complejo, ¿no?

Pues, efectivamente, lo es si uno quiere hacer bien la evaluación y aquí entra en juego un principio muy importante: es mejor no hacerla que hacerla mal. Es mejor no tener información y saber que uno no la tiene que tener mala información y confiar en ella: esto puede conducir a serios errores y, finalmente, destruirá la credibilidad del sistema de información no solo con relación a este tema sino que sembrará dudas incluso sobre lo

¹⁸⁵ De ser este el caso, las diferencias entre las personas en los niveles más bajos y aquellas en los niveles más altos de la escala se amplificaría por efecto de la mediación de la escritura.

¹⁸⁶ Este es el propósito de técnicas como las de computar para cada respondiente un conjunto de «valores plausibles» y no solo un puntaje singular (Von Davier *et al.* 2009).

¹⁸⁷ Postlethwaite (2004) presenta una excelente introducción sumaria a los diversos aspectos claves a tener en cuenta en mediciones de logro académico.

que sí se hace bien. La información de mala calidad no debe ser producida, ya que si se elabora se terminará por desecharla (dado que se suponía que uno quería hacer algo más barato, y algo que se tiene que tirar a la basura no es barato en absoluto, todo lo contrario, es terriblemente oneroso) o, lo que es aún peor, tal información puede conducir a tomar decisiones profundamente desacertadas que afecten la vida de las personas.

Nuevamente cabe preguntarse: ¿Por qué hay gente que tan alegremente sugiere que para tratar los temas educativos no se necesitan cosas hechas en serio? ¿Será porque no tienen la capacidad para entender y manejar enfoques complejos? Entonces, ¿esta actitud es solo una coartada frente a la propia incompetencia?

Por otra parte, también es posible encontrar situaciones en las que la necesaria complejidad de los esfuerzos es esgrimida como una coartada para proteger un interés comercial no siempre legítimo. No olvidemos que detrás de cada iniciativa hay personas que solemos tener diversos tipos de visiones e intereses.

En cualquier caso, lo que cabe siempre es usar el principio de parsimonia al que ya nos referimos anteriormente (ver la nota al pie 95), de modo que se asegure un diseño de investigación de la complejidad mínima necesaria, pero no menor que ella. Es decir, la aplicación de la navaja de Occam es siempre necesaria, pero se debe evitar la sobresimplificación del tipo «más breve, más pequeño, más barato» (Wagner 2003, 2011) que, al introducir un prurito sobresimplificador que no especifica dónde cabría simplificar, en realidad solo crea el riesgo de convertir la navaja de Occam en lo que podríamos llamar la «guillotina de Wagner».

¿Cómo informar mejor el diseño de las intervenciones?

Ahora bien, conocer que hay problemas en la distribución de las competencias deseadas (alfabetismo u otras) de las personas es un buen punto de partida para la formulación de políticas. Sin embargo, no es necesariamente uno suficiente.

En el caso del alfabetismo, saber que un determinado porcentaje de la población debería ser destinataria de programas orientados a desarrollar sus competencias de lectoescritura es importante, pero sería aún mucho mejor tener indicios de qué es lo que dicha población puede o no hacer con textos escritos. Partir del conocimiento previo permite identificar las áreas que pueden ser desarrolladas de modo más inmediato (por encontrarse próximas al conocimiento previo de las personas) y qué se requiere como pasos intermedios¹⁸⁸.

Lo anterior refiere a que un programa educativo bien diseñado es uno que parte de la necesidad de atender la diversidad de situaciones, incluida la de saberes previos. Es decir, las soluciones del tipo «*one size fits all*», si bien pueden servir en el caso de calcetines,

¹⁸⁸ Esta idea que es elemental en las ciencias de la educación contemporáneas se basa en lo que Vygotsky llamó «zona de desarrollo próximo» (1978).

no suelen ser una buena opción educativa. Esto, por supuesto, se aplica a cartillas de alfabetización de formato único que asumen que **todos** sus destinatarios son idénticos en términos de las competencias que (no) poseen solo porque respondieron que «no» a la antes mencionada respuesta censal.

Una evaluación de las competencias de alfabetismo de las personas puede permitir una caracterización más fina de los diversos grupos poblacionales que deben ser destinatarios de los programas orientados a mejorar sus competencias lectoescritoras, ya que posibilita ubicar las diferencias (en términos de competencias y características sociales) entre la que sería la población objetivo de dichas intervenciones. Lo mismo aplica a la medición de otros logros de aprendizaje.

Esto, por su parte, puede potenciarse aún más si los instrumentos de medición de alfabetismo incluyen secciones específicas destinadas a la evaluación de los ya mencionados **componentes de lectura**¹⁸⁹: reconocimiento de letras y de dígitos, pronunciación de letras y de sílabas, reconocimiento de palabras (frecuentes y no frecuentes), pronunciación de pseudopalabras, procesamiento de frases, lectura fluida de pasajes y comprensión lectora; elementos todos que pueden ser tomados en cuenta en una evaluación.

Ahora bien, varias de estas habilidades básicas que preceden a la lectura competente (que tiene un alto grado de automatización de los elementos mecánicos de la lectura) son específicas de la operación de cada idioma y sistema de escritura, por lo que es preciso considerar: (i) si las personas son efectivamente competentes en la lengua de evaluación; y (ii) que los resultados de medir estos aspectos no pueden necesariamente tomarse como comparables a través de diferentes lenguas o sistemas de escritura.

Con relación al primer elemento (competencia en la lengua de evaluación), es posible considerar la inclusión de ítems de vocabulario o de comprensión auditiva que no son parte de la medición del alfabetismo como tal, pero sí un prerrequisito a efectos de determinar si un bajo desempeño en la prueba es el resultado de un bajo nivel de competencias lectoras o de falta de manejo de la lengua, que es algo completamente diferente.

En referencia a lo segundo (no comparabilidad entre lenguas o sistemas de escritura), es posible mencionar algunos ejemplos:

- La decodificación de una palabra en una lengua como el español, que tiene la característica de ser ortográficamente **transparente**, es decir, donde hay una alta correspondencia entre símbolos y sonidos, no necesariamente informa demasiado o, en todo caso, informa de una manera diferente a lo que sucede en lenguas no transparentes, donde una misma sucesión de letras puede pronunciarse de modo muy distinto dependiendo de su significado (por ejemplo, «oo» en inglés puede

¹⁸⁹ Véase Perfetti y Rieben (1991).

sonar de una forma cercana a «o» en castellano –en *flood*– o a «u» en castellano –como en *food*–).

- Igualmente, no es lo mismo operar con lenguas alfabéticas que cuentan con 26 a 28 caracteres (como en el caso de las lenguas basadas en el sistema de escritura romano) que hacerlo con otras que tienen 48 caracteres o escritura logográfica con miles de símbolos disponibles.
- La fluidez en la lectura de una oración tiene una forma diferente de acuerdo a la estructura gramatical propia del idioma: no es lo mismo una lectura que sigue la pauta sujeto-verbo-objeto que una que sigue la pauta sujeto-objeto-verbo.

La introducción de elementos vinculados a los componentes de lectura permite ajustar el diseño de los programas de desarrollo de competencias de lectoescritura al brindar información relativa, por ejemplo, a:

- Si la población rural con lengua materna x (lengua ágrafa con una estructura de frases del tipo sujeto-objeto-verbo) muestra dificultades particulares en el procesamiento de frases y en la lectura fluida de pasajes en la lengua z , que es la lengua oficial.
- Si la población y no muestra familiaridad siquiera con el alfabeto usado en la lengua de evaluación.

Hecha toda esta discusión, la pregunta que resta abordar es: ¿con qué herramientas contamos para hacer todo esto?

Las últimas tres décadas han sido el escenario en el que se han desarrollado algunas experiencias de primera importancia en este terreno y en base a ellas debemos mencionar lo siguiente:

- Varios países han desarrollado encuestas nacionales sobre los niveles de alfabetismo de la población joven y adulta. Entre ellos podemos mencionar a título de ejemplo los casos de los Estados Unidos de Norteamérica (el Young Adult Literacy Assessment de 1985, el National Adult Literacy Survey de 1992 y el National Assessment of Adult Literacy de 2003), Canadá (y su rol central en las evaluaciones conducidas en países de la OECD en la década de 1990 e inicios del siglo XXI), Kenya (2006) y Bangladesh (2008).

Cada una de estas experiencias tiene sus propias características y no es posible brindar una revisión sistemática de sus alcances y limitaciones en este acápite, ya que esto implicaría una discusión conceptual y metodológica detenida sobre cada caso, lo que está, además, limitado por la información metodológica disponible (muy limitada en el caso de los dos últimos estudios mencionados).

- A nivel internacional, la OECD en colaboración con la Oficina Nacional de Estadística de Canadá (Statistics Canada) desarrolló el International Adult Literacy Survey (IALS)¹⁹⁰ en los años 1990 y su sucesora, la encuesta ALL¹⁹¹. Ambas estuvieron basadas en las experiencias previas en Canadá y, especialmente, en las de los Estados Unidos de Norteamérica¹⁹² y constituyen esfuerzos internacionales muy significativos en este terreno. Actualmente, la OECD está desarrollando el Programme for International Assessment of Adult Competencies (Piaac)¹⁹³.

Los esfuerzos de la OECD (como los de los Estados Unidos de Norteamérica y de Canadá) en este terreno han estado marcados por un énfasis en la importancia de las competencias de alfabetismo de las personas en el crecimiento y la competitividad económica de los países. De hecho, resulta significativo que el diseño de estos estudios excluye a la población mayor de 64 años por no ser parte de la población económicamente activa.

Es también importante notar el impacto de estos estudios sobre la forma como la OECD se involucró en la medición de competencias en escolares (población de quince años matriculada en educación secundaria) a través de su programa PISA, originalmente concebido como un complemento de los anteriores, ya que mientras estos se enfocarían en el *stock* de habilidades de la fuerza laboral, PISA daría cuenta de las habilidades de aquellos que están por incorporarse a ella (*intake*). La razón por la que se optó conducir PISA en escuelas es que la mayor parte de personas de quince años en la mayoría de los países de la OECD se encuentra matriculada en la escuela y, por lo mismo, es operacionalmente más fácil evaluarlas en escuelas.

- La Unesco, a través del UIS, inició en 2003 el programa LAMP¹⁹⁴, que también se encuentra basado en las experiencias de IALS y ALL, aunque por su propio diseño y por ser conducido por la Unesco incorpora necesariamente algunos elementos nuevos que lo diferencian de experiencias previas:
 - (i) Como ya se señaló a propósito de la discusión sobre la calidad de la educación, si bien la Unesco valora los efectos económicos de la educación, ve a esta desde una perspectiva más amplia basada en un enfoque de derechos.

¹⁹⁰ Para información de detalle sobre IALS, véase OECD y Canadá-Statistics Canada (2000).

¹⁹¹ Para información de detalle sobre ALL, véase OECD y Canadá-Statistics Canada (2005).

¹⁹² La génesis conceptual y empírica de estos estudios (basados en una exploración hecha con una muestra de jóvenes de diecisiete años en los Estados Unidos a mediados de los años 1980) es descrita por St. Clair (2012).

¹⁹³ El primer informe de Piaac fue publicado en 2013 (OECD 2013b).

¹⁹⁴ Para información en detalle sobre LAMP, véase Unesco-UIS (2009b) y Guadalupe y Cardoso (2011). Algunos problemas complejos de una iniciativa como LAMP son objeto de reflexión en Guadalupe (2015).

- (ii) Las encuestas IALS y ALL fueron conducidas únicamente en lenguas europeas usando el alfabeto romano. LAMP fue validado en diez lenguas pertenecientes a seis familias lingüísticas diferentes, usando tres sistemas de escritura y dos sistemas de numeración diferentes.
- (iii) La Unesco no está interesada en conducir encuestas para servir una agenda propia, sino en promover y facilitar que los países puedan hacerlo. Desde este punto de vista, LAMP no es un «estudio internacional» manejado por la Unesco, sino un esfuerzo de desarrollo metodológico (para poner las herramientas a disposición de los países) y de capacidades (para que los países puedan llevar a cabo autónomamente estos emprendimientos).
- (iv) Lo anterior lleva a que los esfuerzos de la Unesco impliquen generar instrumentos y procedimientos que pertenezcan al dominio público y no que sean propiedad industrial de un subcontratista privado¹⁹⁵.
- (v) LAMP es conducido como un esfuerzo nacional apoyado por la Unesco. Este apoyo implica, asimismo, que la Unesco juegue un papel importante en la validación de todo el esfuerzo cerciorando que se satisfagan estándares de calidad, lo que contribuye a asegurar la credibilidad de los datos.
- (vi) LAMP incluye la medición de competencias de lectura y uso de números, así como un módulo sobre componentes de lectura orientado a la población de menor desempeño.

Como puede notarse, existe ya un cúmulo de experiencias de las que los países pueden beneficiarse para desarrollar sus propios esfuerzos. Estas experiencias permiten contar, en algunos casos, con enfoques validados y con un importante trabajo previo de desarrollo metodológico que podría ser muy oneroso si cada país tuviese que replicarlo de modo independiente.

También las experiencias muestran limitaciones y dificultades de las que se puede aprender para seguir mejorando. En particular, es imperativo llamar la atención sobre las agendas políticas e ideológicas que subyacen al mundo de la evaluación de logros de aprendizaje (Cooper y Dunne 1998, 2000; Darville 1999; Guadalupe 2015; Hamilton y Barton 2000; Hamilton 2001, 2012; Street 1998).

¹⁹⁵ Si bien esta es una práctica legítima en el mundo de los negocios, impone restricciones inaceptables sobre la rigurosidad científica de los esfuerzos realizados, ya que impide la posibilidad de validación externa de procedimientos que no son públicamente accesibles. En este sentido es muy importante el énfasis puesto por la Learning Metrics Task Force en considerar la información como un **bien público** (2013a).

2.4.8.2 Aspectos adicionales de la medición de aprendizajes en la escuela: características socioeconómicas y culturales de los estudiantes y desarrollo de actitudes

Más allá del tema del alfabetismo adulto, la medición específica de aprendizajes en el medio escolar presenta algunas peculiaridades adicionales que es menester abordar.

En principio, se podría postular que la generación de información sobre las tasas de aprobación debería ser suficiente para conocer en qué medida los estudiantes logran los aprendizajes esperados en el año escolar que han cursado. De hecho, se supone que «aprobar» implica lograr los aprendizajes. Sin embargo, esto no siempre es así por diferentes razones (que incluyen el hecho de que los docentes evalúan de modo independiente, sin contar con una pauta única, lo que lleva a que los aprobados en un lugar podrían ser desaprobados en otro dado su rendimiento)¹⁹⁶ y existe evidencia que muestra que los niveles de logro tienden a ser inferiores que los que las tasas de aprobación sugieren.

Por otra parte, las tasas de aprobación no permiten conocer qué materias y cuáles aspectos del *currículum* no son logrados por los estudiantes y, en el mejor de los casos, solo posibilitarían un análisis vinculado a las propiedades de cada grupo de estudiantes y no de cada alumno individualmente.

Así, en las últimas décadas se ha introducido de modo creciente pruebas estandarizadas no destinadas a certificar a los estudiantes individualmente, sino a conocer y dar seguimiento con mayor precisión al comportamiento de los logros de aprendizaje.

En las páginas precedentes ya se ha tocado varios elementos centrales vinculados a la medición estandarizada de tales logros, que incluyen:

- Tipos de evaluaciones: basadas en normas o criterios (ver las páginas 211 a 213).
- Elementos centrales para la construcción de pruebas: constructos, ítems y estímulos, calibración de ítems, complejidad necesaria y riesgos de sobresimplificación, especificidad lingüística, estandarización y equivalencia de resultados, comparaciones y *rankings*, impactos no deseados del conocimiento previo, composición de cuadernillos de evaluación, pilotaje, etc. (ver la página 213 y ss.).

¹⁹⁶ Las «pruebas nacionales de salida de un nivel» buscan enfrentar este problema de modo que exista un momento en el tiempo en el que el desempeño de todos los estudiantes sea juzgado con un criterio común. Su existencia, por otra parte, sirve como guía a los docentes en términos de lo que se ha de evaluar al final del nivel. Esto puede tener un efecto no deseado complejo: que los docentes «enseñen para la prueba» y no para favorecer el aprendizaje. Por último, es importante considerar que la facultad de los docentes para evaluar de acuerdo a su propio juicio profesional puede cumplir un rol clave en atender con propiedad la diversidad de situaciones existente en la medida en que: (i) los docentes sepan hacerlo y (ii) cuenten con pautas claras sobre los aprendizajes que hay que lograr.

Por lo tanto, resta aquí referirse a otros temas específicos de la evaluación estandarizada de logros a nivel escolar.

Un primer tema tiene que ver con conocer **en qué medida las escuelas logran asegurar los aprendizajes independientemente de las características socioeconómicas y culturales de los estudiantes**. Si la educación es un derecho universal, es claro que todos deben alcanzar determinados logros previstos por el *currículum* nacional.

Ahora bien, como ha sido documentado desde hace varias décadas, empezando por el informe Coleman (Coleman *et al.* 1966), los resultados de los estudiantes (y las modificaciones en estos a lo largo del tiempo) no solo dependen de lo que haga el sistema educativo, sino también (y, en muchos casos, mayormente) de las características socioeconómicas y culturales de los estudiantes. Esto equivale a decir que los sistemas educativos no logran contrarrestar el efecto de dichas variables de modo que los aprendizajes sean equitativos. Esta situación lleva a preguntarse en qué medida los sistemas educativos terminan reproduciendo las desigualdades preexistentes y cuál es el grado de relevancia que tienen para producir aprendizajes.

El ejercicio presentado en el acápite sobre financiamiento precisamente reconoce el hecho de que los resultados dependen de satisfacer las necesidades de aprendizaje de los estudiantes (que pueden ser de naturaleza específicamente educativa o no), las mismas que se explican por diferencias socioeconómicas y culturales. Por ejemplo, para lograr aprendizajes equivalentes entre estudiantes que tienen como lengua materna la lengua oficial de enseñanza y otros que no, es claro que los procesos educativos deben ser diferentes y, por lo mismo, la dotación de recursos también. Si este no es el caso, lo más probable es que el sistema educativo reproduzca las diferencias, de modo que castigue a las poblaciones que ya se encuentran en una situación de desventaja¹⁹⁷.

En este sentido, los estudios sobre logros académicos son de crucial importancia para saber si los sistemas educativos contribuyen a desarrollar sociedades más equitativas o no y, como ya se anotó, usar esos estudios principalmente para construir *rankings* de escuelas (como en el caso chileno, tal como lo legó la dictadura) o para entregar «bonos» o incentivos al desempeño resulta ser falaz, ya que las «mejores» escuelas son, usualmente, las que tienen a los estudiantes más favorecidos y sus resultados tienden a ser más una expresión de esto que de sus características propias como escuela. Esto, por otra parte: (i) contribuye a la reproducción de la desigualdad en el plano simbólico, pues

¹⁹⁷ Por cierto, tener como lengua materna una lengua no oficial de enseñanza no debería definir una situación de desventaja, ya que cada quien tiene derecho a su propia cultura (de la que la lengua es una parte central). Sin embargo, las lenguas están siempre entroncadas con estructuras de poder (Ostler 2006) y cada lengua permite un universo de oportunidades distinto. Asimismo, no todas las lenguas tienen sistemas de escritura. Por lo mismo, existen diversas opciones para manejar este tema que suelen entender que lo mejor es asegurar una educación en lengua materna para todos, aunque se incluya también una transición hacia el manejo (incluida la lectoescritura) de la(s) lengua(s) que brindan mayores oportunidades.

fortalece el prestigio de algunas escuelas que no necesariamente son las que tienen mejor desempeño; (ii) impulsa a las escuelas a segregarse a efectos de «mejorar» en los *rankings* mediante el expediente de excluir a los que les harían bajar el puntaje promedio; y (iii) desvirtúa la práctica pedagógica al existir una presión por prestar una atención excesiva a los resultados en la prueba, de forma que modificar el indicador termina sustituyendo al objetivo (asegurar aprendizajes) y esta modificación puede lograrse de diversas formas (no solo mediante mejoras en los aprendizajes), que incluyen el entrenamiento para la prueba, el fraude, etc. Una fracción importante de la profunda y acertada crítica dirigida hacia las formas predominantes de evaluación estandarizada se explica exactamente por este tipo de problemas, que se exacerban cuando la evaluación tiene un carácter punitivo o con incentivos mal definidos¹⁹⁸.

Si se desea hacer una evaluación del desempeño de las escuelas se debería, entonces, considerar no los resultados absolutos de aprendizaje mediante pruebas estandarizadas sino la proporción de estos que explica la escuela. Para este efecto, se ha venido desarrollando estudios de «valor agregado», es decir, que comparan los logros de una escuela contra el logro esperado dada la composición social de los estudiantes.

Este enfoque tiene la virtud de tomar en cuenta los atributos de las personas, pero solo opera si es que las diferencias sociales explican las existentes en los logros educativos. Un sistema educativo que es capaz de lidiar con dichas diferencias en el origen de sus estudiantes no podría estimar el logro «esperado» dada la condición de los estudiantes, pues aquel tendería a ser independiente de esta.

Por otra parte, tal y como se discutió en el acápite sobre calidad de la educación, **los logros académicos no son la única variable que importa para evaluar el desempeño de los programas educativos**. De hecho, existen importantes esfuerzos de evaluación que van mucho más allá de esta variable a efectos de brindar una mirada de mayor alcance a la realidad de las instituciones educativas, generando evidencias que orientan las acciones a tomar¹⁹⁹.

En cualquier caso, no se debe perder de vista el propósito último de los sistemas de información: contribuir al debate público sobre la educación, a los procesos de toma de decisiones y a dar cuenta de modo transparente de un asunto que es de interés público.

¹⁹⁸ Otros elementos certeros de crítica se vinculan a la demasiado fácil asociación entre medición estandarizada y estandarización de la educación (lo que niega la necesidad de atender la diversidad y afirma visiones del mundo dominantes a expensas de expresiones culturales minoritarias o en desventaja). Finalmente, una parte no tan bien fundada de críticas obedece a una mala comprensión de la medición (porque la asocia a los problemas anteriores, como si estos fueran consustanciales a la medición) que usualmente se basa en una epistemología débil, la cual separa de modo insuperable la evidencia «cualitativa» de la «cuantitativa», lo que no es sostenible.

¹⁹⁹ Un excelente ejemplo de evaluación de las instituciones educativas es el que se verifica desde hace ya varios años en Escocia (Escocia-HM Inspectorate of Education 2007).

En este sentido, la construcción de *rankings* obedece a otro tipo de preocupación («brindar información al consumidor para que este elija la escuela a la que envía a sus hijos»²⁰⁰ o subrayar la competencia entre países) que puede ser considerado válido por muchos pero que es diferente de lo anterior.

Así, los estudios sobre logros académicos basados en pruebas estandarizadas pueden cumplir un rol clave en los debates públicos si son diseñados en correspondencia con los propósitos mencionados. Por ejemplo, el análisis comparado de los resultados (que no tiene por qué limitarse a los *rankings*) puede permitir la identificación de:

- Escuelas con desempeños mejores a los esperados dada la composición social de sus estudiantes. Esto supone que el estudio esté diseñado de una determinada manera, es decir, de una que permita obtener conclusiones a nivel escuela²⁰¹.
- Vacíos en los logros que sirvan para identificar qué áreas específicas del *currículum* deberían recibir mayor atención. Este es un aspecto de singular importancia, ya que puede brindar elementos concretos para mejorar el trabajo de los docentes en el aula. Lograrlo supone que el estudio sea diseñado de una manera que se alinee con los contenidos curriculares (lo cual es uno de los problemas principales que impide que PISA pueda dar mucha información relevante)²⁰².
- Variables educativas que se asocian de modo positivo con los resultados. Sin embargo, esta es un área cuyo análisis requiere un particular cuidado para no desdeñar aspectos de gran importancia solo por su falta de asociación con los puntajes.

Al mismo tiempo, es muy importante tener presente que las pruebas estandarizadas suelen estar concentradas solo en algunos aspectos de los logros que se espera que las personas alcancen gracias a su experiencia educativa. Usualmente, estos incluyen aspectos **cognitivos** en áreas como comunicación (o lenguaje), matemáticas y ciencias. Pero los propósitos de la educación no se agotan en estas y, por lo mismo, es importante considerar que es también necesario mirar otros aspectos, como el **desarrollo de actitudes**. En este orden de ideas, la diversa experiencia internacional de la IEA (Husén 1996) resulta de capital importancia, como también lo son los trabajos de Heckman (Heckman y Kautz 2012; Heckman *et al.* 2006).

²⁰⁰ Por cierto, este propósito es muchas veces una ficción, dado que para muchas personas no existen opciones entre las cuales elegir, sea por su lugar de residencia o por los costos directos y de transacción presentes. Así, solo algunos pueden ser «consumidores informados».

²⁰¹ Un ejemplo interesante está contenido en Unesco Santiago (2002a).

²⁰² Para una discusión detenida sobre PISA, véase Pereyra *et al.* (2011). Una mirada crítica a los problemas más políticos detrás de PISA es ilustrada por Loveless (2013a, 2013b).

El ejercicio desarrollado por Unesco Santiago (2008) con relación a operacionalizar los cuatro pilares de la educación puede sentar las bases para esfuerzos de medición de actitudes asociados a dichos pilares, ampliando de modo significativo el cuerpo de información que las pruebas estandarizadas generan. Esto se asocia también a la creciente conciencia acerca de las limitaciones y efectos perversos de enfocarse exclusivamente en saberes instrumentales (Learning Metrics Task Force 2013a; Robinson 2011; TED 2006). Este tema se explora en el siguiente acápite de esta sección.

2.4.9 LA MEDICIÓN DE LOS CONTENIDOS DE LA EDUCACIÓN

La tarea educativa no solo consiste en asegurar logros de aprendizaje en el terreno cognitivo y en algunas materias clave. Siguiendo lo desarrollado por Unesco (1996) y Unesco Santiago (2007), los propósitos de la educación tienen que ver con el desarrollo de la identidad (ser), en colectividad (vivir juntos), lo que está directamente asociado a la construcción de sentidos, al lado del desarrollo de capacidades humanas básicas que permitan actuar (hacer) y continuar desarrollándose (aprender).

Del mismo modo, la calidad de la educación implica la pertinencia de la labor educativa, y medir la pertinencia también representa un desafío mayor.

Ahora bien, ¿cómo medir la situación y el avance en temas como estos?

Lo primero que es preciso hacer es convertir estos grandes temas en aspectos asibles mediante un esfuerzo sistemático de medición, es decir, operacionalizar estos conceptos. Una de las principales contribuciones de Unesco Santiago es exactamente haber hecho esto tal y como se muestra en el anexo 1 (2008: 211 y ss.) de dicha publicación.

Este primer esfuerzo sistemático de operacionalización, si bien perfectible como cualquier otro, brinda un significativo punto de partida para tornar estos aspectos en realidades observables de modo sistemático. Una primera forma de hacerlo es la desarrollada en dicho documento consistente en un análisis documental (de políticas, regulaciones y provisiones curriculares) que busca dar cuenta de en qué medida: (i) esos pilares de la educación se encuentran efectivamente presentes en los elementos intencionales de las políticas educativas; y (ii) los servicios educativos se organizan alrededor de consideraciones básicas que los hagan pertinentes (valoración de la diversidad y participación; flexibilidad y adaptabilidad).

Ahora bien, que estos aprendizajes y prácticas estén presentes en las declaraciones y postulados de política y normativos no es una garantía de que se verifiquen como aprendizajes efectivamente logrados. Así, es conveniente postular un segundo nivel de observación vinculado a los **logros de aprendizaje** de los estudiantes y a las **prácticas educativas** en estos terrenos.

A estos efectos, las pruebas cognitivas brindan algunos elementos iniciales, sobre todo cuando se enfocan en explorar capacidades o competencias (que se vinculan al

saber hacer). En este sentido, las pruebas estandarizadas son muy valiosas y pueden ser situadas en un lugar que resulta más apropiado que el discutible postulado referido a que estas son la evidencia *par excellence* de la calidad educativa.

Sin embargo, es necesario ir más allá y explorar los aspectos de identidad (ser), actitudes relativas a la convivencia (vivir juntos) y disposiciones hacia el aprendizaje (aprender). Del mismo modo, se requiere explorar en qué medida las prácticas hacen efectivos (o no) los principios de valoración de la diversidad, participación, flexibilidad y adaptabilidad.

Una herramienta potencialmente muy valiosa en estos terrenos, y también originada en el mundo de la psicometría, refiere a la construcción de escalas actitudinales²⁰³.

Así, contar con una operacionalización de los cuatro pilares y de los elementos implicados en la pertinencia es un excelente punto de partida para iniciar esfuerzos de medición actitudinal que pueden incorporarse en los operativos de medición de logros académicos mediante instrumentos adicionales²⁰⁴.

Sin embargo, esta no es una tarea sencilla. La construcción de instrumentos para la medición de actitudes así como la administración de los mismos es una tarea de gran envergadura, ya que la confiabilidad de la medición está directamente ligada a: (i) el número de ítems que se utilizan para medir cada aspecto singular; y (ii) el número de casos que se consideran en la muestra.

En efecto, medir una sola actitud requiere una amplia batería de preguntas (ítems) y no es posible ni deseable exponer a cada estudiante a un conjunto demasiado grande de ellas. Por lo tanto, cada operativo muestral puede brindar la oportunidad para medir solo algunos aspectos actitudinales. En ese sentido, los operativos censales de medición de logros (cuya utilidad es relativa especialmente si se elimina el interés en la construcción de *rankings*)²⁰⁵ brindan oportunidades únicas, pues se puede utilizar baterías más amplias de instrumentos que, rotados con propiedad entre la población mediante el uso de bloques (cada estudiante recibe un cuadernillo con un set de ítems que corresponde a un número limitado de bloques de preguntas, pero el conjunto de la población recibe un conjunto diferente de cuadernillos que cubren un número de

²⁰³ Una breve descripción de cómo se construyen estas escalas se brinda en la cuarta parte del anexo 1.

²⁰⁴ Debe notarse que estudios nacionales e internacionales (como PISA o los conducidos por la IEA) incluyen elementos de medición de actitudes, aunque estos tienden a enfocarse en actitudes relativas a los logros cognitivos que ocupan los estudios (por ejemplo, actitudes respecto de la lectura). Lo que aquí se está postulando es la necesidad de medir actitudes por sí mismas, no como subsidiarias a otro tema, ya que el desarrollo de actitudes es parte de la tarea educativa.

²⁰⁵ Su mayor utilidad no está vinculada a la formulación de políticas como tal, sino a la transparencia de la gestión, ya que puede brindar información específica de cada programa educativo (su desempeño absoluto y comparado con otras unidades de gestión o jurisdiccionales) a la comunidad local. Sin embargo, dicha información solo se limita a los logros de aprendizaje y debería valorarse más la necesidad de enfoques más integrales, como el desarrollado en Escocia-HM Inspectorate of Education (2007).

aspectos mayor mediante un número amplio de bloques), pueden brindar información sobre muchos aspectos de modo simultáneo.

Por otra parte, al igual que con los logros cognitivos, el análisis de las actitudes implica preguntarse por la medida en la cual estas son el resultado de la experiencia escolar o de las vivencias en otros ámbitos de la vida social. Por lo mismo, es necesario prever el uso de instrumentos complementarios (por ejemplo, relativos al clima escolar y a las pautas democráticas y/o no democráticas con las que se organiza la escuela para evaluar si estas impactan, por ejemplo, en los aspectos de convivencia) que ayuden a aislar el efecto de la escuela, identificando los aspectos de cultura y prácticas institucionales que impactan sobre el desarrollo de las actitudes.

Del mismo modo, es posible considerar que de modo paralelo a la medición de logros cognitivos y actitudes se conduzcan esfuerzos de investigación más centrados en la comprensión de la interacción entre las personas mediante observaciones directas de la dinámica escolar. Así, etnografías de la vida escolar pueden brindar luces clave para entender con propiedad la información que se genera mediante otros procedimientos de investigación.

El trabajo realizado por Carnoy *et al.* (2007) a efectos de mejor entender por qué Cuba muestra diferencias importantes en logros cognitivos frente a los demás países latinoamericanos es un excelente ejemplo de cómo los métodos de observación pueden brindar información muy valiosa que hace palidecer la que algunos pueden generar corriendo, de modo fetichista, innumerables modelos matemáticos usados con muy poca perspicacia.

Finalmente, abordar el desafío de generar información sistemática sobre estos aspectos (midiendo actitudes a través del uso de procedimientos estadísticos y sumando los esfuerzos de la investigación etnográfica) contribuirá de modo importante a diluir la falsa y usualmente exacerbada brecha entre estudios «cualitativos» y «cuantitativos»²⁰⁶.

2.4.10 LOS DOCENTES

El tema docente es central en la operación de los sistemas educativos así como en el aseguramiento, o no, del derecho de las personas a la educación. Sin embargo, es poco lo que en general se sabe sobre los docentes y, en particular, sobre sus prácticas profesionales.

²⁰⁶ Esta frontera se basa en una tradición que diferencia de modo radical la **explicación racional** de la **comprensión interpretativa**, asumiendo que la primera es válida (y «objetiva» en el sentido decimonónico de la palabra) para los hechos naturales y la segunda (aunque inevitablemente «subjetiva») lo es para los asuntos humanos. La filosofía del conocimiento del siglo XX liquidó esta diferenciación al entender que toda empresa de conocimiento es una actividad humana basada en conceptos, representaciones e ideas que pertenecen al campo de la cultura (de la «subjetividad»). Sin embargo, nuestros no siempre bien informados mundos pretendidamente «académicos» siguen reproduciendo un discurso decimonónico carente de sustento.

Al mismo tiempo, existen discursos muy diversos sobre el tema de los docentes y la docencia, que van desde aquellos centrados en tratar a los docentes como un «recurso» (un insumo en el modelo tradicional) y se enfocan en aspectos vinculados a la dinámica del mercado laboral docente, hasta aquellos más bien marcados por discursos reivindicativos o sindicales que ponen el énfasis en «la situación de los docentes» y no en cómo estos contribuyen, o no, a través de su acción a garantizar, o no, el derecho de las personas a la educación que es, o debería ser, el foco del tratamiento público de los temas educativos.

Así, este acápite se centra en brindar elementos que ayuden a situar la discusión sobre la generación de información acerca de los docentes y la docencia en un marco que gira en torno a sus prácticas. Como toda acción humana, la docencia es un tema de gran complejidad y, por lo mismo, se requiere un enfoque lo suficientemente comprensivo y robusto que sienta las bases para su mejor tratamiento. Esto no es de ninguna manera equivalente a elaborar un listado sobre los diversos aspectos posibles que «afectan» a los docentes, sino más bien equivale a buscar un modelo de interpretación de su accionar como sujetos. Así, más que en proponer indicadores, este acápite se centra en tratar de definir el problema y sus componentes respecto de los cuales habría que avanzar en esfuerzos de medición.

La falta de información sobre los docentes y sus prácticas se explica no solo por la dificultad para generarla (por ejemplo, la observación sistemática de las prácticas docentes es un tema de gran complejidad) sino también por el debate acerca de hasta qué punto las autoridades públicas tienen derecho a hacerlo, ya que esto podría afectar la autonomía de la colectividad profesional. Así, esta discusión se enmarca en debates políticos sobre la necesidad de impulsar la transparencia y rendición de cuentas (*accountability*) por parte de los docentes en su accionar profesional y sobre cómo esta necesidad afecta o no el estatuto profesional de los docentes como colectivo profesional que requeriría, para un adecuado ejercicio, de autonomía y autorregulación.

Esta discusión, sin embargo, es una con una carga ideológica y política muy fuerte y no está exenta de importantes equívocos o desaciertos conceptuales. Por ejemplo, Ávalos (2004) sugiere que los cambios en las políticas educativas en Chile llevan a un movimiento de una situación inicial marcada por una identificación de los docentes como servidores públicos a una nueva en la que se trata de profesionales que deben rendir cuentas. Este argumento pierde de vista que la rendición de cuentas es, en primerísimo lugar, un atributo básico de la función pública, mientras que justamente la autonomía y la autorregulación son la bandera central de las colectividades profesionales.

En efecto, si existe una «carrera pública magisterial» esto implica necesariamente transparencia y rendición de cuentas (de lo contrario no es pública). Por otra parte, incluso si los docentes tuviesen un régimen laboral equivalente al del sector privado, su accionar debe ser objeto de escrutinio público, tanto por el hecho de ser financiado con el dinero de los contribuyentes, como porque la educación es un bien público.

Así, es claro que este texto, al enfocarse primariamente en el derecho de las personas a la educación y entender esto como un asunto de interés público, considera central que se genere evidencia sistemática sobre los docentes y la docencia de forma tal que el debate y la formulación de políticas sobre educación no pierdan de vista un elemento clave de toda la operación del sistema (la acción docente) en beneficio de una autonomía profesional mal entendida (se trata de un servicio público) o del interés de una corporación que, por muy legítimo que sea, no puede primar sobre el interés público y los derechos fundamentales de las personas.

Esta situación ilustra también un problema complejo de manejar, que es el potencial conflicto entre diversos intereses que pueden ser todos legítimos. No existe ninguna razón inmanente que haga que todo lo que pueda parecer deseable sea necesariamente compatible, por lo que una de las tareas de la política consiste en tomar decisiones sobre estos potenciales conflictos: lo importante, en este caso, es que dichas decisiones sean primordialmente informadas por el derecho fundamental de las personas y no por mecanismos de poder (como la capacidad de presión).

De hecho, un grupo importante de problemas complejos de la educación es el resultado justamente de este desbalance de poder en perjuicio del derecho de las personas. Ilustraciones de este desbalance se encuentran, por ejemplo entre otros muchos ejemplos posibles, en:

- (i) la independencia existente entre la dinámica de crecimiento de la fuerza laboral docente y la matrícula observable en varios países en las últimas décadas²⁰⁷;
- (ii) las normas sobre creación de nuevos puestos docentes que primaba en México hasta que el «Pacto por México» introdujo reformas en 2012, que se traducían en que el 50% de las puestos que se creaban se entregasen al sindicato docente para que este los asignase sin que exista ninguna regulación ni supervisión al respecto²⁰⁸;
- (iii) las normas que indican que la carrera docente debe iniciarse en instituciones educativas ubicadas en lugares como zonas rurales o urbano-marginales, ya que estas no hacen sino sancionar que tales lugares son poco deseables (para los

²⁰⁷ Por ejemplo, en el caso del Perú en el período 1906-2006 la matrícula en las escuelas primarias públicas creció en aproximadamente 22 a 25 veces, mientras que el número de docentes lo hizo en 47 veces (Guadalupe 2009). La información más reciente muestra una exacerbación de ese comportamiento, reforzando la tesis que afirma que las políticas de contratación de docentes son independientes de la prestación del servicio educativo.

²⁰⁸ Evidencia no sistemática sugiere que dichos puestos eran asignados fundamentalmente en base a dos criterios: (i) a hijos de docentes en ejercicio y (ii) mediante subasta. Cualquier comentario que uno pueda hacer respecto de esto (a partir de una preocupación por el interés público) palidece ante el abrumador carácter premoderno de este tipo de políticas y las pruebas de la corrupción que se han hecho públicas desde 2012.

- docentes) y, por lo mismo, son usados para «pagar derecho de piso» o «noviciado», implicando que es deseable dejarlos (ya que eso indica «progreso» en la carrera), cuando probablemente las instituciones educativas que atienden dichas poblaciones sean las que requieren de docentes más experimentados o, en todo caso, con mayores (o específicas) capacidades para lidiar con situaciones más demandantes;
- (iv) las normas que definen bonos salariales para los docentes que laboran en instituciones educativas con mayor volumen de matrícula, ya que esto no tiene ninguna relación con la carga de trabajo de los docentes (que haya más alumnos en la institución no significa que el docente tenga que atender más alumnos, eso depende del tamaño de las clases) y usualmente favorece a quienes laboran en zonas urbanas (donde se encuentran las instituciones con mayor matrícula), haciendo que las instituciones educativas que probablemente necesiten más sean aún menos atractivas (para los docentes). Estas normas son particularmente nefastas si van de la mano con las mencionadas en el punto anterior;
 - (v) las normas que brindan un número dado de días que los docentes pueden tomar como «libres» para atender sus propios asuntos, sin que medie una sola disposición acerca de cómo esas ausencias serán cubiertas a efectos de no ocasionar pérdida de horas lectivas a los estudiantes.

Es decir, los estatutos docentes muestran muchas evidencias de cómo la dinámica de negociación sindical (que en sí misma es completamente legítima) ha logrado «conquistas» que, por no existir un contrapeso que se preocupe por el interés público (al parecer los gobiernos no son muy efectivos en este terreno y no hay un sindicato de estudiantes que reclame de modo sistemático que se debe priorizar la garantía al derecho a una educación de calidad)²⁰⁹, se traducen en la existencia de «derechos» que más bien resultan ser prerrogativas corporativas contrarias al ejercicio del derecho fundamental de las personas a la educación.

Los párrafos previos deben servirnos para ubicar el difícil contexto en el que se tendría que desarrollar un esfuerzo significativo para contar con información sobre docentes y sugieren cuál es el tipo de barreras que se puede encontrar en el desarrollo de esta tarea.

Dicho lo anterior, cabe incluir en esta sección una reflexión conceptual que ayude a definir las bases de un sistema de información sobre docentes y docencia²¹⁰.

Definir el alcance y las características del conjunto de información a generar sobre docentes y docencia implica clarificar las diferentes dimensiones relativas tanto a los docentes como a la docencia que sean centrales para garantizar una educación de calidad

²⁰⁹ Una excepción notable es el caso ya mencionado de Chile en 2006.

²¹⁰ Los siguientes párrafos se basan en Guadalupe (2006, 2009), donde se encuentra el detalle de la revisión de la literatura y un análisis sobre la situación peruana.

para todos. Esto, por su parte, está relacionado con las diversas maneras como el tema docente es enfocado conceptualmente. Una revisión crítica de la literatura sobre docentes y docencia sugiere la existencia de tres grandes tipos de aproximaciones: (i) aquellas que, desde diferentes perspectivas, buscan identificar los **determinantes de la acción docente**; (ii) aquellas que se centran primariamente en la condición de los docentes como un **colectivo profesional**; y (iii) aquellas que buscan esclarecer los temas docentes combinando elementos que vienen de un análisis orientado a la dinámica institucional con elementos relativos a **las experiencias, vidas y percepciones** de los docentes.

Exploremos de modo sumario cada uno de estos tipos de enfoques a efectos de extraer lecciones que alimenten el diseño de los sistemas de información en lo relativo al tema docente.

- (i) **Determinantes de la acción docente.** Existe un conjunto amplio de autores que aborda el tema docente a partir de una preocupación subyacente común que puede reseñarse de la siguiente manera: si las acciones de los docentes son importantes a efectos de explicar los resultados que logran los estudiantes y las escuelas, resulta crítico identificar los elementos de la docencia que han de ser afectados por las políticas, de forma tal que se garantice un adecuado nivel de desempeño docente. Identificar lo que los docentes saben y pueden hacer se torna, entonces, un elemento clave para la formulación de políticas.

A partir de este razonamiento subyacente, diferentes autores han sido capaces de identificar elementos centrales relativos a la dinámica del mercado laboral y a la conformación de la fuerza de trabajo docente y aspectos institucionales sancionados a través de los marcos regulatorios. Este conjunto de trabajos subraya la importancia del reclutamiento (por ejemplo, ¿cuáles son las características socioeconómicas y académicas de aquellos que son reclutados y retenidos?) y de las reglas de carrera, incluyendo los salarios y las condiciones de trabajo, tanto en sí mismas como en comparación con otros grupos ocupacionales.

- (ii) **La condición profesional de los docentes.** Un segundo conjunto de trabajos se centra en el carácter profesional de la docencia. Al hacerlo, consideran dos perspectivas profundamente imbricadas en el debate político: (a) los temas de colegialidad y autonomía como atributos claves que definen a una profesión y (b) cómo abordar los temas del desempeño y la rendición de cuentas.

Estas opciones se encuentran presentes de modo particular en los debates que se suscitan en países con una fuerte tradición de autonomía local y docente, la misma que aparece afectada por la introducción de esquemas de gestión más centralizados. Un ejemplo típico es el Reino Unido donde, por ejemplo, la introducción de un *curriculum* nacional es un fenómeno relativamente reciente (de los años 1980). Por otra parte, también existen reflexiones más centradas

en la crítica a los sistemas educativos por estar organizados más en función de intereses corporativos que en función del interés público²¹¹. Así, estos debates están directamente vinculados con temas que involucran conflictos de poder.

- (iii) **La combinación del análisis institucional con aspectos de las experiencias, vidas y percepciones de los docentes.** Este tercer grupo de trabajos busca darle mayor «concreción» a las descripciones más generales, usualmente centradas en los aspectos institucionales o estructurales, mediante la captura de los elementos subjetivos que son necesarios para una mejor interpretación de la manera como los docentes decodifican los atributos institucionales o estructurales y los incorporan en su propia práctica. Al hacerlo, estos trabajos toman en cuenta que las personas siempre tenemos un conocimiento de nuestro entorno y del por qué actuamos de una forma dada.

Un elemento clave en esta literatura subraya la importancia de la identidad docente para entender su acción y el porqué de esta. Asimismo, considera a la identidad como un proceso permanentemente inacabado de continua conformación en un contexto particular.

La revisión de estos tres tipos de enfoque permite identificar algunos temas centrales que deberían ser abordados por los sistemas de información:

- (i) Las dimensiones institucionales, las dimensiones subjetivas y la interacción entre ambas a través de la práctica son dominios ineludibles para entender la práctica docente.
- (ii) Lo anterior se traduce en la necesidad de combinar elementos que pueden ser llamados «estructurales» con aquellos más directamente vinculados a las percepciones de los propios docentes.
- (iii) Los elementos «estructurales» pueden ser observados utilizando información macro o de contexto, así como mediante información sobre el perfil sociodemográfico de los docentes, mientras que las percepciones requieren de un esfuerzo de interacción directa con los docentes a través de métodos de investigación que subrayen los elementos interpretativos. De esta forma, los primeros son aquellos más fácilmente abordables por los sistemas de información educativa, mientras que las segundas pertenecen al reino de la investigación educativa que los sistemas de información deben conocer y con la cual deben dialogar a efectos de mejorar de forma continua.

²¹¹ Ciertamente, esta es una forma de privatización de la educación pública usualmente obviada en las discusiones sobre la educación en la región. Si bien es debatible la conveniencia o no de permitir la operación de agentes privados con fines de lucro en la educación, es claro que esta otra forma de privatización es, fuera de toda dudas, perniciosa.

Estas consideraciones sugieren pistas importantes para generar de modo sistemático información sobre los docentes y la docencia. Tal información puede describirse a partir de tres temas: (i) la constitución de los docentes como agentes; (ii) los arreglos institucionales; y (iii) la práctica docente– de la siguiente manera²¹²:

(i) Elementos de información vinculados a la constitución de los docentes como agentes

- (a) **Características sociodemográficas de los docentes.** Bajo este rubro se incluye aspectos como el estatus socioeconómico de los docentes y de los postulantes a tales (¿de qué sector de la sociedad se recluta docentes?) y el nivel de escolaridad de los padres de los docentes (¿cuál es la trayectoria social de los docentes?).

Cabe preguntarse acerca de la configuración de la fuerza laboral docente en contextos, por ejemplo, donde son mayoritariamente la primera generación con acceso a educación terciaria, residentes en zonas en las que la única oferta de educación terciaria es la docencia o donde la docencia aparece como la opción de educación terciaria menos exigente en términos académicos, etc.

- (b) **Estatus social y condiciones de vida.** Ingresos individuales por trabajo a tiempo completo²¹³ comparados con: uno, el ingreso percibido por otros grupos ocupacionales según nivel de escolaridad alcanzado; dos, los niveles de riqueza media del país y línea de pobreza. También el ingreso de los hogares con jefe docente comparado con los otros hogares (según nivel de escolaridad del jefe del hogar); asimismo las percepciones de los docentes

²¹² Cada aspecto de la realidad puede implicar un universo muy amplio de aristas y detalles; a efectos de seleccionar y organizar de modo significativo cuáles son los aspectos centrales que tienen que observarse, la discusión conceptual es de capital importancia. La revisión previa justamente permite organizar los siguientes párrafos. De otra forma, se podría componer un listado no estructurado de temas mediante la agregación de inquietudes, por ejemplo, las presentes en la acción reivindicativa de los gremios docentes, como las reproducidas en Robalino (2005), Vaillant (2005) y Murillo (2005) con limitado valor analítico.

²¹³ Es importante que se sea cuidadoso con relación a: (i) ingresos anuales versus ingresos mensuales; en varios países diversos grupos ocupacionales perciben un número diferente a doce remuneraciones mensuales por año, por tanto, es mejor estandarizar estos valores, sea anualizando o mensualizando los ingresos totales, y no tomar una única cifra sin prestar atención a estos posibles escenarios; (ii) ingresos brutos y netos; ¿qué debe contabilizarse como ingresos y qué no? Por ejemplo, las contribuciones al fondo de pensiones pueden representar un ingreso adicional para los docentes no disponible para otros grupos ocupacionales (dado el nivel de precariedad del empleo), aunque no sea un ingreso de libre e inmediata disponibilidad (en la práctica puede ser una especie de ahorro forzoso que si otros sectores ocupacionales quisieran tener deberían deducirlo de su ingreso regular); (iii) jornada de trabajo y días efectivos de trabajo al año; ¿qué cuenta como tiempo completo?

respecto de su estatus social y trayectorias sociales; finalmente, el prestigio de la profesión docente.

- (c) **Formación docente preservicio.** Nivel de escolaridad requerido para ser docente según nivel educativo del programa y posibles diferencias entre programas gestionados por el sector público o agentes privados; duración típica de los estudios requeridos para ser docente; orientación de los programas de formación docente (pedagogía y contenido específico).

Estos aspectos son objeto de intensos debates que impactan en las decisiones de política, referidos a: ¿qué nivel de certificación debe tener un docente?, ¿cuántos años de formación inicial son necesarios?, ¿cuál habría de ser la mejor combinación de aspectos vinculados al conocimiento de la materia y cuestiones pedagógicas?, ¿cuál habría de ser la mejor combinación de aspectos teóricos y elementos más directamente vinculados a la práctica?, ¿la formación inicial docente debería priorizar la formación en la práctica –mediante trabajo en aula y sistemas de acompañamiento o mentoría– o la formación pedagógica académica?²¹⁴

Por otra parte, estas discusiones no son ajenas a otros elementos que tienen que ver con el estatus y reconocimiento social de la profesión docente y, por lo mismo, con los niveles de satisfacción de los docentes. Asumir que la docencia requiere, por ejemplo, formación universitaria con un alto contenido académico equivalente al de otras profesiones sugiere afirmar que los niveles de ingreso y prestigio social asociados a la docencia debieran ser similares a los de otros grupos ocupacionales con formación tenida por equivalente. ¿Qué pasa si esto no es así?

Asimismo, es importante explorar si la certificación se corresponde con los niveles efectivos de conocimientos, habilidades y destrezas logrados por quien recibe la certificación.

- (d) **Expectativas de carrera.** Es un lugar común referirse a la docencia, al menos en el sector público, como una «carrera»; sin embargo, ¿cuál es el significado de esto?, ¿cuáles son las perspectivas de trayectoria profesional u ocupacional de los docentes?, ¿existe un vínculo entre dichas trayectorias y aspectos de desempeño, formación en servicio, antigüedad u otros factores?, ¿cuál es la relación entre opciones pedagógicas y de gestión educativa?, ¿existe una jerarquía vinculada al tipo de programa educativo en el que se presta servicios?

²¹⁴ Un ejemplo de abordaje de estos temas se encuentra presente en Lewin y Stuart (2003). Se trata de un informe que presenta la síntesis de un conjunto de investigaciones conducidas en diversos países del mundo. Los diferentes informes y otros detalles de esta investigación se encuentran disponibles en <<http://www.sussex.ac.uk/cie/projects/completed/muster>>.

Estas preguntas son importantes, ya que constituyen un conjunto de «señales» que informan al docente acerca de lo que el sistema de regulaciones considera valioso en la docencia. Por ejemplo, normas acerca de que el inicio de la carrera debe darse en zonas rurales o urbano-marginales indican, de modo implícito, que en estas zonas se requiere menor experiencia y que el curso «normal» de la carrera consiste en salir de esos lugares hacia otros que resultan más «deseables». Esto, evidentemente, tiene consecuencias complejas sobre los niveles en los que efectivamente se brinda atención a las necesidades educativas de las personas que es preciso ponderar.

Del mismo modo, considerar que la docencia de aula es un primer peldaño en la carrera y que los peldaños «superiores» se vinculan más bien a la gestión (ser director de centro sin carga lectiva o funcionario de la administración educativa) supone el mismo contenido implícito que torna la docencia de aula en menos «deseable» o de menor nivel de reconocimiento que otras opciones vinculadas a la gestión. Asimismo, esto tiene supuestos y consecuencias respecto del tipo de formación que se requiere para diferentes funciones.

Este conjunto de señales informa las representaciones que los docentes se forman sobre el entorno institucional en el que se desenvuelven y, por lo mismo, impactan en la acción de estos tanto en términos aspiracionales como de valoraciones de determinadas prácticas y acciones.

Es posible encontrar que las regulaciones de la carrera pueden tener relaciones favorables o desfavorables con el ejercicio del derecho a la educación o, incluso, pueden no formar un cuerpo coherente de mensajes.

En una investigación llevada a cabo a finales de los años 1990, el BID encontró en diversos países de América Latina cuerpos regulatorios que contribuyen a formar una realidad que Navarro describe en los siguientes términos:

- La profesión tiende a atraer predominantemente [...] a los [...] de menores calificaciones académicas.
- Para un buen número [de los aspirantes a docentes] la docencia no fue la primera preferencia [...].
- Con pocas excepciones, los maestros finalizan su adiestramiento profesional sin tener calificaciones adecuadas para enseñar [...].
- El Ministerio de Educación no realiza una buena selección entre el conjunto de candidatos aspirantes a ocupar plazas docentes [...].
- [...] una vez dentro de la profesión, la gestión de personal se caracteriza por una ausencia casi completa de instrumentos de evaluación y gestión del recurso docente [...].
- Los maestros trabajan con poco o ningún apoyo humano y/o material [...].
- Se observa una proliferación de actividades de capacitación en servicio [...al parecer estas son] la respuesta [...] para tratar de reparar las fallas originadas en la formación inicial [...].

- [...] quienes abandonan la profesión más temprano tienden a ser los mejor calificados, con lo cual el efecto neto de la estabilidad laboral es promover largos períodos de ejercicio profesional por parte de un grupo de maestros que tiende a ser de menor nivel de calificación a medida que avanza el tiempo.
- [...] una alta proporción [de los docentes] tiene un segundo trabajo para complementar el ingreso [...].
- [...] un grado considerable de compromiso del personal docente con su profesión y con los niños [...].
- [la regulación sobre jubilaciones] podría hacer insostenible el presupuesto del Ministerio de Educación [en el futuro cercano] (Navarro 2002: 2-4).

(ii) **Elementos de información vinculados a los arreglos institucionales**

- (a) **Regulaciones vinculadas a las condiciones de trabajo.** Duración prescrita y efectiva de la jornada de trabajo²¹⁵, cuántos empleos tiene un docente, si la docencia es un trabajo de tiempo completo; duración del año escolar y períodos vacacionales para los docentes, licencias, beneficios sociales (fondo de pensiones, protección de salud, acceso a créditos, número de remuneraciones mensuales por año, etc.), estabilidad laboral. Todos estos aspectos deben ser vistos en comparación con otros grupos ocupacionales de acuerdo al nivel de escolaridad y de ingresos percibido.
- (b) **Políticas vinculadas a las condiciones de trabajo.** ¿Qué variables definen el número de personas a reclutar?, ¿cuáles definen quiénes son reclutados?, ¿quiénes reclutan docentes?, ¿cuáles variables determinan las remuneraciones: desempeño, antigüedad, formación, etc.?, normativa sobre la salida del servicio (¿quién y por qué se despide a un docente?) y el número de personas que dejan la docencia según motivos (jubilación, despido, renuncia voluntaria, transferencia a función administrativa, etc.).
- (c) **Formación docente continua.** Junto con la formación inicial de los docentes, existe una creciente tendencia a subrayar que esta, por buena que sea, siempre es insuficiente y que es preciso, como en cualquier otra profesión,

²¹⁵ ¿Qué debe considerarse como un trabajo de tiempo completo? Si bien cada ocupación puede tener una norma distinta sobre esto que debe ser tomada en cuenta, es también preciso considerar que en el caso de la docencia se suelen tener en cuenta horas de preparación de clase u otras. Por otra parte, debería verificarse si esto es efectivo y comparar la carga horaria efectiva (corrigiendo el impacto de las «horas pedagógicas», donde las hubiere) con las de otros grupos ocupacionales. En un escenario extremo, la carga laboral anual docente puede ser equivalente solo al número de horas efectivas de clase (digamos unas 800 horas al año), cuando en otros grupos ocupacionales la norma puede bordear las 48 horas semanales (es decir, en cuatro meses se trabaja una cantidad de tiempo equivalente a la correspondiente a un año en un docente). Esta diferencia puede, perfectamente, estar a la base de los contrastes salariales.

que los docentes tengan experiencias regulares de formación que acompañen el ejercicio profesional.

Estas experiencias de aprendizaje continuo pueden tener relación con la introducción de modificaciones en las políticas educativas generales (por ejemplo, un cambio de contenidos o enfoques curriculares) que llevan a percibir que los docentes deben ser capacitados para poder implementarlas; pero también pueden más bien tener un carácter de contribución general al ejercicio de la docencia en diversas áreas. El primer caso tiene un propósito más instrumental, el segundo tiende a percibir al docente como un profesional autónomo al que se le brindan herramientas para que, según su propio juicio, las adquiera y utilice en su quehacer profesional.

En cualquiera de los dos casos, siempre cabe preguntarse por la naturaleza de las acciones de formación (asociadas o no a iniciativas gubernamentales, pedagógicas, de gestión, etc.) y por la vinculación (o su falta) entre el desarrollo profesional que la formación lleva aparejada y la carrera formal. Asimismo, resulta interesante preguntarse por qué la formación profesional continua tiende a aparecer más como una obligación gubernamental que como una preocupación profesional autónoma de los docentes individuales o de sus colectivos profesionales, como es el caso de otras profesiones. Esta reflexión permite dimensionar el lugar (muchas veces central) que la acción estatal tiene en la formación de las identidades docentes.

- (d) **Cultura institucional.** Las regulaciones institucionales tal y como son leídas por los agentes, así como el posicionamiento de cada uno de ellos frente a su contexto, sus pares y otros agentes, devienen en la conformación de una cultura institucional que marca el entorno en el que se desarrollan los aprendizajes.

Dado lo anterior, factores clave de la cultura escolar (definidos con relación a los contenidos de la educación) requieren ser conocidos y documentados, tanto en general como con relación a su impacto específico en el quehacer docente: ¿qué tipo de cultura(s) institucional(es) tienden los docentes a construir?, ¿cuáles son los mensajes entroncados en dicha cultura respecto de lo que se espera de los docentes?

Un ejemplo que podría sonar manido (aunque tiende a obviarse en el mundo de la educación) es el que vincula negativamente las normas laborales indiferenciadas con desempeños negativos. Un entorno institucional donde un buen desempeño es remunerado igual que un mal desempeño, donde la promoción solo depende de la antigüedad o de relaciones clientelísticas (con la gestión gubernamental o con los sindicatos), donde la permanencia en el

trabajo es independiente del desempeño, donde las faltas no son sancionadas, etc. es un entorno laboral que favorece la mediocridad, ya que:

- i. Los docentes **con mejores desempeños** son sancionados mediante un mensaje que indica que ese no es un comportamiento que reditúe económicamente y que incluso termina incomodando a los demás y, por lo mismo, el buen maestro termina siendo visto por la mayoría como un personaje «incómodo».
- ii. Aquellos **con un sólido juicio moral** terminan siendo asqueados por el manejo corrupto de las promociones, rotaciones, contratos, etc.
- iii. Estos **buenos docentes**, por tanto, terminan buscando otras opciones laborales a las que los docentes incompetentes y/o corruptos no pueden aspirar, de donde la probabilidad de permanecer en la docencia se torna más elevada para este segundo grupo.

Por el contrario, una cultura institucional (reforzada por regulaciones) que reconoce el desempeño y sanciona la incompetencia o la corrupción es una que tiende a retener a los mejores docentes y a expulsar a los que no deberían estar al frente de los aprendizajes de los estudiantes.

Es un lugar común, que al parecer no se toma muy en serio a la hora de definir las normas sobre docentes, afirmar que hay muchos aspectos (especialmente los actitudinales y valorativos) que se aprenden más mediante la interacción con el contexto, los ejemplos y las prácticas que caracterizan a la escuela como espacio de socialización que mediante el «dictado» de contenidos curriculares. Claro está que distintas culturas institucionales, como las descritas en estos párrafos, favorecen diferentes aprendizajes y que, desde el punto de vista de la educación a la que tienen derecho las personas, hay ciertos atributos de la cultura institucional que son indeseables. Los sistemas de información pueden ayudar a identificar situaciones deseables a fortalecer y no deseables a enfrentar.

- (e) **Trayectorias laborales.** Anteriormente se hizo mención a las normas sobre carrera magisterial, queda aquí mencionar que también es preciso verificar los comportamientos reales (y no solo los normativamente sugeridos) de los docentes. De hecho, la comparación entre las trayectorias y la normatividad puede echar luces sobre inconsistencias o patrones que sugieran modificaciones (para fortalecer o eliminar) normativas existentes.
- (f) **Satisfacción laboral.** Esta es una variable muy importante como elemento que se encuentra en la base del desempeño (junto a la competencia y a las condiciones de trabajo) y que tiene un peso actitudinal muy importante

que afecta el entorno del aprendizaje. Un docente motivado (competente y respaldado por las condiciones de trabajo) tiene una probabilidad mayor de abordar exitosamente las necesidades de aprendizaje de sus estudiantes. En este terreno, resulta importante distinguir aspectos que pueden tener un impacto positivo o negativo sobre la motivación (como el reconocimiento profesional o su falta) de aquellos factores cuya presencia puede no motivar, pero cuya ausencia puede tener un efecto devastador sobre la moral de las personas. Por ejemplo, contar con sanitarios adecuados no es algo que motive a nadie, pero no contar con ellos es un elemento que hostiga a las personas.

Resulta de capital importancia determinar los potenciales impactos (positivos o negativos) que estos arreglos pueden tener sobre el derecho de las personas a la educación.

Por ejemplo, en el País B los docentes tienen derecho (de acuerdo al estatuto docente) a no asistir a su trabajo sin exposición de causa (ya que el estatuto asume algunas causas consideradas legítimas, como asistir a reuniones gremiales, atender problemas familiares o de salud, etc.) hasta por un total de diez días en un ciclo lectivo anual. La pregunta que surge entonces es: ¿cuáles son las provisiones legales que existen en paralelo para asegurar que dichas inasistencias no se traduzcan en que los estudiantes dejen de percibir el servicio al que tienen derecho? De no existir ninguna disposición al respecto (como es el caso en el País B), debe concluirse que el interés corporativo de los docentes es tenido como más importante que el derecho fundamental de las personas a la educación, lo que constituye una clara inversión de prioridades en las políticas públicas, lo que seguramente obedece a que el gremio docente tiene una mayor capacidad de presión que la de aquellos interesados en defender el derecho a la educación (si es que hubiese alguien interesado en esto en dicho país).

(iii) Elementos de información vinculados a la práctica docente

Finalmente, los distintos aspectos antes mencionados se entrelazan en contextos específicos en los que los docentes ejercen sus labores. Este ejercicio es el que, en último término, contribuye a que las personas aprendan (o no) a definir qué y cómo aprenden.

Ahora bien, el ejercicio de la docencia implica al menos tres niveles diferentes:

- (a) **La práctica pedagógica.** ¿Qué es lo que efectivamente hacen los docentes con relación a los aprendizajes de los estudiantes? Responder a esta pregunta implica desde aspectos formales de cumplimiento (horas efectivas de clase, de preparación y de evaluación frente a las prescritas para dichos propósitos)

hasta aspectos vinculados a la transformación del *curriculum* prescrito en el *curriculum* efectivamente administrado en el aula, tanto con relación a los objetivos cognitivos como a los sentidos mayores de la tarea educativa.

Es de capital importancia, para mejorar los sistemas educativos, preguntarse por la medida en la cual los docentes garantizan que los estudiantes cuenten con las oportunidades de aprendizaje a las que tienen derecho y definir cuáles son los factores que facilitan o entorpecen que este sea el caso. Algunos de estos factores se vincularán a la competencia profesional de los docentes, mientras que otros tendrán relación con las condiciones en las que ejercen su profesión; diferenciar ambos aspectos es clave, ya que no hacerlo puede llevar a ineficacia en las acciones, sea porque se espera que los docentes hagan algo que no saben hacer (por ejemplo, al introducirse nuevos contenidos curriculares que los docentes no conocen) o que no pueden hacer por carecer del soporte pedagógico o administrativo o de materiales para ello.

Asimismo, es importante identificar las oportunidades de aprendizaje profesional que tienen los propios docentes y verificar en qué medida estas se aprovechan y los factores que facilitan o entorpecen dicho aprovechamiento.

- (b) **La práctica como colectivo profesional.** Si bien la práctica profesional en cualquier ámbito contiene un elemento clave de desempeño y responsabilidad individual, también es cierto que existen espacios claves para la mejora que están dados por la existencia del intercambio entre pares mediante equipos docentes y colectivos profesionales más amplios. ¿Cuentan los docentes con esos espacios de desarrollo y mejora profesional? ¿Cuáles son los factores que favorecen o limitan su existencia y operación?

En muchos países, los docentes están asociados en gremios fundamentalmente para propósitos reivindicativos o de gestión de algunas prerrogativas corporativas. Sin embargo, esto no es idéntico a contar con organizaciones o espacios de **colegialidad profesional**.

No se está sugiriendo acá la necesidad de crear de modo forzoso «colegios profesionales» que en muy pocos casos cumplen funciones profesionales y se limitan más bien al espacio corporativo fijando barreras de acceso a la profesión para proteger a los agremiados²¹⁶.

²¹⁶ ¿En qué medida los comités de ética de los colegios profesionales protegen a las personas o, más bien, son un espacio que evita ventilar y sancionar las responsabilidades derivadas de la mala práctica profesional? O, dicho de otra forma, ¿cuántos casos de suspensión de la licencia profesional por mala práctica han sido sancionados por los colegios profesionales? ¿Son estos todos los que debieran ser? Allsop y Mulcahy (1998) muestran claramente la operación de un **espíritu de cuerpo** entre los médicos que se corresponde con esta lógica corporativa.

Lo que sí se sugiere es que un elemento clave de cualquier profesión está dado por el intercambio entre pares. ¿En qué medida este existe en la docencia? ¿Quién lo promueve?, ¿las instancias de gobierno?, ¿los propios docentes? ¿Por qué se promueven o no se promueven?

- (c) **La participación en la formulación de políticas.** Un tercer elemento de la práctica docente que también toca a los docentes como sujeto colectivo es el relativo a la formulación de políticas. En este terreno cabe primero que nada preguntarse por el sentido de esa participación. La educación es un asunto público y, por lo tanto, las políticas deben ser definidas en espacios públicos de debate en los que el colectivo profesional docente –dada su especialización profesional– puede hacer una contribución (como también la pueden hacer otros sectores).

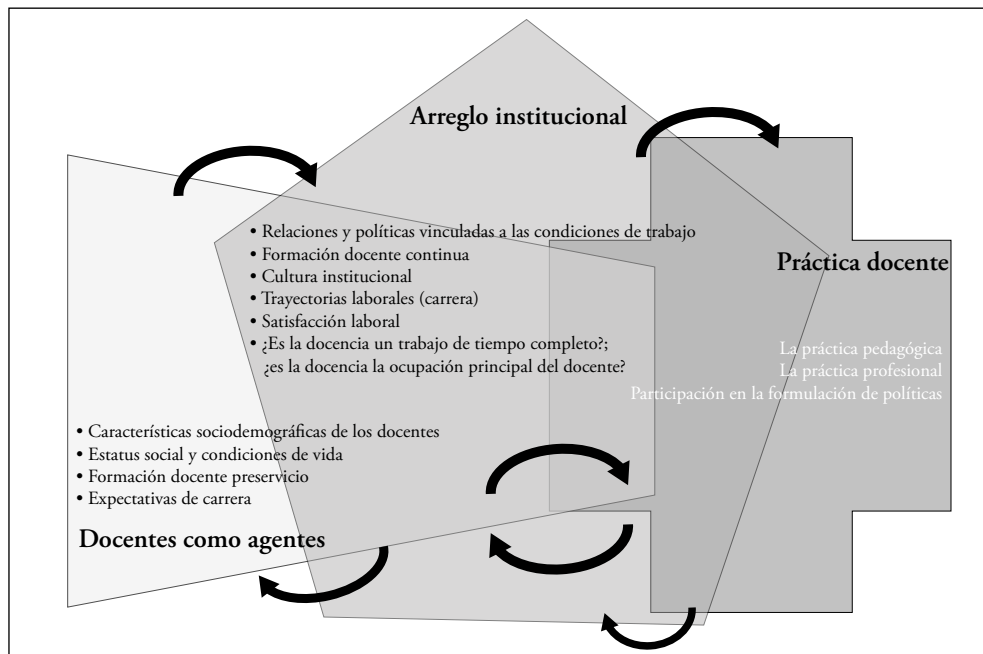
Ahora bien, existe un potencial problema derivado de que dicha participación puede no ser profesional (de los docentes como portadores de un conocimiento especializado), sino en tanto corporación que porta un conjunto particular de intereses. Si bien estos pueden ser perfectamente legítimos, existe el riesgo de que se sobrerrepresenten en la política, al punto de que opaquen la centralidad del interés público en el aseguramiento del derecho a la educación.

Como se mencionó anteriormente, un problema central es que muchas veces la política educativa está abocada a gestionar escuelas y contratar maestros y no necesariamente a contribuir a que las personas aprendan. En ese sentido, una sobrerrepresentación del interés corporativo magisterial constituye un peligro en tanto privatiza (a favor de la corporación) aquello que es de interés público²¹⁷.

La discusión anterior permite postular que un sistema de información sobre docentes podría organizarse de la manera propuesta en el gráfico 22.

²¹⁷ Es necesario un análisis detenido de los estatutos docentes a efectos de determinar en qué medida constituyen en términos reales (y no a nivel de los discursos) las instancias más claras de privatización de la educación pública.

Gráfico 22. Representación gráfica de un modelo para desarrollar información sobre los docentes y la docencia



Fuente: Adaptado a partir de Guadalupe (2006, 2009).

SECCIÓN 3

MEJORANDO EL SISTEMA

La introducción de mejoras, es decir, de cambios que pretenden incrementar la capacidad de servicio del sistema, supone una evaluación cuidadosa de los costos y beneficios de las mismas. Todo cambio supone el riesgo de interrumpir series históricas, necesidad de capacitación adicional u otros costos que incluyen una cuidadosa documentación así como sus potenciales impactos en las publicaciones de datos sobre las que se debe alertar consistentemente.

Asimismo, el establecimiento de mecanismos de control y verificación en toda la «cadena de producción» de información es un elemento básico que implica la definición de procedimientos y normas para tratar con situaciones problemáticas de un modo consistente. El manejo de la omisión, de la inconsistencia interna, de la inconsistencia temporal, de la validez o veracidad de los datos, etc. son temas que hacen parte del sistema de controles que es necesario establecer a cada nivel.

La presente sección busca brindar una visión panorámica de diferentes aspectos vinculados al mejoramiento de los sistemas de información. La selección de temas (ya que podrían abarcar un universo muy extenso) se ha hecho sobre la base de lo observado como preocupaciones en la experiencia de diversos países, principalmente de América Latina.

3.1 CONSISTENCIA

Un primer elemento a considerar tiene que ver con los niveles de consistencia de la información. Una información sólida es consistente con la realidad que busca representar y existe un conjunto de procedimientos que ayudan a abordar este tema.

Por otra parte, cuando se habla de consistencia, se hace referencia a diversos ámbitos en los que se espera que la información sea coherente. A continuación se discuten los tres ámbitos centrales a los que alude esta discusión: consistencia interna, temporal y entre diversas fuentes.

3.1.1 CONSISTENCIA INTERNA

Los diversos elementos de información que derivan de **un único evento de generación de información** deben ser coherentes entre sí.

Los controles más simples que los sistemas de información suelen incluir caen en esta categoría: asegurar que los totales marginales de tablas de dos entradas sumen los mismos totales generales; asegurar que totales de diversas tablas independientes pero que refieren a magnitudes comunes sean iguales; asegurar que los controles lógicos se hayan mantenido (por ejemplo, que un programa educativo femenino no reporte matrícula masculina; que un centro educativo que opera en un local ubicado en la ciudad no aparezca como rural, etc.).

El momento propicio para anticipar estos controles es durante el diseño de los instrumentos de generación de información (cédulas, manuales e instrumentos de carga de datos). Asimismo, es conveniente minimizar las oportunidades en las que tales inconsistencias se puedan presentar evitando preguntar lo mismo en más de una oportunidad, a menos que esta sea una decisión deliberada destinada a identificar posibles sesgos en el reporte de datos. En cualquier caso, el procedimiento para manejar redundancias (y posibles discrepancias que aparezcan) debería ser establecido en el mismo momento en el que se decide introducirlas.

Usualmente, los procedimientos de crítica y codificación de instrumentos así como los programas de carga de datos son los momentos en los que típicamente se introducen estos controles.

Cabe destacar la importancia de las labores de crítica que están destinadas justamente a identificar potenciales problemas antes de proceder a la carga de datos. Estos procedimientos de crítica son de capital importancia para el manejo de potenciales inconsistencias y, mientras más temprano se hagan en el tiempo (es decir, mientras

menor la distancia temporal entre el momento de reporte y el momento de la crítica) y más cercana en el espacio sea la conducción de este proceso con el reporte, mejor (es decir, idealmente la crítica debería ser un proceso descentralizado que facilite la devolución de instrumentos con problemas a la unidad de observación para un nuevo reporte –si son datos declarados– o una nueva administración, de ser el caso).

Los procesos de carga pueden incluir algunos controles redundantes respecto de los contenidos en la crítica (lo que a su vez permite verificar la confiabilidad de esta) y controles adicionales que suponen cálculos más complejos. Sin embargo, es importante notar que en algunos casos la inclusión de los controles de consistencia en el procedimiento de carga de datos puede producir un entrapamiento en este, ya que tal proceso se puede ver estancado por la necesidad de esperar la resolución de los conflictos. Una opción es que los controles de consistencia solo registren los problemas, pero no bloqueen la carga. Esto, si bien agiliza el proceso de carga, implica el riesgo de que la base de datos mantenga información «sucias» cuya consolidación y limpieza final pueda ser diferida de modo indefinido.

En general, los controles de consistencia interna pueden agruparse en dos grandes categorías: (i) aquellos que hacen referencia a imposibles lógicos y que, por lo mismo, muestran información que no puede ser correcta desde el punto de vista de la observación y registro de la realidad; (ii) aquellos respecto de los cuales no es posible definir su nivel de representación de lo real.

Por ejemplo, contar con dos tablas de matrícula con diferentes totales para una misma observación es un imposible en la realidad (y, por lo mismo, un imposible en términos lógicos); por lo tanto, al menos uno de los dos valores debe ser falso. Sin embargo, esta comparación por sí misma incluso si no mostrarse inconsistencia (ambos datos iguales) no permite definir si los dos valores observados son una adecuada representación de la realidad.

Así, desde el punto de vista exclusivo de la consistencia interna, es posible encontrar datos consistentes entre sí, pero inconsistentes con la realidad. Es por esta razón que el análisis de la consistencia interna es insuficiente para garantizar la consistencia de la información y se requiere otros enfoques, como los que se describen a continuación.

3.1.2 CONSISTENCIA TEMPORAL

Se refiere a que los elementos de información observados en una serie temporal tengan un comportamiento **consistente con las tendencias reconocibles** en la realidad a lo largo del tiempo.

Así, los comportamientos de un indicador a lo largo del tiempo deben mostrar patrones conceptualmente identificables y empíricamente razonables dado el conocimiento que se tiene de la realidad mediante otra evidencia.

Por ejemplo, si en el año t se cuenta con 325.000 estudiantes de seis años de edad matriculados en el grado 1, no resulta razonable encontrar más de 325.000 estudiantes de siete años en el grado 2 en el año $t + 1$. De hecho, la matrícula de siete años en el grado 2 (asumiendo que no hubo matrícula precoz en t) será igual a 325.000 - (defunciones + abandonos + repetidores en el grado 1 de siete años). La única excepción a esta regla lógica estaría dada por la observación de un saldo migratorio positivo a favor de la unidad de observación en dicha edad y en dicho período que compense (en un sistema cerrado, ya que podría incluso superar la cantidad de matrícula en un sistema abierto) las otras fuentes de potencial pérdida de matrícula.

La consistencia temporal no implica necesariamente que los datos sean estables en el tiempo, sino que sus cambios no vayan en un sentido lógicamente imposible, como el ejemplo mostrado, o en contra de lo que otra evidencia sugiere. Por ejemplo, un incremento de 15% en la matrícula entre dos años consecutivos en un nivel dado (cuando el incremento típico observado para dicho nivel en los últimos diez años ha sido de 3% anual) puede, y de hecho debe, llamar la atención, pero no por ello se tiene que asumir inmediatamente que es un error. De hecho, si existe evidencia de un cambio de política que puede explicar dicho incremento, ese valor puede resultar perfectamente razonable. De modo análogo, un valor que no muestra cambios no necesariamente indica ausencia de dificultades, esto dependerá de la existencia de factores que pueden sugerir una variación que, por alguna razón, el sistema de información no está captando.

Así, la consistencia temporal no solo implica observar variaciones inesperadas, sino también falta de variaciones que no debería darse.

Por ejemplo, en un país en el que las tasas de cobertura para las edades de la educación primaria son altas y no se espera que varíen mayormente, donde se ha hecho un esfuerzo reciente por disminuir las tasas de repetición y donde el volumen de la población en edad de ingreso se mantiene estable o empieza a declinar, es de esperar que la matrícula total del nivel tienda a disminuir por el efecto conjunto de la reducción de la repetición que cambia el perfil de la matrícula y del volumen absoluto de ingreso en descenso.

¿Qué pasaría si las series de datos no muestran ninguna caída, sino más bien magnitudes estables? Un sistema de información mal preparado no reaccionaría frente a la ausencia de cambios y no identificaría ningún problema. Un sistema preparado empezaría a indagar qué es lo que está sucediendo y podría encontrar que, por ejemplo, los docentes y directores de los programas educativos han empezado a reportar una matrícula abultada manteniendo los números previos o han empezado a incorporar estudiantes de menores edades para no «perder» matrícula (en realidad, para no perder sus puestos de trabajo o bonificaciones por tamaño de matrícula). Esta es una situación perfectamente posible en situaciones en las que existen bonos salariales determinados por el volumen de la matrícula. No se sugiere aquí que esos bonos sean impropios (podrían serlo, pero esa no es la discusión en ese punto), sino que estos se convierten

inadvertidamente en una fuente de problemas que el sistema de información debería ser capaz de detectar.

3.1.3 CONSISTENCIA ENTRE DIVERSAS FUENTES

Este tipo de consistencia se refiere a que la información originada por un evento específico (censo, encuesta, etc.) sea **coherente con otras observaciones independientes** del mismo fenómeno o de eventos relacionados.

Supongamos, por ejemplo, que se cuenta con información sobre la llamada tasa de supervivencia al grado final de educación primaria de los últimos quince años y se observa que los valores del indicador (calculado mediante el modelo de cohortes reconstruidas) han permanecido más o menos estables oscilando entre 83% y 85%. Al mismo tiempo, la información sobre la proporción de la población de quince a diecinueve años que ha culminado la educación primaria muestra incrementos a lo largo de los últimos quince años y presenta valores superiores al 90%.

El análisis de los dos datos en forma relacionada debería llamar la atención por dos razones:

- La tasa de supervivencia debería estar asociada a la conclusión, por lo tanto, ambos fenómenos deberían covariar. Si este no es el caso, se tiene problemas con al menos uno de los dos indicadores.
- El porcentaje de personas que culmina la educación primaria debería ser inferior al de estudiantes que alcanzan el grado final (lo que aquí no es el caso), ya que:
 - la tasa de supervivencia solo considera a los que ingresaron al sistema y no a toda una población dada; en el supuesto de que toda la población ingrese, la tasa de supervivencia establece (teóricamente) el límite superior posible de la proporción de personas que alcanza un grado dado; y
 - los que culminan la educación primaria son aquellos que alcanzan el grado final y continúan sus estudios hasta al menos la culminación de la primaria, de donde la tasa de supervivencia establece (teóricamente) el límite superior de este valor.

Entonces, hay un problema de consistencia entre datos diferentes (pero relacionados) y originados en fuentes distintas (censo escolar y modelo de cohortes; dato poblacional de un censo o encuesta). La determinación acerca de cómo resolver este problema pasa por identificar cuál de los dos datos que se están comparando resulta más sólido o confiable. Como se mostró previamente, en el caso de América Latina hemos podido observar esta misma inconsistencia y determinar que los datos de las encuestas de hogares, tomados en varios momentos en el tiempo, resultan ser consistentes entre sí y lo son también

con los datos censales, lo que sugiere sólidamente que las tasas de supervivencia tienden a subestimar el fenómeno que pretenden observar. Esta es una de las razones que ha llevado a la necesidad de revisar los supuestos y criticar los alcances de los modelos de cohortes reconstruidas tal y como se plantea en la segunda sección, acápite 2.4.2 (ver la página 147 y ss.).

El análisis de consistencia (especialmente la que importa en último término: la consistencia con la realidad) supone tener una comprensión profesional compleja de los fenómenos que son objeto de medición por el sistema de información. No basta con la técnica estadística, se necesita el conocimiento de la realidad educativa y esto implica prestar permanente atención a la investigación y política educativas, así como a los procesos sociales que afectan a la educación²¹⁸.

²¹⁸ Un caso ilustrativo de problemas en este terreno viene del informe de tendencias en el desempeño lector medido por PISA en 2010 (OECD 2010). Al observar los cambios experimentados entre 2001 y 2009, este nota distintas variaciones entre países y, consciente de que parte de estas se puede explicar por cambios en la población objetivo, procede a «ajustar» dichas estimaciones de cambios. En el caso peruano (el Perú era el país con mayor nivel de avance en puntos en la escala de lectura), el informe concluye que el progreso observado debía ajustarse levemente porque las condiciones socioeconómicas de los estudiantes peruanos en 2009 eran peores que las de los estudiantes de 2001 (OECD 2010: 50). Esta afirmación resulta llamativa, dado que todos los indicadores socioeconómicos del país sugieren lo contrario. Así, al revisar la fuente de esta afirmación, se encontró que en todos los indicadores usados para estimar cambios socioeconómicos no había, para el caso peruano, cambios significativos entre 2001 y 2009 (lo que sugiere que no se están midiendo bien), excepto con relación a la edad media de los estudiantes que en 2009 era ligeramente menor. Ahora bien, ¿cómo así una mejora en los niveles de atraso escolar sugiere que las condiciones socioeconómicas son más adversas? Claramente, los autores siguieron mecánicamente un procedimiento e interpretaron el resultado sin contrastarlo con la realidad. Una constatación de ese tipo hubiese llevado a concluir que la forma como se mide las condiciones del entorno no resulta consistente con la realidad. Este tema es discutido en detalle en Guadalupe y Villanueva (2013).

3.2 OMISIONES Y ESTIMACIONES; PROYECCIONES

3.2.1 OMISIONES Y ESTIMACIONES

En la primera sección se habló extensamente sobre el tema de la omisión, planteando que esta es un hecho regular en los sistemas de información, el cual, si bien se trata de minimizar, usualmente tiene una magnitud diferente de cero que es preciso conocer y manejar.

Existen dos formas distintas de omisión. La primera refiere a la ausencia de un número dado de unidades de información en un conjunto dado de información; por ejemplo, programas educativos que no reportan ninguna información en un censo escolar. La segunda refiere a que, si bien existe información de la unidad de información para el conjunto dado, ella no está completa porque el proceso mediante el cual se generó omitió algunos elementos por diferentes razones posibles.

En el primer caso es preciso, en primerísimo lugar, ser capaces de definir el subconjunto del universo de unidades de información que ha sido omitido. En las circunstancias en las que corresponde, la mejor manera de hacer esto es mediante el uso de directorios, tal y como fue descrito en la primera sección (página 49 y ss.); por ejemplo, contar con un directorio actualizado de programas educativos permite verificar cuáles se encuentran omisos al reporte de información en un censo escolar. Si es posible proceder a esta identificación, entonces el problema de la omisión de unidades queda resuelto, pues el universo total resulta conocido; si ello no es posible, entonces es necesario estimar la cantidad de unidades de información que se encuentran en condición de omitidas.

La estimación de unidades omitidas sin contar con un directorio de partida es una tarea muy compleja, ya que cualquier estimación necesita reposar en información relevante. Así, en ausencia de un directorio, es necesario apelar a otras variables que puedan ayudar a hacer una estimación razonable y provisional del universo de unidades de información. Por ejemplo, si se cuenta con datos que sugieren que existen centros educativos en las mismas jurisdicciones en las que existen instalaciones de salud, es posible recurrir al directorio de instituciones de salud e identificar las jurisdicciones en las que estas existen pero se carece de registros de programas educativos. De modo análogo, si existen normas que indiquen que debe establecerse un programa educativo en localidades que cuenten con al menos un número de habitantes determinado, un catálogo de centros poblados con información de población puede contribuir a dicha tarea. Así, es posible encontrar otros fenómenos (para los que se cuente con información) y usarlos como pistas para estimar omisiones. Sin embargo, como ya se anotó, este es un primer paso provisorio

que debe ser seguido por la verificación de lo estimado y, con ello, llegar a un directorio final más adecuado.

Una vez que el universo de unidades de información está plenamente definido²¹⁹, queda la tarea de estimar información para las unidades omitidas, para lo cual se sigue procedimientos similares a los que se aplican a la estimación de información faltante para unidades que cuentan con información incompleta, lo que nos lleva al segundo caso mencionado.

En el segundo caso, sobre estimación de información faltante, es preciso desarrollar un conjunto de procedimientos que, en términos generales, siguen los siguientes pasos:

- (i) Definir cuál es la información imprescindible que debe ser estimada. Esto supone decisiones acerca de los elementos de información prioritarios respecto de los cuales una opción razonable es hacer una estimación (que por muy bien hecha que esté es siempre una estimación).

Por ejemplo, contar con números finales de matrícula total del país necesita que se estime omisos a efectos de no enturbiar el dato agregado final con la variación en las tasas de respuesta. Además, en la medida en que la fracción de la información total que se estima no es demasiado alta, los errores de estimación no afectan significativamente el agregado general.

Por otra parte, si lo que se necesita es conocer el número de alumnos por clase y las omisiones se distribuyen homogéneamente a lo largo de todas las unidades de información, estimar los elementos necesarios para calcular esta razón no sería estrictamente necesario, ya que el resultado final sería el mismo.

- (ii) Analizar los patrones de comportamiento de la información a estimar y de las unidades de información que muestran omisión al respecto. Esto es de especial importancia para determinar si las omisiones de información siguen algún patrón que haga necesario seguir pautas de estimación diferentes para subgrupos de unidades de información.

Son dos tipos de factores que se tienen que considerar en este punto: (a) si la omisión se distribuye aleatoriamente a lo largo de la población o si más bien hay algún tipo de sesgo en la omisión al reporte de datos²²⁰; y (b) si la variable omitida se comporta de modo homogéneo a lo largo de las unidades de información.

Por ejemplo, el primer caso corresponde a variables que son omitidas, digamos, por programas educativos de gestión privada. ¿Por qué omiten esta información?

²¹⁹ Nótese que siempre es posible que haya ciertos márgenes de incertidumbre en el resultado final, por lo que se debe asegurar en todos los casos lograr el mejor resultado posible y ser, al mismo tiempo, consciente de que este puede ser perfectible.

²²⁰ Tómese en cuenta que se usa el término «población» no para referirse a personas necesariamente, sino al conjunto total de unidades de observación que componen el universo bajo observación.

¿Existe alguna motivación o estímulo derivado del entorno regulatorio que impulsa en esta dirección? ¿Tal impulso tiene un sentido particular hacia el subregistro de la variable?

El segundo caso, por ejemplo, corresponde a variables que, si bien pueden ser omitidas por unidades pertenecientes a distintos subgrupos poblacionales (por ejemplo, escuelas urbanas y rurales), no tienen un comportamiento similar en cada situación concreta, ya que se vinculan con otros fenómenos que operan de modo diferenciado a través de los subgrupos de unidades de información.

- (iii) Construir un modelo para cada variable a estimar y para cada subconjunto de la población que muestre un comportamiento diferenciado respecto de dicha variable.

Los modelos pueden ser simples o de gran complejidad dependiendo de la naturaleza de las variables a estimar. Los programas estadísticos más avanzados permiten, incluso, manejar modelos para hacer imputaciones múltiples.

Hay variables como la matrícula total que por ser más estables en el tiempo (siempre y cuando se cuente con evidencia para afirmar esto) son más fáciles de estimar, por ejemplo, mediante una simple imputación de la tasa de variación anual observada por unidades de información que pertenecen a un subconjunto lo suficientemente homogéneo respecto del comportamiento de tales variables. Así, si las variaciones interanuales de la matrícula son analizadas y se descubre que existen diferencias entre niveles, áreas de residencia y gestión del programa educativo y que estos subgrupos son homogéneos a su interior respecto de dicha variación, entonces se puede calcular las tasas para cada uno de esos subgrupos y aplicarlas a las unidades de los mismos que tengan la información sobre matrícula faltante.

Pero existen variables cuya estimación requiere modelos más complejos que pueden combinar tendencias históricas o un conjunto complejo de relaciones con otras variables conocidas siguiendo la lógica descrita cuando se habló del modelamiento en la primera sección (ver la página 89 y ss.).

Para desarrollar estos análisis existen diferentes herramientas estadísticas que incluyen, entre otros, los siguientes procedimientos: (a) cálculo de los comportamientos típicos y la dispersión en cada subgrupo (como promedios y desviaciones estándar de, en el ejemplo, las tasas de crecimiento interanual de la matrícula); (b) análisis de conglomerados (*clusters*); (c) análisis tipológicos basados en diversas formas de análisis de variables latentes. En cualquier caso, el tema clave es determinar subgrupos de unidades que sean diferentes entre sí, pero homogéneos a su interior.

- (iv) Estimar los valores e imputarlos a cada unidad de información. El procedimiento descrito implica estimar a nivel de cada unidad y no de agregados. A efectos

de evitar mayores distorsiones originadas en el uso de diferentes formas de estimación para cada variable, es siempre mejor imputar datos a nivel de cada unidad, incluso en los casos en los que la información a este nivel se considere demasiado frágil como para publicarla como estimación individual y solo se use para estimar agregados. Esta es una de las ventajas que resulta de la capacidad actual de los sistemas de registro y procesamiento computarizados de datos (que deriva en el uso de los llamados microdatos).

3.2.2 PROYECCIONES

Una proyección es una estimación de un valor futuro de una variable o conjunto de variables dadas. Es decir, una mirada anticipatoria o prospectiva respecto de una realidad que aún no está plenamente manifestada. Proyectar permite anticipar y no solo reaccionar.

Dado lo anterior, las proyecciones constituyen herramientas de singular importancia para anticipar necesidades o para explorar tendencias que se desea reforzar o modificar.

Las proyecciones se basan en diversas aproximaciones o modelos, entre ellos: (i) tendencias verificables en el tiempo; y (ii) simulaciones.

El primer grupo de proyecciones es relativamente más sencillo y «seguro», ya que solamente consiste en asumir que un comportamiento observado se mantendrá bajo los mismos parámetros en el futuro. El problema mayor de este enfoque es que justamente la política se suele formular de manera que se busca afectar la tendencia para que el progreso hacia determinados objetivos se verifique, por ejemplo, a un ritmo más elevado que el registrado en el pasado reciente. Es en este terreno donde el segundo grupo de proyecciones, las simulaciones, juega un papel clave, ya que involucra asumir determinados impactos futuros de medidas de política.

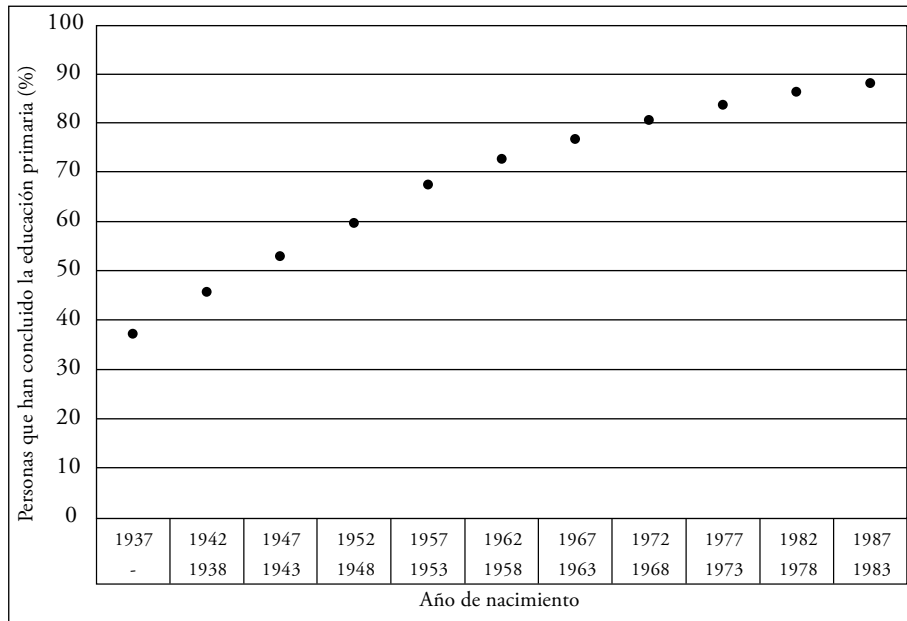
3.2.2.1 ¿Cómo proyectar una tendencia histórica?

- En primer lugar, es necesario contar con información que muestre dicha tendencia; es decir, se necesita una serie temporal de datos para la variable a proyectar. Esta serie debe ser consistente y hay que someterla a un riguroso análisis de modo que se pueda tener seguridad de que brinda una adecuada representación del comportamiento de la variable en el tiempo. Lamentablemente, dados algunos ejemplos vistos en la experiencia, resulta necesario subrayar que una serie de datos debe estar compuesta por un número significativo de observaciones, porque proyectar en base a, por ejemplo, solo dos puntos de datos es, para decirlo diplomáticamente, poco robusto.

A continuación, en el gráfico 23, se presenta una serie de datos sobre los niveles de conclusión de la educación primaria en dieciocho países de América Latina que fue generada con información de encuestas de hogares del año 2002. Esta serie de

datos será usada través de las siguientes páginas para ejemplificar cada uno de los pasos que se presentarán.

Gráfico 23. Personas que han concluido la educación primaria en dieciocho países de América Latina, nacidas hasta 1987 (en porcentajes)



Fuente: Unesco Santiago (2004: 51); elaboración propia.

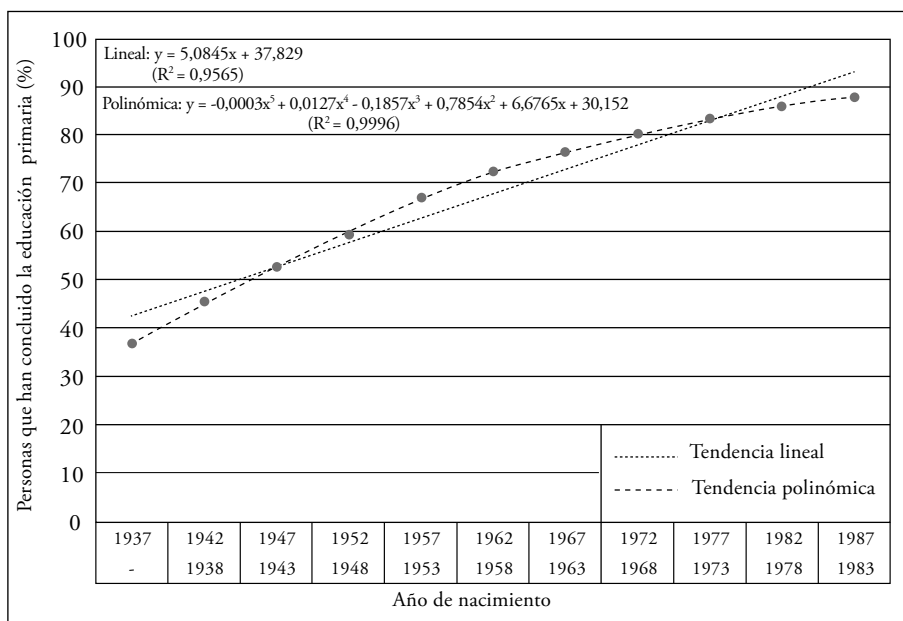
- En segundo lugar, es necesario definir un modelo estadístico a utilizar para la proyección. Así, debe discutirse si la variable a proyectar es una para la cual se puede esperar: (i) un comportamiento lineal; (ii) una curva de crecimiento con un límite definido (por ejemplo, la conclusión de un determinado nivel educativo no puede ser mayor a 100% para un grupo poblacional dado); (iii) un comportamiento análogo al observado con relación a otros fenómenos; etc. El *software* de hojas de cálculo suele incluir varios de estos procedimientos como opciones que son calculadas de modo más o menos automático para mostrar una tendencia y proyectar valores. El problema principal que hay que tomar en cuenta no es qué es lo que se puede hacer con el *software* disponible (es decir, hojas de cálculo y *software* estadístico), sino precisamente qué es lo que no se debe hacer con él: el *software* no piensa (por lo menos hasta ahora) y, por lo tanto, la decisión sobre el modelo a usar pasa por una comprensión sustantiva del fenómeno a proyectar.

El ejemplo con el que venimos trabajando es uno en el que se verifica una tendencia sostenida al crecimiento del valor del indicador. Sin embargo, no resulta razonable asumir un comportamiento lineal, ya que para este tipo de variables el nivel de

esfuerzo requerido para un incremento porcentual dado es usualmente mayor en los tramos más altos posibles (es decir, es más fácil pasar de, digamos, 20% a 30% en los niveles de conclusión, que de 95% a 97%), ya que los «últimos» tramos de la población suelen ser los más difíciles de atender.

- Una vez definido el modelo a utilizar para la proyección –entre las tres opciones de comportamiento señaladas–, se debe proceder al cálculo de la línea de tendencia buscando que la misma tenga un importante grado de ajuste a los valores observados. El ajuste se refiere al nivel en el que la línea de tendencia representa las observaciones y se estima mediante el cálculo de las distancias entre los valores observados y la línea que los representa. Usualmente, un coeficiente de correlación es calculado para medir dicho ajuste²²¹. Una vez que se cuente con la función que define la línea de tendencia, se procede a calcular los valores que corresponden a los períodos para los que se hace la proyección (ver el gráfico 24).

Gráfico 24. Personas que han concluido la educación primaria en dieciocho países de América Latina, nacidas hasta 1987 (líneas de tendencia; en porcentajes)



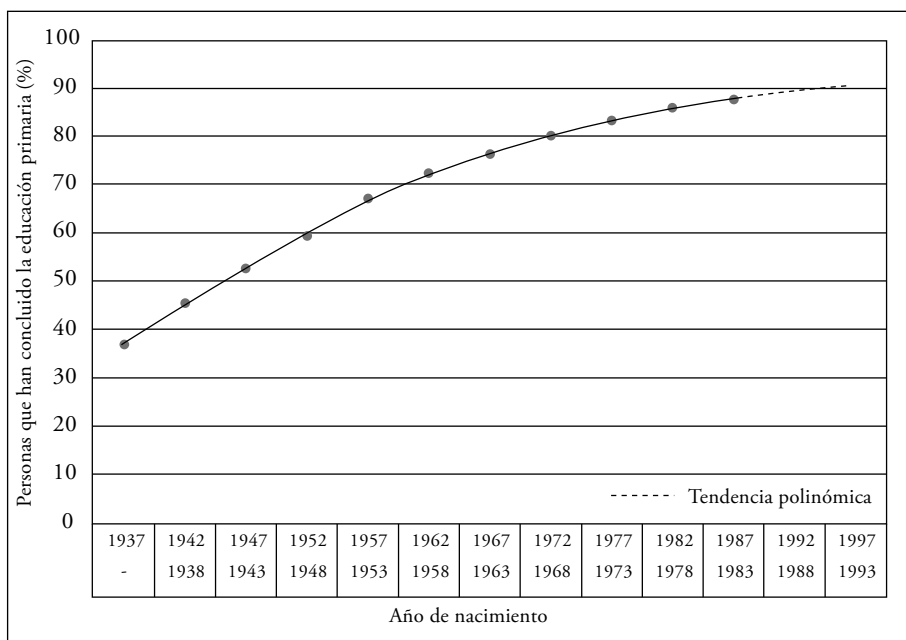
Fuente: Unesco Santiago (2004: 51); elaboración propia.

²²¹ Como se verá en el anexo 1, estos coeficientes de asociación entre variables arrojan valores entre 0 y 1, donde 0 indica la no asociación (o el no ajuste en este caso) y 1 una asociación perfecta (un ajuste total entre la línea de tendencia y los puntos observados en este caso). Dependiendo de la naturaleza de las variables, la asociación puede tener una dirección; así, esta puede ser positiva o directa (cuando ambas variables varían en el mismo sentido: ambas aumentan o ambas disminuyen), o bien negativa o inversa (cuando los incrementos en una se ven acompañados por disminuciones en la otra); esto se representa por el signo del coeficiente.

El gráfico 24 muestra, a manera de ilustración, dos posibles líneas de tendencia: (i) una tendencia lineal y (ii) una tendencia polinómica. Como queda claro, ambas tienen un buen ajuste (véase los valores de R^2 próximos a 1); sin embargo, la función polinómica tiene un mucho mejor nivel de ajuste, lo que se observa en la línea de tendencia calculada que toca prácticamente todos los puntos observados. Asimismo, debe notarse que la función lineal muy rápidamente va a tender a superar fácilmente el valor de 100%, lo que no es consistente con la realidad en general porque, como ya se mencionó, la estimación lineal asume que los ritmos de progreso (pendiente de la línea) se mantienen constantes.

Así, usando la función calculada que más se ajustaba a la tendencia observada, ha sido posible proyectar un comportamiento esperado que en el gráfico 25 es representado por el segmento punteado de la línea, el mismo que indica cuáles son los valores esperados para la conclusión de la primaria que serían observables hasta el año 2017 (cuando los nacidos en 2002 tengan quince años, es decir, una edad en la que se espera que ya hayan culminado la educación primaria).

Gráfico 25. Personas que han concluido la educación primaria en dieciocho países de América Latina, nacidas hasta 1987 (en porcentajes) y valores proyectados hasta las nacidas en 1997 (función polinómica)

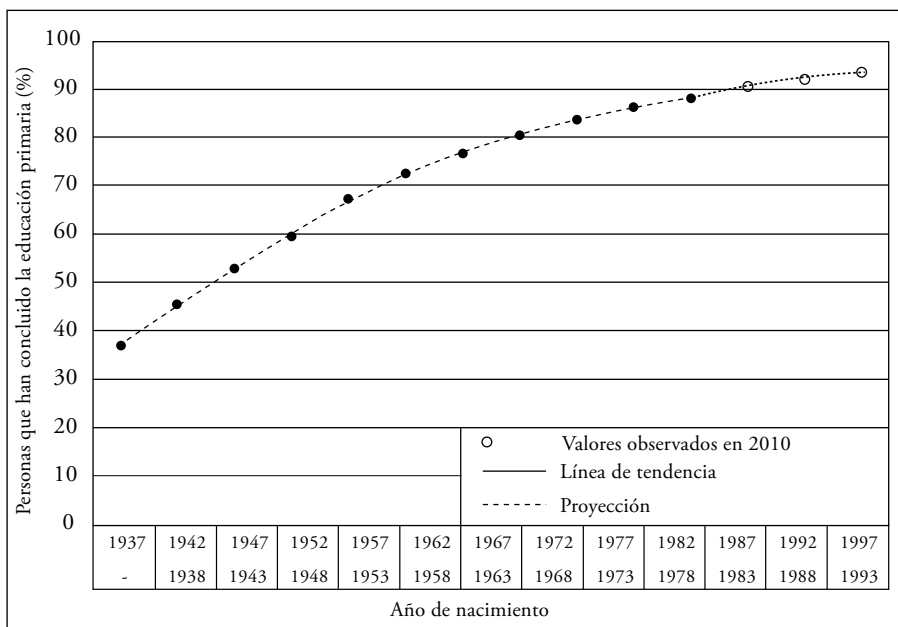


Fuente: Unesco Santiago (2004: 51); elaboración propia.

Esta proyección (hecha con información de encuestas de hogares de alrededor de 2002) es próxima a los valores observados con información computada con encuestas de hogares de alrededor de 2010, que muestran que la conclusión de la primaria para los nacidos entre 1990 y 1995 es de 90,2% (Unesco Santiago 2013: 66). Es decir, es una proyección validada por su consistencia con las mediciones independientes realizadas en los años posteriores a la generación de los datos usados para la proyección.

- Como ya se anotó, también es posible usar otro procedimiento, basado este en la observación empírica de algunas pautas que se consideren adecuadas para el caso, es decir, usando la opción (iii) mencionada anteriormente (en la página 265). En efecto, en vez de calcular una línea de tendencia, se pueden calcular, por ejemplo, las tasas de crecimiento observadas en los distintos países que se encuentran detrás de esta serie regional agregada para los distintos tramos de conclusión.

Gráfico 26. Personas que han concluido la educación primaria en dieciocho países de América Latina, nacidas hasta 1997 (en patrones observados en la región y tendencia polinómica)



Fuente: Unesco Santiago (2004: 51; 2013: 66); elaboración propia.

Como se detalla en el documento del que el ejemplo ha sido extraído (Unesco Santiago 2004: 47-48), este es el tipo de procedimiento que se siguió en dicha

publicación a efectos de considerar la dificultad creciente que se experimenta a medida que el indicador se aproxima al 100%, para lo cual se tomó como patrón de referencia los valores observados en los propios países de la región para los progresos observados más allá del umbral del 80%. Así, como muestra el gráfico 26, se obtuvo un resultado levemente más optimista que el mostrado con la función polinómica (aunque todavía próximo a los ya citados datos observados empíricamente con información de 2010).

- Finalmente, sea cual fuere el procedimiento utilizado, se debe proceder a su documentación, de modo que cualquiera pueda replicarlo y discutir sus supuestos.

Otro ejemplo de proyección de tendencias históricas nos permitirá mostrar algunas consideraciones adicionales que hay que tener en cuenta.

El gráfico 27 ha sido tomado del informe de avances sobre las metas del milenio que elaboró en 2004 el Gabinete Social del gobierno de la República Bolivariana de Venezuela con la colaboración de las agencias de las Naciones Unidas representadas en dicho país (República Bolivariana de Venezuela-Gabinete Social 2004).

Gráfico 27. Tasa neta de escolaridad en educación básica registrada y proyectada, Venezuela, 1990-2015



Fuente: República Bolivariana de Venezuela-Gabinete Social (2004: 44); los valores de la curva se toman literalmente de la fuente, en la cual se indica que el PNUD es el origen de los datos sin mayor especificación.

Nota

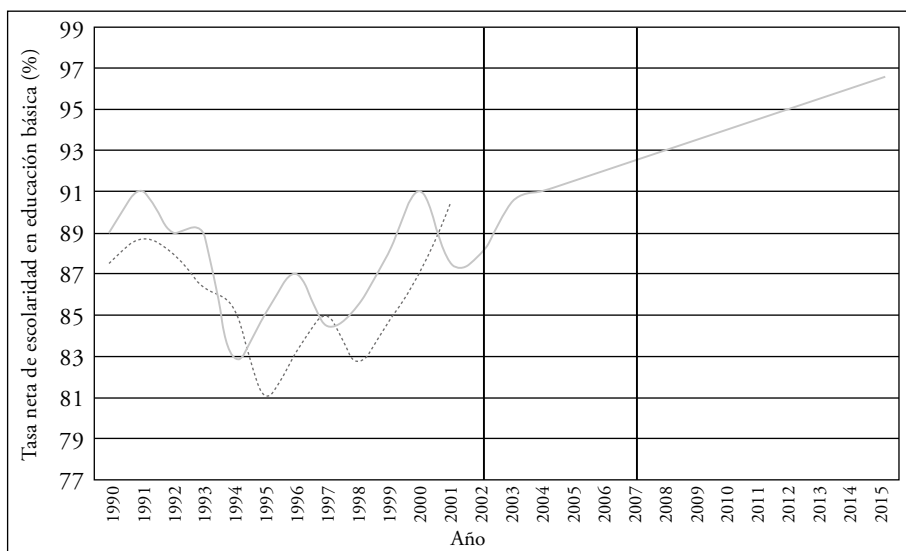
Se ha agregado al gráfico original los enunciados del lado izquierdo y de abajo.

El gráfico 27 permite algunos comentarios que son relevantes para nuestro interés²²²:

- En primer lugar, se mide la meta de la conclusión universal usando la tasa neta de matrícula, lo que, como ha sido discutido en la segunda sección, no es una forma suficiente para abordar dicha meta.
- En segundo lugar, se hace referencia a la educación básica y no a la educación primaria, que no son lo mismo. En cualquier caso, esto no representa un problema mayor (a menos que se quiera comparar con la situación de otros países), ya que cada país tiene el derecho de plantearse sus propias metas.
- En tercer lugar, en el gráfico se afirma: «La Tasa (sic) alcanza 100% antes del año 2015», lo que en realidad se contradice con el mismo gráfico, que muestra una proyección según la cual en 2015 se alcanzará un valor de aproximadamente 96%. De hecho, el gráfico ni siquiera contiene el valor 100 en su escala.
- En cuarto lugar, el texto señala: «Puede afirmarse que muy probablemente, si se logra mantener el ritmo de crecimiento de la matrícula observado en los últimos cuatro años, la meta podría lograrse hacia el año 2007» (República Bolivariana de Venezuela-Gabinete Social 2004: 43; esta es una cita de un texto identificado como PNUD [2004], cuya referencia no está disponible). Ello tampoco es consistente con el gráfico planteado, que muestra, para 2007, una tasa neta de matrícula estimada de aproximadamente 92%.
- En quinto lugar, el anexo de datos de este informe contiene (República Bolivariana de Venezuela-Gabinete Social 2004: 92) las tasas netas de matrícula registradas, las mismas cuyo comportamiento no es el presentado en el gráfico precedente; por esta razón aquí se modifica dicho gráfico a efecto de incluir esos datos, como se muestra en el gráfico 28.

²²² Nótese que todos los comentarios que aquí se plantean se refieren exclusivamente a los aspectos metodológicos de la proyección y a la presentación de la misma. En ningún caso constituyen opiniones sobre la situación educativa en dicho país.

Gráfico 28. Tasa neta de escolaridad en educación básica, Venezuela: registrada y proyectada, 1990-2015; periodo 1990-1991 a 2001-2002 (según el anexo de la fuente)

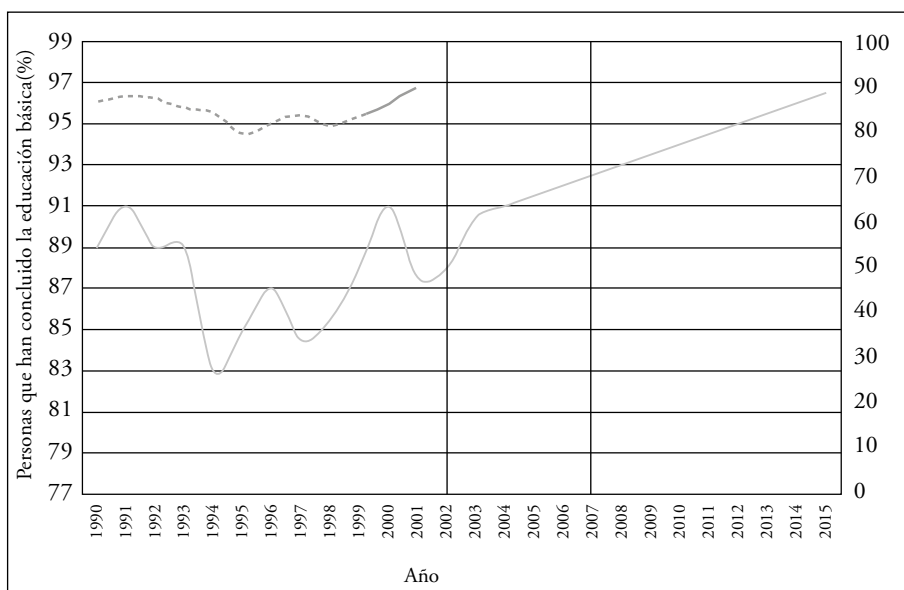


Fuente: República Bolivariana de Venezuela-Gabinete Social (2004: 44); los valores de la curva se toman literalmente de la fuente, en la cual se indica que el PNUD es el origen de los datos sin mayor especificación; los datos de la línea punteada se toman del anexo de la fuente (2004: 92).

- En sexto lugar, el recorte de la escala produce un efecto de exageración de la tendencia, lo cual si bien es un procedimiento aceptable en algunos casos (por ejemplo, para facilitar el reconocimiento visual de las diferencias entre dos series de datos en un mismo gráfico), aquí no contribuye a dicho fin²²³. Veamos lo que sería representar los mismos valores observados en una gráfica con la escala correcta (para ello, se usa el eje secundario, de modo que se pueda representar las dos curvas equivalentes en un mismo gráfico, el 29).

²²³ El tema del cambio de la escala para forzar tendenciosamente una impresión es uno de los ejemplos más conocidos que se describen en la excelente introducción a la estadística *How to Lie with Statistics* que Darrell Huff escribió en 1954 (1991).

Gráfico 29. Tasa neta de escolaridad en educación básica, Venezuela, 1990-2015



Fuente: República Bolivariana de Venezuela-Gabinete Social (2004: 44); los valores de la curva se toman literalmente de la fuente, donde se indica al PNUD como origen de los datos, sin mayor especificación; los datos de la línea punteada se toman del anexo de la fuente (2004: 92).

Nota

Ambas líneas representan los mismos valores, sin embargo, la línea punteada se lee en el eje secundario que va de 0 a 100.

La curva tiene los mismos valores que la línea punteada, solo que expresados en una escala no recortada de modo arbitrario e inadvertido (la escala completa de 0 a 100 se encuentra en el eje secundario, al lado derecho). El solo cambio de la escala crea una impresión distinta. ¿Existe la clara tendencia al aumento de la tasa que sugiere el gráfico original? La respuesta es clarísima: no.

- En séptimo lugar, el procedimiento de proyección no se encuentra documentado en el texto. Allí se sugiere que se ha procedido a una proyección lineal del crecimiento observado en los cuatro últimos años con datos. ¿Por qué cuatro y no toda la serie?, ¿solo porque el periodo coincide con una determinada gestión gubernamental?
- En octavo lugar, la línea de tendencia que se incluye en el gráfico original ha sido calculada **después** de proyectar los valores, lo que es metodológicamente insostenible. Las líneas de tendencia se computan con los datos observados para poder proyectar la tendencia, no con datos proyectados para graficar una tendencia pasada inexistente.

- En noveno lugar, una proyección basada en toda la serie, y no solo en los últimos cuatro años, no sería tan optimista. De hecho, proyecciones lineales, logarítmicas o exponenciales mostrarían una tendencia a la reducción del valor del indicador (sea con los valores del gráfico original o con los del anexo de datos).
- En décimo lugar, si uno se fija exclusivamente en los datos registrados, notará un comportamiento errático del indicador. ¿Es ese un comportamiento real? ¿No será el resultado de variaciones en las tasas de respuesta de los eventos que generaron los datos? La serie original no parece ser lo suficientemente sólida como para justificar una proyección.
- Por último, el gráfico original incluye dos líneas verticales que demarcan el período 2002-2007, al que se titula «Proyección», sin que medie mayor explicación. ¿El período 2008-2015 no es una proyección? ¿Cuál es el propósito de delimitar el período 2002-2007? ¿Tiene algo que ver con el enunciado –ni explicado ni documentado– de que la meta se logrará al 2007?

Ahora el lector entenderá cabalmente por qué se presentó este ejemplo: es muy difícil encontrar un gráfico que con una sola variable en una serie de tiempo presente de modo simultáneo más de una decena de desaciertos técnicos. De hecho, este ejemplo tiene casi tantos errores como puntos de datos (trece).

Si esta discusión metodológica aún deja dudas, una forma de demostrar de modo incontrovertible el valor negativo de proyecciones como esta consiste en mirar la realidad. Pasados los años, es posible ver en qué medida la proyección resiste una verificación empírica (como hicimos en el caso de las proyecciones de la Unesco Santiago 2004, las mismas que, efectivamente, se verificaron al menos a nivel regional).

A efectos de evitar cualquier suspicacia, en el gráfico 30 contrastamos la proyección del informe venezolano de 2004 con los datos incluidos en el más reciente informe oficial sobre los objetivos del milenio (República Bolivariana de Venezuela 2013: 47).

Si se compara la proyección elaborada sin ningún rigor metodológico que fuera presentada en el gráfico 27 con la información empírica reportada en 2013 (gráfico 30), se notará que aquella no se corresponde en lo más mínimo con lo que finalmente aconteció en el país en los años siguientes. ¿Cuál es entonces el valor de una proyección como tal?

Gráfico 30. Tasa neta de escolaridad en educación primaria, Venezuela, por periodos escolares, 1990-1991 a 2011-2012



Fuente: República Bolivariana de Venezuela (2013: 47); elaboración propia.

Nota

Ambas líneas representan la misma serie de datos. La curva es la misma que se presenta en el informe citado (República Bolivariana de Venezuela 2013: 47) y se lee en el eje principal (izquierda), que se encuentra recortado. La línea punteada se lee en el eje secundario (derecha), que presenta la escala completa. La fuente de datos citada en el informe es «Memoria y cuenta 2011-2012» del Ministerio del Poder Popular para Educación.

Por otra parte, la información del informe más reciente es destacable (a pesar del desacierto en cortar la escala), pues es presentada sin ser objeto de tortura (al decir de Coase) para que diga algo que no puede decir. De hecho, una de las cosas que llama la atención de tal informe es que el capítulo en el que este gráfico se presenta se inicia con la siguiente afirmación:

Venezuela logrará la universalización de la educación de la educación primaria antes del año 2015. Esto se refleja en el aumento de la tasa neta de escolaridad y en la proporción alta de niños que han alcanzado el sexto grado (República Bolivariana de Venezuela 2013: 47)²²⁴.

²²⁴ Las tasas de supervivencia al grado final de la primaria son presentadas en la página siguiente del informe y han pasado de 60% a 85% en el mismo período (1990-2012). Ahora bien, el texto afirma que estas tasas solo se refieren a personas que logran llegar a sexto grado sin atrasos (en seis años) y que, si se considera a todos los que alcanzan el sexto grado, esta cifra (que no es presentada en ninguna parte del informe) asciende a 97%. Al no haber ni una sola nota metodológica que explique qué es lo que se está presentando,

Esta afirmación no se sustenta en ninguna de las dos piezas de evidencia que se mencionan ya que, en primer lugar, las tasas netas de matrícula se encuentran estancadas en alrededor del 92% y, en segundo lugar, las tasas de supervivencia no superan el 85% (o el 97% según el texto). ¿De qué manera esos números respaldan la afirmación?

Este informe es un ejemplo en el que, claramente, la evidencia estadística solo cumple un lugar ritual y donde la afirmación que desea hacerse prescinde de la evidencia de modo absoluto (solo se ignora a los números que no dicen lo que se quiere). La ganancia metodológica (no repetir las «proyecciones» anteriores) observada entre este informe y el precedente debe ser encomiada, pero el vínculo entre la evidencia y el discurso político sigue siendo algo ajeno al rigor y al respeto por el derecho de las personas a discutir sin arbitrariedades sobre un asunto de interés público.

Más allá de la legítima preocupación que se puede tener por los derechos ciudadanos violentados de esta forma, lo que interesa destacar en esta sección es cómo estas acciones (proyecciones metodológicamente insostenibles, divorcio entre el discurso y la evidencia) solo llevan a debilitar la credibilidad de lo que se presenta. Por esta primera razón, debe evitarse a toda costa proyecciones de dicho tipo, pues cualquier lector consciente de los problemas aquí presentados (o de, al menos, una parte de ellos) podría preguntarse si los procedimientos usados en el resto de la publicación son similares. Esta sola pregunta siembra dudas inmediatas sobre la credibilidad de todo el esfuerzo desarrollado. En este caso (el informe de 2004), seguramente tal esfuerzo ha estado basado en la buena voluntad y en el aval que la masiva participación del sistema de las Naciones Unidas en el país brinda. Lamentablemente, ni el esfuerzo, ni la buena voluntad o el involucramiento de agencias de las Naciones Unidas presentes parecen haber sido suficientes para que se garantice niveles mínimos de rigor técnico en lo que al elemento singular aquí mencionado respecta.

3.2.2.2 ¿Cómo proyectar en base a un modelo de simulación?

Una forma diferente de proyectar es mediante modelos de simulaciones. Estos modelos básicamente consisten en el establecimiento de un conjunto de relaciones que definen el fenómeno a proyectar, un conjunto de supuestos y un conjunto de datos de partida.

Típicamente, los modelos de simulación de la evolución de matrícula pertenecen a este universo, así como todos aquellos que, partiendo de las simulaciones de matrícula, proceden a estimar necesidades de, por ejemplo, docentes, aulas, materiales educativos, etc.

Usualmente, los modelos de simulación y de proyecciones en educación operan sobre la base de un esquema similar al usado en demografía para las proyecciones de pobla-

no hay forma de evaluar estas afirmaciones. En todo caso, una tasa de supervivencia al sexto grado de incluso 100% no es sinónimo de conclusión universal, ya que dicha tasa no dice si el acceso es universal (pueden terminar todos los que acceden, pero si estos no son todos –y las tasas netas de matrícula sugieren que no lo son– no hay conclusión universal).

ción. Es decir, se construye una «tabla de vida» (en el caso de la educación, un modelo de cohortes como el descrito en la segunda sección [acápite 2.4.2] o indicadores de eficiencia como los que se ha presentado en este texto) que define el flujo temporal de la población escolar (asumiendo diversas hipótesis respecto de las variables de flujo) al que se suman supuestos sobre el ingreso al primer grado y la información sobre la población en edad de ingresar para distintos momentos en el tiempo.

A partir de lo anterior, se proyecta un flujo dado de matrícula. La proyección de la matrícula es un primer resultado de estos modelos. Dada esta proyección, es posible proceder a proyectar otros elementos asociados a la matrícula: (i) necesidades de docentes (asumiendo determinados valores para la ratio alumnos por docente); (ii) volúmenes globales de inversión (asumiendo determinados valores para la inversión por alumno); (iii) volúmenes específicos de inversión en infraestructura (asumiendo determinadas ratios de depreciación y necesidades de mantenimiento y de nueva construcción, las mismas que se vinculan a la matrícula de modo lineal o usando unidades discontinuas)²²⁵.

Estas ideas básicas pueden luego complejizarse si se toma en cuenta que las necesidades no son iguales en diversas zonas del país. Sin embargo, en estos casos los modelos de proyecciones basados en modelos de cohortes tradicionales se vuelven vulnerables debido a sus limitaciones para manejar transacciones entre diversas unidades de información²²⁶.

Estos modelos de simulación han cobrado tal difusión que en 2008 diversas organizaciones crearon una red y un sitio web específico al respecto: la Red Interagencial sobre Modelos de Simulación en Educación²²⁷.

²²⁵ Modelos simples asumen una relación lineal simple entre matrícula y, por ejemplo, necesidades de aulas (por ejemplo, un aula cada 30 alumnos, y para números que no son múltiplos de 30 simplemente se calcula la fracción y se redondea al entero más próximo). Modelos más complejos consideran rangos, de modo que la construcción de un aula adicional es identificada a partir de cuando se alcanza un número dado de estudiantes que equivale al límite superior de un rango y no antes. Así, no se define una ratio de 30 alumnos por aula, sino un aula por cada 20 a 40 estudiantes, de modo que cuando se alcanza el valor 41 se necesitarán dos aulas (una con veinte y otra con veintiún estudiantes). Lo mismo es de aplicación a las proyecciones de volúmenes de docentes necesarios u otros elementos que se refieren a grupos de estudiantes y no a cada uno de modo individual (o, en realidad, a cualquier variable discreta).

²²⁶ Como se mostró en la segunda sección, los modelos de cohortes asumen unidades de observación cerradas (no controlan migraciones o transferencias de matrícula entre diversas unidades) y con puntos rígidos de entrada y salida del sistema (solo se ingresa en el primer grado y cada salida es siempre definitiva). Así, cuando estos modelos se aplican a subunidades en las que se observa migraciones (por ejemplo, si las escuelas rurales tienden a ser incompletas –no ofrecen todos los grados–, la prosecución de estudios implica migrar), se empieza a obtener resultados inconsistentes, ya que la «deserción» en unas zonas (desde las que los estudiantes «emigran») se transforma en una no controlada «deserción negativa» en otras zonas (hacia donde los estudiantes «inmigran»).

²²⁷ La Inter-Agency Network on Education Simulations incluye a las siguientes instituciones: el secretariado de la iniciativa Fast-Track de la EPT (hoy Global Partnership for Education), el Banco Mundial, el PNUD, el Unicef, la Academy for Educational Development (AED, hoy FHI360), la Unesco; el Instituto Internacional de Planificación Educativa de la Unesco (Unesco-IIEP) y el Pôle de Dakar de Analyse Sectorielle en Education (véase: <<http://inesm.education.unesco.org/>>).

3.3. METADATOS E INTERPRETACIÓN

Tan importante como publicar la información es que esta se vea acompañada de la documentación suficiente para que sea clara y no induzca a errores. Así, junto con los atributos usuales de la información (período de referencia, definición de la variable usada, etc.) es de capital importancia que los usuarios puedan consultar lo que se ha venido en llamar **metadatos**, es decir, el conjunto de documentación que explica con detalle los procedimientos seguidos, los supuestos utilizados y los atributos de las fuentes de datos.

Un dato sin metadatos es potencialmente muy riesgoso, ya que se puede interpretar de una manera completamente ajena a lo que el dato dice. Este riesgo se amplifica cuando su definición (e incluso su nombre) carece de precisión, de modo que termina sugiriendo algo que no corresponde.

Del mismo modo, tan riesgoso como la publicación de datos sin metadatos es la práctica de utilizarlos sin prestar atención a dichos metadatos o sin siquiera requerirlos si no están inmediatamente disponibles. Una práctica completamente ajena a la rigurosidad intelectual consiste en tomar la información acríticamente y renunciar a la tarea obligada en un análisis riguroso en el cual se requiere «ensuciarse las manos» con los detalles que están detrás de la información que se usa. Publicar los metadatos minimiza los riesgos de mala utilización de datos, pero no los elimina pues permite, sin garantizarlo, que los usuarios tengan un comportamiento riguroso.

La publicación de información, como norma, debería incluir al menos los siguientes atributos:

- **Definiciones.** Es decir, qué es lo que se está midiendo efectivamente, incluyendo el procedimiento de cálculo y el significado de la información. Eventualmente, se debe advertir sobre significados erróneos y/o errores típicos.
- **Fuentes.** Refiere al origen de los datos con los cuales se ha elaborado la información: si es censal o muestral, quién fue responsable de la misma; y referir a posibles publicaciones que brinden mayores detalles.
- **Unidad de observación.** De qué o de quién se está hablando con esta información. Es decir, la unidad geográfica, de gestión, etc. a la que refiere la información. Imagine el lector que se encuentra con una tasa de matrícula que no indica de qué país es, esto sería completamente absurdo e inútil, sin embargo, cuando no se presenta el año, o la cobertura real de datos no nos encontramos en una situación muy diferente de aquella.

- **Período de referencia.** A qué momento en el tiempo refiere la información. Dado que la realidad es dinámica, publicar un dato sin fecha es como publicar uno que no hace referencia a ninguna unidad de observación.
- **Procedimiento de cálculo.** Es de fundamental importancia que los datos que son el resultado de un cálculo cuenten con información que permita que cualquier usuario lo replique. Los procedimientos de cálculo (o de estimación, donde fuera el caso) deben ser transparentes, de modo que el usuario pueda juzgar la pertinencia y rigurosidad de los procesos seguidos.
- **Supuestos utilizados.** La información originada en modelamientos no debe incluir solo claras referencias a los procedimientos seguidos sino a los supuestos que se utilizaron en el modelo. En un modelo de proyección de matrícula, por ejemplo, se suele hacer supuestos sobre la evolución de las tasas que expresan el flujo de la misma (como la repetición, la deserción y la promoción) o sobre los niveles de ingreso al primer grado.
- **Parámetros de cálculo.** En el caso de algunas variables, como las vinculadas a los logros de aprendizaje, se calculan parámetros (nivel de dificultad, discriminación, adivinación) para cada uno de los ítems de evaluación. Estos parámetros se computan en base a modelos que deben ser documentados; ellos son usados para estimar los niveles de competencia de las personas y deben ser conocidos.
- **Instrumentos de generación de información.** Es importante que tales instrumentos sean de dominio público para facilitar un mejor uso de la información y la reflexión crítica sobre el proceso que la genera. La única excepción a esta regla son los casos en los que la diseminación de los instrumentos de generación de información pueda tener un efecto negativo, como es el caso de las pruebas de logro o competencias, si es que se pretende reutilizar dichos instrumentos. En cualquier caso, una fracción de estos siempre puede hacerse de dominio público, al igual que sus especificaciones o formatos modelo.
- **Procesos de consistencia y «limpieza» de datos.** Normalmente los datos publicados no son idénticos a los primariamente generados, ya que es preciso que estos pasen por un proceso de revisión que asegure su calidad. Es importante que estos procesos se encuentren documentados, de modo que dicho proceso de «limpieza» no termine siendo una **caja negra** de arbitrariedad.
- **Revisiones metodológicas y actualizaciones.** Una de las cosas más frustrantes en este terreno consiste en obtener datos sobre el mismo fenómeno y de una misma fuente en dos momentos diferentes en el tiempo y descubrir que hay discrepancias. Las publicaciones de datos deben estar propiamente fechadas y cualquier modificación posterior tiene que ser documentada, sea que se trate de una actualización de datos o de un cambio en los mismos por una modificación en la metodología de cálculo utilizada.

- **Vigencia de las series.** Los cambios metodológicos pueden afectar las series históricas y esto debe documentarse con detalle.

Estos elementos se corresponden que un principio básico de la información al que se hace referencia en este documento y que consiste en entender que la información sobre educación generada por las agencias públicas es un **bien público**. Este principio ha sido destacado recientemente por la Learning Metrics Task Force (2013a) y es de primera importancia para evaluar, en especial, áreas de información, eventos y procedimientos de generación de datos que están (en muchos casos de modo perfectamente legítimo) bajo un régimen de propiedad particular. Por ejemplo, la industria de la aplicación de encuestas o de pruebas (test) no es solo un área de conocimientos académicos, sino un espacio en el que operan industrias y negocios. Si bien muchos de estos negocios son legítimos, no es conveniente que la información de interés público, en especial la que atañe a los derechos fundamentales de las personas y que es generada con recursos públicos, sea manejada de una manera que, en aras de la protección de la propiedad industrial, resta transparencia.

3.4 CENSOS, MUESTRAS Y PERIODICIDAD

Una parte de los temas que requieren ser objeto de reflexión es la identificación de los elementos de información que deben ser producidos censalmente (o mediante muestreo), así como la de la periodicidad en la generación de la información.

Muchas veces se tiende a asumir que los censos son intrínsecamente «mejores» que las observaciones por muestreo. Esto no es así, precisamente por el volumen de casos que se observa según cada procedimiento. Tratar con un menor número de casos (como es el caso del muestreo) permite tener un mucho mayor control del proceso de generación de datos, de modo que la información muestral puede, muchas veces, ser más «limpia» que la censal. Por otra parte, el muestreo permite, justamente por ser menor el número de observaciones, la administración rigurosa de instrumentos más complejos que, por lo mismo, permiten abordar más temas que los usualmente posibles de observar censalmente. La masividad de los censos obliga a no poder profundizar demasiado en cada tema cubierto y a que el número de estos sea reducido.

Por lo mismo, es preciso prestar una atención detenida y reflexionar profundamente acerca de qué información es producida de mejor manera (o incluso únicamente) mediante censos y qué información debe ser producida muestralmente.

La respuesta a esta pregunta no es inmediata, ya que necesita considerar los propósitos para los que se requiere la información. Por ejemplo, si las pruebas de medición de logro académico de los estudiantes tienen como propósito monitorear el estado general de la educación y desarrollar análisis que brinden pistas para la formulación de políticas, es mejor contar con operativos muestrales. Si, por otra parte, se busca que estas pruebas sirvan como un elemento de rendición de cuentas a la colectividad, no en el plano nacional sino en el plano local, es decir, si se quiere que esta información sirva para que los padres de familia cuenten con elementos de información sobre las escuelas a las que van sus hijos, es necesario que los operativos sean censales.

Dicho lo anterior, queda reiterar que no es posible dar una receta única acerca de cuándo debe preferirse un censo a una observación muestral. Lo que sí cabe es llamar la atención sobre la necesidad de reflexionar sobre esto y utilizar como elemento decisivo una detenida consideración de los propósitos para los que se genera información y de las ventajas, limitaciones y posibles consecuencias no deseadas que cada tipo de evento de producción de información puede tener.

La generación de información base que posteriormente será usada para diferentes fines, incluyendo el diseño de operativos muestrales, requiere de información censal. No hay forma de conducir una observación por muestreo que permita inferencias sin

contar con un marco muestral, es decir, sin un dimensionamiento cabal del universo de unidades de observación a considerar.

Otro aspecto crucial a considerar tiene que ver con la periodicidad con la que se requiere generar cierta información. En este caso, es la naturaleza del fenómeno a observar la que dicta la pauta a considerar; cuando se trata de fenómenos que no presentan grandes variaciones en períodos breves de tiempo (por ejemplo, si se está pensando en un levantamiento censal acerca del estado de la infraestructura educativa), no cabe pensar en levantamientos anuales o en ciclos más breves.

Aquí cabe indicar que la rutina burocrática tiende a favorecer la repetición invariable de los mismos eventos de generación de información año a año con el argumento relativo a la necesidad de construir series para observar tendencias. Sin embargo, nada obliga a construir una serie con observaciones anuales, aunque iniciarla en períodos más dilatados, evidentemente, requiere un tiempo inicial mayor.

Lo que resulta aconsejable en este caso es no solo prestar detenida atención a la variabilidad esperable del fenómeno en diferentes períodos temporales, sino también a la posibilidad de combinar diferentes estrategias de generación de información, de modo que se optimice el resultado general.

Por ejemplo, es posible pensar en un censo escolar de ciclo anual que genere alguna información básica todos los años, pero que tenga secciones modulares que se apliquen, por ejemplo, cada cinco años. De esta forma, por ejemplo, en el primer año se puede generar información básica que se preguntará todos los años más alguna información de detalle sobre infraestructura; en el año dos, la misma información básica más información de detalle sobre docentes; en el año tres, la sección modular puede centrarse en temas de gestión y participación; en el año cuatro, sobre estrategias de aprendizaje; en el año cinco, sobre nutrición y salud de los estudiantes; y, a partir de ahí, repetir el ciclo²²⁸.

Del mismo modo, se puede combinar eventos censales con eventos muestrales en un ciclo temporal. Por ejemplo, si a efectos de observar problemas de retardo en el crecimiento asociados a desnutrición se organiza un censo de peso y talla de estudiantes, este se puede repetir en intervalos decenales y, para el período entre dos censos, es posible conducir observaciones muestrales cada año que permitan dar un seguimiento a la tendencia. Incluso, procedimientos de estimación para áreas menores permitirían usar toda esta información para estimar comportamientos a niveles menores que los normalmente posibles de acuerdo al diseño muestral (Rao 2003).

Finalmente, es posible contar con eventos censales que en realidad sean una combinación inmensa de eventos muestrales simultáneos. Por ejemplo, imaginemos que se decide realizar un operativo censal de medición de logros académicos en matemáticas y

²²⁸ Un esquema como este es seguido desde hace varios años por la UEE del Ministerio de Educación de El Salvador.

lengua, en este contexto sería perfectamente posible plantear la conducción simultánea de una investigación sobre aspectos actitudinales también desarrollados como aprendizajes. Por ejemplo, como ya se mencionó, es posible construir herramientas para medir actitudes vinculadas a los cuatro pilares de la educación, pero estas herramientas serían muy complejas ya que la medición de cada aspecto a considerar requiere una batería de muchos ítems. Por lo mismo, en vez de aplicar todas las baterías posibles a todos los estudiantes, es posible aplicar algunas a diferentes grupos (es decir tomar muestras diferentes y aplicar un conjunto de instrumentos diferente a cada muestra) o, mejor aún, aplicar a las diferentes muestras instrumentos diferentes pero vinculados mediante bloques de preguntas que son rotados entre los diferentes grupos. Esto permitiría una cobertura temática muy grande cuya información, si además se vincula a la parte de los logros académicos, tendría un inmenso valor.

3.5 DATOS AGREGADOS Y MICRODATOS

El tema anterior (sobre censos y muestras) se vincula también a las formas de agregación con las que se requiere la información. El desarrollo tecnológico ha posibilitado pasar de sistemas de información basados en la agregación de datos por niveles (los datos de cada sección son agregados a nivel de programa educativo, los de estos a escala local, las cuales se suman a nivel subnacional y, finalmente, los datos nacionales resultan de la agregación final de los anteriores) a sistemas que compilan la información a nivel desagregado, es decir, a microdatos (estudiantes, secciones o programas educativos).

Sin embargo, ¿en qué medida esta información desagregada –microdatos– es utilizada? ¿Cuál es el valor agregado de contar con microdatos que usualmente son más difíciles de compilar que las agregaciones que solían marcar la actividad estadística en el pasado?

En efecto, si bien los análisis que consideren a las secciones o los programas como unidad básica son hoy posibles, no se aprecia abundancia de los mismos. Por ejemplo, actualmente se puede indagar acerca de los niveles de heterogeneidad de los programas educativos respecto de diferentes variables, incluyendo cuestiones tan básicas como su tamaño, si existe un patrón en la variabilidad del tamaño de los programas educativos o la forma que tiene dicha variabilidad. Y se puede llegar a información relacionada muy específica, por ejemplo, ¿cuán similares o distintas son las escuelas rurales entre sí?, ¿y las urbanas?, ¿a qué se deben estos patrones diferenciados, si los hubiera?, ¿cómo esta diversidad debería ser tomada en cuenta en el momento de definir pautas para asignar recursos?

El detalle de los microdatos permite análisis más finos que, lamentablemente, no se ven muy frecuentemente. Por ejemplo, es posible pensar en tipologías de programas educativos que brinden pautas más detalladas para la asignación de recursos, las cuales hacen posible tomar acciones que lleven a superar la tendencia aún presente de asignación plana o igualitaria (es decir, no equitativa) de los recursos públicos en educación.

Analizar la homogeneidad o heterogeneidad de las unidades de observación respecto de diversas variables clave es de fundamental importancia para entender que los sistemas educativos no solo tienen algunas características generales, sino que las mismas son el resultado de una combinación compleja de un inmenso universo de diversidad de situaciones y características que la información agregada no puede (ni tiene como propósito) mostrar.

Esto tiene directa relación con los intentos de construir sistemas de información que, finalmente, se basen en el registro individual de cada estudiante. Sin embargo, estos intentos a veces parecen conducidos más por su factibilidad tecnológica que por

su capacidad de servicio. Si bien los sistemas basados en el registro individual han de permitir análisis mucho más complejos (por ejemplo, flujos y dinámica temporal sin las grandes limitaciones de los modelos de cohortes reconstruidas), ¿en qué medida la necesidad de hacer estos análisis está presente a la hora de emprender esfuerzos para elaborarlos?

3.6 CARTOGRAFÍA EDUCATIVA

En los últimos años ha habido también una tendencia importante a la introducción o el desarrollo de la cartografía educativa. Sin embargo, se ha observado también que los propósitos que están detrás de estos esfuerzos no son del todo claros. Una inversión importante de recursos como la demandada por la cartografía digital no puede servir solamente a la inquietud de un consultor o al interés por contar con mapas de colores muy lindos que quedan bien en la oficina de un funcionario del ministerio. Para pintar mapas se pueden usar lápices de colores y eso podría ser más que suficiente.

¿Cuál es el propósito de contar con una cartografía digital? ¿Cuáles son los requisitos para que esta sea útil y sostenible?

La cartografía digital tiene varios propósitos posibles, que incluyen:

- (i) Apoyar en forma gráfica la representación de determinados fenómenos. El hecho sancionado por la expresión que afirma que **una imagen vale más que mil palabras**, significa que, por ejemplo, inequidades por áreas de residencia pueden ser representadas cartográficamente llevando a un mayor impacto visual.
- (ii) Facilitar el desplazamiento territorial del personal de gestión de la educación, sea durante sus visitas de apoyo pedagógico o en los propios eventos de generación de información.
- (iii) Servir de herramienta de apoyo a la gestión local en la medida en que permite un seguimiento espacialmente referenciado de las variables educativas. Esto deriva en que los gestores cuenten con un instrumento de representación de la información con menor nivel de abstracción que los informes o cuadros estadísticos, información que, a su vez, se encuentra asociada dinámicamente a otras variables vinculadas a la localización de la población, los servicios y las vías de comunicación disponibles.
- (iv) Apoyar la planificación nacional, dado que permite análisis generales sobre las características espaciales de la realidad educativa. Más allá de los mapas impresos, la posibilidad de calcular variables vinculadas a la fricción (es decir, a las dificultades reales de los traslados en el espacio implicados en asistir a una escuela, por ejemplo) y aspectos comunes observados en unidades espaciales permite identificar diversidad de situaciones y, por lo mismo, definir estrategias para lidiar con ellas.

Para que estos usos sean posibles es preciso, sin embargo, tener distintos elementos, que van más allá del *software* de sistemas de información geográficos.

En primer lugar, es preciso contar con una cartografía digital con **suficiente nivel de detalle** para los diferentes análisis requeridos. Los dos primeros usos mencionados requieren un grado menor de detalle, pero soportar los dos últimos supone una base informativa de gran envergadura, que incluye:

- El perfil topográfico del territorio a una escala suficiente como para no obviar accidentes naturales significativos al analizar variables vinculadas al traslado de las personas en el espacio. Por ejemplo, las curvas de nivel cada cien metros en un mapa de escala 1:250.000 pueden perfectamente hacer que la mayor parte de las cañadas o pequeños valles sean invisibles, volviendo completamente inútiles los análisis que con esta información se realicen.
- Cuerpos de aguas: desde ríos, riachuelos y lagos hasta bofedales y nevados, etc.
- Vías de comunicación: tanto carreteras como trochas y senderos.
- Toponimia: de todos los pueblos, ríos, cerros, etc.
- Centros poblados de todo nivel.
- Locales escolares, asociación entre estos y los programas y entre los programas y la información de matrícula, etc.

En segundo lugar, **toda** esa información debe **estar armonizada**, ser consistente entre sí y ser actualizada de modo regular. Por ejemplo, los responsables de mantener la cartografía vial de un país pueden trabajar con estándares metodológicos diferentes de aquellos usados para la construcción de la cartografía topográfica. ¿Existe una institución nacional encargada de la normalización de las diferentes cartografías digitales?

En tercer lugar, la cartografía digital para ser útil requiere la conjugación de **varias capas de información** que son producidas por diversas instituciones. Quién es responsable de qué partes y cómo se coordina institucionalmente el esfuerzo son preguntas centrales que, si no son respondidas o consideradas en la planificación de los esfuerzos, pueden llevar a situaciones de inoperancia o a la no sostenibilidad de estos.

Una cartografía detallada y de calidad puede constituir una herramienta clave de lo que ha venido en llamarse microplanificación. Pero, ¿para qué microplanificar en un esquema institucional altamente centralizado?

3.7 EDUCACIÓN NO FORMAL

En ningún momento en el desarrollo de este texto se ha hecho referencia a los programas educativos no formales. Esta aparente omisión obedece a una razón fundamental: no existe ninguna motivo para restringir lo aquí dicho a los llamados programas formales de educación. Un sistema de información robusto, como se señaló en la primera sección cuando se discutía acerca de lo que definía la calidad de estos (página 34 y ss.), tiene que **ser suficiente** en cuanto a su alcance (página 38); es decir, cubrir lo que debe cubrir.

Sin embargo, existe una tendencia a imaginar la discusión anterior con relación a lo que constituye el universo de programas «regulares» o a los que, de cierta forma, se considera que presentan la trayectoria educativa «deseable», frente a otras opciones que, en cierto sentido, son consideradas prácticamente como un subestándar.

Más allá de los aspectos potencialmente discriminatorios que se encuentran debajo de las prenociones descritas en el párrafo precedente, la discusión sobre el alcance de los sistemas de información educativa –incluyendo a la llamada educación no formal– supone considerar al menos los siguientes dos elementos: (i) ¿qué es la educación no formal? y (ii) ¿cuál es el efecto posible de extender el alcance de los sistemas de información educativa de modo que cubran la educación no formal donde corresponda?

Un primer elemento a considerar es contar con una definición clara de lo que se quiere abordar. Desafortunadamente, los términos educación formal, educación no formal y educación informal no son usados de la misma manera en todos los países, ni tampoco, necesariamente, en su interior.

Por lo tanto, destinemos unos minutos a tratar de clarificar este tema.

3.7.1 LOS CONCEPTOS DE LA CINE 1997

La revisión 1997 de la CINE (Unesco 1999a) proporciona un conjunto de elementos que nos pueden servir como punto de partida para esta discusión: en primer lugar, el glosario de la CINE contiene definiciones expresas de educación formal y no formal:

Educación formal (o educación inicial o educación ordinaria escolar y universitaria)
Educación impartida en el sistema de escuelas, facultades, universidades y demás instituciones de educación formal que constituyen una «escalera» de educación de tiempo completo para niños y jóvenes, que suele comenzar entre los cinco y siete años y continuar hasta los 20 ó 25. En algunos países las secciones superiores de esta «escalera» están constituidas por programas organizados de trabajo a medio tiempo y participación a medio tiempo en el sistema ordinario escolar y universitario: estos

programas se denominan «sistema dual» o con términos equivalentes en los respectivos países (Unesco 1999a: 47).

Educación no formal

Toda actividad educativa organizada y sostenida que no corresponda exactamente a la definición anterior. Por lo tanto, la educación no formal se puede impartir dentro o fuera de un establecimiento docente y está destinada a personas de cualquier edad. Según los distintos contextos nacionales, puede abarcar programas de alfabetización de adultos, educación básica para niños no escolarizados, formación práctica, capacitación laboral y cultura general. Los programas de educación no formal no siguen necesariamente el sistema de «escalera» y su duración es variable (Unesco 1999a: 47).

Estas primeras definiciones ameritan algunas reacciones que pueden ayudarnos a comprender mejor el tema:

- El párrafo 12 de la CINE 97 (Unesco 1999a: 10) señala expresamente que este glosario es «provisional», lo cual introduce un elemento importante de incertidumbre sobre las definiciones provistas. En este contexto, decir que algo es provisional puede ser entendido como que los autores de la CINE no estaban completamente convencidos de estas definiciones. La revisión 2011 de la CINE (a la que nos referiremos más adelante) subsana este y varios de los problemas que se destacan aquí con el propósito de elaborar una discusión que resulta relevante ya que, más allá de las mejoras introducidas en esta revisión, muchos de los problemas que acá se identifican siguen vigentes en la discusión educativa general.
- No se incluye en las definiciones a la educación «informal», cosa que ayudaría a acotar los límites de lo formal y no formal. De hecho, en la CINE 97 no hay una sola mención a la educación informal.
- Más allá de los dos elementos anteriores que son más de forma por su naturaleza, existe un tema sustantivo por abordar en estas definiciones: como resulta claramente insinuado por la frase entre paréntesis que acompaña el título «educación formal», esta se equipara a la educación «regular» u «ordinaria». Entonces, ¿qué es educación regular? Ello no se define en la CINE 97.
- Este no es un tema menor ya que, además, la educación no formal es «definida» como todas aquellas actividades educativas organizadas y sostenidas que no se **correspondan exactamente** a (sic) la definición anterior²²⁹. ¿Qué quiere decir «no

²²⁹ Nótese también que la traducción al español deja mucho que desear. Por ejemplo, la CINE habla sistemáticamente de **enseñanza** en vez de **educación**, y en casos como el del glosario habla de **centros docentes** por decir **centros educativos**. Este problema de interferencia del francés en las traducciones al español en los documentos de la Unesco ya se observó anteriormente a propósito de otro documento (ver la nota 79 (en la página 111)). Por cierto, los problemas que aquí se discuten no son de expresión sino de concepto y se presentan de igual manera en la versión original en inglés de la CINE 97.

exactamente»? La verdad es que ese no es un uso muy preciso del idioma que ayude en una definición.

Resumiendo, en primer lugar, la CINE 97 afirma que la educación formal es la educación regular (lo que sea que esto signifique) y que la no formal es todo lo demás en la medida en que se trate de actividades educativas organizadas y sostenidas. No se trata, claramente, de un tratamiento muy riguroso ni muy afortunado del tema.

Sin embargo, y en segundo lugar, a pesar de las debilidades se subraya un elemento clave: un programa educativo es una actividad **organizada y sostenida**. Sin embargo, la definición incluye un elemento adicional que define por qué los programas son organizados y esto se refiere a la **intencionalidad educativa** de la acción. De acuerdo a la CINE 97 –Unesco 1999a: 10–, organizado quiere decir **planificado y con objetivos explícitos e implícitos**. Por otra parte, **sostenida** (aunque en la versión en español se usa también **continuada** –Unesco 1999a: 10–, que no es lo mismo) alude a que la experiencia tiene duración y continuidad. Esto, sin embargo, no es definido con precisión, ya que, para empezar, toda experiencia tiene una duración (el tiempo es un atributo del mundo en el que vivimos y resulta un oxímoron imaginar una experiencia que no tenga duración alguna); además, este es un tema que se deriva para ser tratado en un «manual operacional» (que nunca fue producido) y parece referirse exclusivamente al establecimiento de duraciones mínimas de los programas; finalmente la idea de **continuidad** no es definida en lo absoluto.

La **intencionalidad** de la acción educativa que deriva en la constitución de un programa (es decir, en un cuerpo organizado de acciones con un propósito educativo) es lo que diferencia entonces el aprendizaje informal, accidental y/o aleatorio del aprendizaje desarrollado a través de un programa educativo (formal o no formal, eso lo veremos en un momento).

Así, ya hemos logrado un primer elemento para acotar el ámbito de observación y excluir el aprendizaje informal o aleatorio. Sin embargo, esto no es suficiente para diferenciar lo formal de lo no formal y, hasta este punto, solo contamos con la no definida idea de «educación regular» como criterio a utilizar.

Así, en tercer lugar, tenemos dos opciones: (i) asumir que la educación formal es efectivamente equivalente a la «educación regular» y proceder a definirla; o (ii) explorar un poco más este tema para tener una mejor comprensión de las cosas y, a partir de dicha comprensión, tomar una mejor decisión. Total, si la definición de la CINE 97 fue auto-declaradamente «provisional», ¿por qué convertirla en definitiva no siendo, en realidad, concluyente? Esto se torna imperativo si leemos con detenimiento el parágrafo 12:

La CINE abarca tanto la educación inicial de los primeros años, antes de que la persona entre en el mundo del trabajo, como la formación a lo largo de toda la vida.

En consecuencia, a efectos de la CINE, la educación incluye una gama de programas y tipos de educación que se designan en el contexto nacional, por ejemplo, educación ordinaria, educación de adultos, educación formal, educación no formal, educación inicial, formación permanente, educación a distancia, educación a lo largo de toda la vida, educación a tiempo parcial, sistemas duales, aprendizaje de oficios, educación técnica y profesional, capacitación y educación de niños con necesidades especiales. Al final del presente documento figura un glosario provisional (Unesco 1999a: 10).

Como resulta claro a partir de este texto, la CINE 97 no es consistente en su definición de educación formal (y esta debe ser una de las razones para autodeclarar la provisionalidad de las definiciones), ya que mientras en el glosario se usa la equivalencia con «regular», aquí aparecen como dos categorías diferentes (nótese que aquí «regular» se tradujo como «ordinaria»).

De hecho, este párrafo resulta interesante, ya que para ser consistentes las categorías de una variable (en este caso, tipo de programa) deben ser exhaustivas y excluyentes y este no es el caso acá. En realidad, lo que se tiene es una serie de etiquetas que se usan en diversos contextos y que no están organizadas de ninguna forma. Se podría pensar, por ejemplo, que la educación a tiempo parcial es diferente de la educación de adultos cuando estas categorías obedecen a dos criterios de clasificación diferentes.

Lo que se tiene aquí es una confusión de términos que no toma en cuenta uno de los primeros elementos abordados en este volumen: la necesidad de definir y diferenciar conceptos clave (ver las páginas 43 y ss.).

La CINE es una clasificación de **programas** educativos y aquí se confunde este tema con la existencia de diversas **modalidades** de prestación de los servicios educativos. Un programa educativo es definido de acuerdo a su contenido (un programa de educación primaria, por ejemplo) y puede ser brindado de acuerdo a distintas modalidades que se estructuran, a su vez, de acuerdo a diversos criterios: organización del tiempo (completo o parcial), población objetivo (niños, adultos, niños con determinadas características –en zonas rurales, con lengua nativa diferente de la oficial, etc.–), organización de las actividades (presencial o a distancia, etc.)

Entonces, en cuarto lugar, lo que encontramos en la CINE 97, y en realidad en el uso corriente en muchos lugares, es que la educación formal es hecha equivalente a cierto tipo de educación llamada allí «regular», que implica no solo una estructuración en niveles que se corresponden con una «escalera» de calificaciones, sino también con modalidades consideradas «regulares» o más bien, consideradas «óptimas» o «normales» **desde una determinada perspectiva cultural y social.**

Dicho con crudeza, lo que este uso trasunta son afirmaciones como las siguientes: la educación **no** «regular» es aquella que hay que ofrecer a los niños pobres que tienen que trabajar (por lo que hay que brindarles otro tipo de programa distinto del «regular»), a niños que tienen una lengua materna diferente de la oficial (por lo que hay que tener

para ellos otro tipo de programa, distinto del «regular»), a jóvenes y adultos que no concluyeron exitosamente su educación «regular» muy probablemente porque el sistema educativo no fue capaz de brindarles adecuadamente el servicio (por lo que hay que ofrecerles otro tipo de programa distinto del «regular»), etc.

¿Existe acaso una razón para que un programa de alfabetización de adultos **no** pueda ser formal? ¿Por qué la educación universitaria, que de hecho incluye adultos y puede ser a distancia o presencial, a tiempo completo o parcial, etc. es necesariamente formal y otras formas de educación con los mismos atributos **no** lo son?

Es decir, el tratamiento del tema (tanto en la CINE 97 como en el habla corriente) trasunta prejuicios culturales, étnicos y/o de clase social que entran en juego, especialmente ante la ausencia de claridad conceptual. Así, la educación «regular» es el óptimo que se ofrece a los sectores más favorecidos de un país y la «otra» es el subóptimo que se brinda a aquellos que viven algún tipo de marginación. Una vez más, las definiciones «estadísticas» no son accidentales y portan, intencionalmente o no, una determinada visión de las cosas que es sostenida por sus autores (por cierto, no podría ser de otra manera, ya que las definiciones son construidas por personas).

3.7.2 LOS CONCEPTOS DE EUROSTAT²³⁰

Así, nos guste o no, debemos abordar una discusión más detenida sobre el concepto de educación formal y educación no formal.

En este sentido, un excelente punto de partida es el manual *Classification of Learning Activities* (*Manual de clasificación de las actividades de aprendizaje* de Eurostat; European Commission 2006) desde su propio título, que expresa la intención subyacente: el foco se pone **en** las actividades de aprendizaje, es decir, **no** en la prestación del servicio de enseñanza.

El documento mencionado empieza afirmando entre las razones para haberlo desarrollado la falta de claridad conceptual con relación a lo informal y lo no formal (tal y como venimos exponiendo en este acápite).

El manual define al inicio qué es lo que ha de incluir como su ámbito, el cual es definido como **actividades de aprendizaje** que se pueden presentar a lo largo de la vida. Estas actividades de aprendizaje:

- Han de ser **intencionales**, ya que las personas se involucran en ellas a efectos de mejorar sus comportamientos, información, conocimientos, entendimiento, actitudes, valores o destrezas²³¹.

²³⁰ Oficina estadística de la Comisión Europea.

²³¹ Esta definición es consistente con el párrafo 9 de la CINE 97 (Unesco 1999a: 9).

- Tienen un **carácter sostenido en el tiempo**, lo que significa que no son incidentales o aleatorias, es decir, son organizadas independientemente de su duración (esto, finalmente, dota de contenido a la idea de «sostenida» del parágrafo 11 de la CINE 97 [Unesco 1999a]).
- Pueden ser **formales o no formales**, es decir, no importa la modalidad (así, la identificación entre formal y una determinada modalidad es explícita).
- Son financiadas, pero **no se considera quién las financia**.
- Lo son, **independientemente de la forma como se brinda el servicio**.
- Abarcan a **toda la población sin importar la edad**²³².

Así, Eurostat empieza justamente diferenciando el aprendizaje informal del estructurado u organizado e intencional, separando los temas de contenido de las modalidades o formas de prestación de los servicios e incluyendo a toda la población. Posteriormente, distingue entre actividades singulares de aprendizaje (European Commission 2006: 10-11) y aquellas estructuradas en un programa, es decir, diversas actividades singulares organizadas en un conjunto o secuencia (European Commission 2006: 11-12).

A partir de aquí, define que los aprendizajes pueden darse en cuatro ámbitos: formal, no formal, informal y aleatorio (European Commission 2006: 12 y ss.). Sin embargo, en esta parte, el manual cita las definiciones «provisionales» de la CINE 97 para educación formal y no formal y agrega las de educación informal (como intencional pero «menos» –¿cuánto menos?– organizada y estructurada) y educación aleatoria²³³.

Más adelante, Eurostat distingue que la «educación puede ser formal o no formal; mientras que el aprendizaje puede darse en cualquiera de estos ámbitos, siendo que la educación ha de ser institucionalizada» (European Commission 2006: 14).

Esto, de alguna manera, nos devuelve al inicio de la discusión, dado que ya habíamos definido un foco en el aprendizaje intencional y organizado (que es el objeto primario de la acción pública en educación, [ver la página 118 y s.]), pero no nos ayuda a subsanar los problemas identificados en las definiciones de la CINE 97 al interior de lo que es considerado educación.

Ahora bien, el mismo manual avanza un poco más y distingue la educación formal de la no formal (más allá y, en cierta medida, contradiciendo la limitada definición de la CINE 97) y lo hace a partir del **tipo de certificación** que una y otra producen tras la conclusión exitosa del programa. Esto, felizmente, nos regresa al tema de los contenidos del programa.

²³² Tales elementos son tomados por este manual (European Commission 2006: 8) citando a un grupo especial de trabajo de Eurostat (Education and Training Statistics, «Report of The Eurostat Task Force on Measuring Lifelong Learning [TF/MLL]» [Pilos 2001]).

²³³ Esta debilidad hace particularmente importante remitirse a la CINE 2011 (como se hará más adelante), que retoma el trabajo de Eurostat.

Eurostat vincula estas certificaciones a los marcos nacionales de calificaciones: un programa formal brinda calificaciones reconocidas en dicho marco, un programa no formal no lo hace. Ahora bien, estos marcos, si bien están extendidos en Europa, no son una práctica común en otras latitudes. Así, si bien la solución provista por Eurostat a efectos de diferenciar lo formal de lo no formal es completamente consistente (y muy sensata), es dependiente de un instrumento con el que no se cuenta en otras latitudes.

Por lo tanto, ¿cómo proceder ante la ausencia de un marco nacional de calificaciones?

Una solución pragmática consistiría en tomar como marco de referencia al esquema de calificación provisto por el sistema «regular» y, además, cualquier programa que brinde calificaciones equivalentes debe ser tenido (al igual que la educación regular) por «formal».

Esto, si bien no termina de resolver el problema de fondo, permite postular con bastante claridad que una fracción de los programas «no regulares» tiene el mismo valor educativo que los «regulares». Lo cual permite, por lo menos y esto es muy importante, contrarrestar la tendencia a considerar como necesariamente de menor calidad o, incluso como no deseables, a las opciones que usualmente sirven para atender a los segmentos menos favorecidos de la población, dada la necesaria flexibilidad que el sistema educativo debe tener (la flexibilidad ya fue destacada como un elemento clave para que la educación sea pertinente y contribuya a la equidad, dos criterios claves para hacerla de calidad [ver la página 120 y ss.]).

Así, por ejemplo, un programa no presencial, para adultos, que brinda calificaciones equivalentes a las de un programa «regular» de educación secundaria vocacional debería ser considerado equivalente a este y, por lo mismo, ambos programas son igualmente formales independientemente de la modalidad de prestación del servicio.

Lo mismo aplica a un programa de alfabetización de adultos que asegura aprendizajes equivalentes a la conclusión de toda o una parte de la educación primaria.

Nótese, sin embargo, que aquí existe un potencial problema: usualmente los programas educativos **certifican** los aprendizajes y, como ya se ha visto, esta certificación no es suficiente para afirmar que las personas han desarrollado las calificaciones requeridas. Sin embargo, este punto es de idéntica aplicación tanto para los programas «regulares» como para los «no regulares» y, por lo mismo, no es relevante para resolver el problema acá planteado.

Así, podemos concluir esta discusión afirmando que:

- (i) Todos los programas que ofrecen calificaciones que se corresponden con el marco nacional de calificaciones correspondiente (o lo que se tome como referencia para definir esto, por ejemplo, el *currículum* nacional de cada nivel educativo) han de ser considerados formales independientemente de su modalidad.
- (ii) Los sistemas de información educativa deberían abarcar, por lo menos, a todos estos programas.

- (iii) Es necesario promover una mayor claridad conceptual respecto de estos temas, ya que existe una no justificable y discriminatoria práctica de equiparación de lo no «regular» con lo no formal.

3.7.3 LA REVISIÓN DE CONCEPTOS DE LA CINE 2011

Como ya se anticipó, la revisión 2011 de la CINE (Unesco-UIS 2013a) aborda estos temas con mayor detalle y rigor que la revisión 1997. Al hacerlo, se nutre de la experiencia europea y plantea un enfoque renovado en el que cabe destacar algunos elementos centrales vinculados a esta discusión:

- (i) En primer lugar, se define a sí misma como un esfuerzo destinado a contar con un marco para la clasificación de **actividades educativas** determinadas en el marco de programas educativos (Unesco-UIS 2013a: § 2).
- (ii) Mantiene la idea principal de clasificar a los programas según su contenido y campo de estudio. A esto añade una clasificación análoga para los niveles de escolaridad logrados por las personas, lo cual está directamente vinculado al tema de la certificación de conocimientos, habilidades y destrezas en un marco de calificaciones.
- (iii) Afirma que los contenidos del programa han de traducirse (o se espera que se traduzcan) en conocimientos, habilidades o competencias que son **certificados** por el programa (mediante algún tipo de credencial, certificado, grado, diploma, etc.). La certificación como el elemento que permite el tránsito a lo largo del sistema educativo y a través de las diferentes modalidades de educación es una preocupación principal para esta revisión, de ahí el interés en clasificar también la escolaridad lograda por las personas y no solo los programas. Es evidente que la certificación formal presenta muchas debilidades como evidencia certera de contar con los correspondientes saberes, destrezas o competencias que los programas ofrecen, tanto por sus posibles debilidades (y de las instancias públicas que deben verificar que cada programa realmente brinde lo que ofrece), como por las posibles pérdidas de habilidades, destrezas, competencias y conocimientos a lo largo del tiempo. Sin embargo, es también evidente que la CINE, que es consciente de este problema (Unesco-UIS 2013a: § 24), no es el lugar en el que puede abordarse.
- (iv) El interés en las certificaciones permite a la revisión 2011 de la CINE estar abierta a diversos mecanismos de certificación existentes, por ejemplo, a la del aprendizaje basado en la experiencia. En la medida en que la autoridad educativa haya establecido mecanismos para dicha certificación, la CINE 2011 está lista

- para reconocerla en la clasificación de escolaridad lograda como equivalente a la conclusión exitosa de un programa formal que conlleve a la misma certificación.
- (v) Con estas consideraciones en mente, la frontera entra la educación formal y la no formal no puede ser definitiva. De hecho, la CINE 2011 (ver en especial el anexo V en Unesco-UIS 2013a: § 287-296) muestra dicha dificultad y revierte parte de la discusión a criterios adicionales, como: (a) lo formal tiene que ver, principalmente, con la educación «inicial» de tiempo completo, es decir, con aquella en la que las personas nos involucramos antes de nuestro ingreso al mercado de trabajo; (b) por lo mismo, lo no formal se vincula más a la necesidad de ofrecer oportunidades de aprendizaje a lo largo de la vida con mayor flexibilidad (menor duración, no necesariamente estructurada alrededor de una institución educativa, etc.); y (c) lo formal es lo que el país define y trata como tal.
 - (vi) Sin embargo, la CINE 2011 insiste correctamente en la necesidad de enfocarse en el contenido de la educación no formal y en los aprendizajes esperados (traducidos en una certificación) para ver si son o no equivalentes a lo ofrecido en programas «formales». Aquí es donde la distancia entre lo formal y lo no formal parece más difusa, ya que si tenemos dos programas que ofrecen contenidos y certificaciones equivalentes, ¿por qué llamar a uno «no formal»? ¿simplemente porque se ofrece de un modo que no se corresponde con lo «típico»?
 - (vii) Esto se vincula al hecho de que la CINE 2011 postula que la educación no formal «mayormente» ofrece calificaciones que no son reconocidas como equivalentes a las contenidas en el marco de calificaciones o a las ofrecidas en el sistema «formal». Desde este punto de vista, la diferencia podría resultar más clara. (Si un programa educativo no ofrece una certificación reconocida es no formal, pero, ¿qué pasa cuando sí la ofrece?).

A pesar de lo avanzado en este terreno (como en otros), la CINE 2011 sigue planteando la necesidad de «desarrollo adicional» en este campo y, vinculado a esto, que los esfuerzos internacionales de producción de información estadística comparable se enfoquen en el ámbito formal. Claro está, esta recomendación final (Unesco-UIS 2013a: § 296) no tiene por qué ser de igual aplicación en el ámbito nacional.

3.7.4 EFECTOS POTENCIALES DE LA AMPLIACIÓN DEL ALCANCE DE LOS SISTEMAS DE INFORMACIÓN

En atención a la segunda de las dos preguntas que planteamos al inicio de este acápite sobre los potenciales efectos de ampliar el alcance de los sistemas de información que hoy solo se enfocan en la educación «regular», es necesario decir que ampliar el alcance

de un sistema de información representa, siempre, incrementar el volumen de sus operaciones. Esto es de aplicación no solo al caso de intentar cubrir programas que no se están cubriendo en un momento dado, sino también a la introducción de nuevos operativos, a la ampliación del alcance de operativos existentes, etc.

Esta situación implica considerar que para ampliar el alcance es necesario hacer un ajuste en la dotación de recursos y capacidades profesionales del sistema de información, de modo que dicha ampliación sea manejable. Este ajuste puede implicar aumentar los recursos materiales disponibles, reorganizar su uso actual de modo de hacerlos más eficientes, hacer un mejor uso de ellos dadas las nuevas circunstancias, ampliar la capacidad profesional de los equipos (incluyendo *expertise* con la que no se contaba y/o mayor capacidades operativas), etc.

Lo anterior es de capital importancia no solo a efectos de asumir exitosamente las nuevas tareas, sino también para preservar los niveles de calidad con los que se vienen desarrollando las tareas precedentes. Dicho de otra forma: **incrementar las responsabilidades sin un ajuste en los recursos y capacidades puede mermar la operación general del sistema con relación a lo nuevo y también a lo que se venía haciendo bien.** Ampliar el ámbito de operaciones puede, en realidad, llevar a reducirlo por la merma general que se puede producir si no se hace cuidadosamente.

Un segundo elemento respecto a la misma pregunta tiene relación con los términos en los que dicha ampliación se puede operar: cualquier expansión en los sistemas de información debe, para ser sostenible y contribuir al desarrollo del sistema de información, ser consistente con el cuerpo de tareas, acciones y herramientas con las que se está operando.

Por ejemplo, construir un sistema de información sobre la educación no formal en una plataforma tecnológica diferente, con un cuerpo de definiciones distinto, desvinculado y potencialmente contradictorio con el que está en uso puede debilitar el sistema más que ampliarlo.

SECCIÓN 4 ANÁLISIS

El análisis de la información es una tarea clave que, entre otras cosas, permite convertir la evidencia en discursos significativos. Ahora bien, existen diversos tipos de análisis y las unidades a cargo de los sistemas de información deben juzgar con propiedad cuáles son las tareas analíticas que les toca realizar y cuáles pueden conllevar el riesgo de distraer recursos en tareas que otros agentes podrían conducir de una mejor manera.

Especialmente en la gestión pública es siempre clave tener presente que los recursos (los cuales pertenecen a la comunidad) que se destinan a un fin están dejando de dirigirse a muchos otros que también son importantes. Por lo tanto, una adecuada gestión pasa justamente por asignar los recursos prioritariamente a las tareas que son propias y que ningún otro agente podría realizar.

A manera de introducción, podemos organizar las posibles tareas de análisis en tres grandes ámbitos: (i) análisis de la calidad de la información que se genera; (ii) análisis destinado a respaldar los procesos de toma de decisiones; y (iii) análisis científico o académico de interés general.

Es posible afirmar que los sistemas de información públicos están obligados a desarrollar el primer tipo de análisis a efectos de mejorar de modo permanente su trabajo. Sin embargo, otros actores pueden hacer este mismo tipo de análisis y, mediante ello, contribuir a las tareas de los sistemas de información. Es de esperar que investigadores rigurosos se «ensucien las manos» tratando de entender cabalmente los atributos de la información que utilizan y que, al hacerlo, encuentren elementos que brinden pistas para la mejora de los sistemas de información.

Del mismo modo, los sistemas de información pública deben desarrollar algunas tareas de análisis destinadas a respaldar los procesos de toma de decisión, ya que al hacerlo logran una mayor comprensión de las necesidades públicas de información, con lo que cuentan con elementos para hacer ajustes y mejorar su capacidad de servicio.

Es claro que este es un terreno que parece ir más allá de las tareas de los sistemas de información, pero no es posible juzgar la relevancia de lo que se hace si no se entra en diálogo con las necesidades de información de la agenda pública en educación.

La idea profundamente entroncada en culturas marcadas por el empirismo acerca de la necesidad de «dejar que los datos hablen solos» no solo es, como ya se anotó, falaz desde un punto de vista epistemológico, sino que también conduce a la sordera tecnocrática respecto de las necesidades de mejora de la información.

Estas tareas de análisis pueden desarrollarse de modo autónomo, si es que se cuenta con recursos suficientes como para instalar unidades de análisis en las oficinas responsables de los sistemas de información, o de modo conjunto con otros agentes dentro o fuera del propio sector público.

El tercer terreno (análisis académico) es uno en el que los sistemas públicos de información solo han de brindar labores de respaldo facilitando el acceso a los datos y metadatos. Sin embargo, este tipo de análisis no debe menoscabarse, ya que también puede impactar de modo significativo tanto sobre el debate público de los temas educativos como en la identificación de oportunidades para mejorar la información.

4.1 MODELOS DE ANÁLISIS

El análisis de la información siempre consiste en utilizarla para construir un discurso significativo sobre un tema que importa. En ese sentido, y como cualquier otra tarea de generación de conocimiento, el análisis está siempre asentado sobre un marco de referencia conceptual que define o constituye el objeto sobre el que se quiere decir algo. Este marco conceptual puede ser o no explícito, pero está siempre presente haciendo posible la labor de conocimiento.

Por ejemplo, un sistema de información que define que solo debe desarrollar labores de análisis para mejorar la calidad de sus productos adopta, implícita o explícitamente, un punto de vista acerca del alcance de sus tareas y uno acerca de qué quiere decir con «calidad de sus productos».

Por otra parte, una definición de la calidad que solo afirma que esta consiste en satisfacer las demandas de los usuarios puede que se limite a simplemente observar el nivel de satisfacción de estos, lo cual sería suficiente desde una definición como la dada. Mientras tanto, una definición más amplia de la calidad, que considere, por ejemplo, las dimensiones descritas en la primera sección (acápito 1.2.4) necesitará verificar los atributos de cada una de esas dimensiones, es decir, conducirá a un análisis amplio de la dinámica institucional y de los aspectos técnico-metodológicos imbricados en la producción de información y la propia capacidad de servicio (que va más allá de la sola satisfacción de los usuarios)²³⁴.

²³⁴ Por cierto, esta idea de «satisfacción de los usuarios» se basa también en un supuesto básico de la economía positiva: si los consumidores (usuarios) son soberanos, es decir, pueden decidir independientemente cuáles son sus preferencias, no hay lugar para una tarea normativa que declare cuáles «deberían ser» dichas preferencias, basta con la manifestación de las mismas a través de un mecanismo que lo permita sin mayores distorsiones (el mercado). Esta visión de las cosas, si bien asume que los actores tienen un conocimiento sobre sus propias expectativas y preferencias que es preciso valorar y no desdeñar desde una perspectiva moralista que termina siendo profundamente autoritaria (alguien se supone que sabe mejor que la propia gente lo que a esta le interesa o conviene, v. gr. Rousseau), puede terminar desconociendo otros elementos que forman un cuerpo principal de problemas en la literatura económica: las diversas asimetrías (de información, de poder, etc.) que hacen que el mercado de competencia perfecta solo exista en los libros de texto como, en el mejor de los casos, un mecanismo heurístico, pero en ningún caso como una descripción de la realidad. Por otra parte, los mercados son mecanismos que permiten lograr equilibrios a partir de la acción de múltiples actores, pero no hay forma de argumentar que esos equilibrios representen, desde algún punto de vista normativo, un «óptimo». Por ejemplo, en un mercado abierto con una demanda poco solvente es posible encontrar equilibrios que permitan brindar acceso a un producto o servicio de bajo costo y muy mala calidad. Eso pasa, por ejemplo, en el caso de la telefonía (lo que se complica además por la tendencia a formar monopolios u oligopolios en este sector), en el abastecimiento de algunos servicios públicos (por ejemplo, el caso del agua «potable» en zonas rurales peruanas actualmente) o en el mismo caso

En cualquier caso, es de capital importancia tratar de hacer explícitos los modelos que se adoptan para el análisis, ya que estos **definen** los problemas a observar. Un potencial problema derivado de la falta de documentación explícita sobre los modelos adoptados puede llevar a desencuentros entre usuarios y productores de información, ya que estos no necesariamente comparten una misma forma de entender un fenómeno. El ejemplo de las páginas iniciales de la primera sección acerca del desajuste entre la pregunta sobre cobertura y la respuesta con ratios brutas de matrícula es solo una ilustración de este potencial desajuste.

Así, las tareas de «análisis de la calidad de los datos» están presididas por una concepción de dicha calidad que, al hacerse explícita, no solo permite el análisis, sino también comunicar sus resultados definiendo con claridad su alcance y significado.

En ese sentido, este volumen postula la conveniencia de abordar el análisis de la calidad de la información (primer tipo de análisis antes identificado) usando un marco analítico como el descrito en el acápite 1.2.4.

Del mismo modo, el análisis destinado a respaldar los procesos de toma de decisión (el segundo tipo mencionado) también requiere un marco analítico explícito. La sección segunda de este volumen se aboca en detalle a este tema (en particular los acápites 2.2 y 2.3) y resulta claro que este texto se adhiere a un esquema conceptual como el postulado en Unesco Santiago (2007, 2008). Asimismo, es necesario considerar el alcance de estas tareas de análisis y de sus correspondientes marcos analítico-conceptuales. Este texto se ha enfocado explícitamente en la información que respalde la formulación y ejecución de las políticas públicas, incluyendo el debate democrático a propósito de estas, sin entrar en detalles sobre aspectos más vinculados a la gestión subnacional, local o de programas específicos.

Finalmente, los análisis académicos (el tercer tipo de análisis señalado) son igualmente informados por marcos analíticos que se corresponden con las visiones y preocupaciones de quienes los desarrollan. En algunos casos, estos marcos pueden ser muy similares a los que se usan a efectos de desarrollar el segundo tipo de análisis aquí mencionado.

En las últimas dos décadas, sin embargo, el análisis de la educación desde una perspectiva profundamente influida por los modelos económicos ha logrado una presencia prácticamente hegemónica sobre la que es preciso reflexionar a efectos de no asumir acríticamente sus aspectos débiles, ni desdeñar prejuiciosamente sus virtudes. En los análisis

de la oferta privada de educación en sus distintos niveles: más de un país que ha abierto oportunidades para una mayor oferta privada en educación terciaria se encuentra, unos años después, con que gran parte de dicha oferta atiende una demanda de educación de bajo costo y baja calidad, lo que es, exactamente, lo que cabe esperar en un mercado con las características propias de sociedades con un amplio acceso a la educación secundaria, bajos logros de aprendizaje, escasa capacidad de gasto de las familias y una oferta deficiente de educación estatal. Por este tipo de consideraciones, la «satisfacción» de los usuarios (como la garantía del derecho a la educación) puede ser una medida engañosa de la calidad de las cosas, especialmente cuando se refiere a realidades que tienen una *raison d'être* fundamentalmente normativa.

influidos por la economía de la educación, por tanto, no debe sorprender el énfasis tanto en los retornos a la inversión y los resultados como en los procesos que conducen a dichos resultados, así como una preocupación por definir la «oferta» y la «demanda» educativas, aunque estos términos, muchas veces, resulten equívocos. Este tipo de análisis brinda importante información sobre la efectividad de las intervenciones, los impactos de largo plazo de la inversión en educación, las inequidades que se producen con relación a resultados educativos, etc. Es decir, produce información muy importante para entender los temas asociados a la garantía del derecho a la educación. Sin embargo, y no podría ser de otra forma, es un análisis menos sensible a las dinámicas de los agentes del sistema educativo, a los aprendizajes que van más allá de los saberes más instrumentales²³⁵, a la formación de representaciones (que es crucial para entender lo que hace o deja de hacer un sistema educativo), a los aspectos pedagógicos, etc. Dichos aspectos son justamente los que conforman un cuerpo importante de reflexión académica sobre la educación, el cual está fuera de la economía. Esto, por cierto, tiene que ver con los intereses y énfasis de cada disciplina y el error mayor consiste en esperar que, por ejemplo, la economía prescriba políticas pedagógicas o que la etnografía determine patrones de inversión de los recursos públicos en educación. Cada disciplina tiene su alcance y sus límites y el universo de las personas que toman decisiones hace bien en escuchar (siempre con profundo espíritu crítico) distintos aportes y proceder en consecuencia (entendiendo también que cada agente opera desde diversas visiones e intereses).

En forma sumaria, el marco conceptual que preside la organización de la información es el que estructura su análisis. De no ser este el caso, se producen inconsistencias y sinsentidos, dado que no existe correspondencia entre lo que el análisis quiere responder y aquello para lo que se construyó la información. Por ejemplo, el desajuste entre las preocupaciones por la cobertura y la conclusión y los sistemas de información más concernidos por los volúmenes de matrícula, que obedecen a preocupaciones centradas en la expansión y el financiamiento. Por tal razón, la discusión de la segunda sección toca directamente a las posibilidades analíticas de la información y a sus significados, lo que, en último término, no es sino una forma de explorar la capacidad de servicio de la misma, es decir, uno de los atributos claves de la calidad discutidos en la primera sección.

²³⁵ Aunque autores como James Heckman estén justamente preocupados por estos temas. En adición a sus textos, de los que una pequeña muestra es tomada como referencia en este trabajo (Heckman y Kautz 2012; Heckman *et al.* 2013; Heckman 2000, 2008), puede consultarse el sitio web dedicado a su obra (<<http://heckmanequation.org>>).

4.2 HERRAMIENTAS Y PROCEDIMIENTOS DE ANÁLISIS

Las herramientas y procedimientos de análisis a utilizar son de diversa naturaleza y han de corresponderse al propósito que se desea realizar y ser consistentes con la información a utilizarse.

El manejo de la información permite organizar **descripciones** sistemáticas de una situación a partir de la identificación de las características relevantes que definen el tema a abordar (y eso se deriva del marco de análisis) y de la aplicación de técnicas específicas en base a la información disponible.

Las descripciones de un fenómeno pueden ser muy básicas: de volúmenes, de magnitudes relativas, de la forma como se distribuye una variable en particular, de la forma como se distribuyen varias variables vistas de modo conjunto, etc.

A efectos de presentar la descripción de una situación particular, se cuenta con un conjunto de herramientas (muchas de las cuales son presentadas en el anexo 1) que refieren a las formas de presentar y representar distribuciones usando tablas y gráficos:

- valores absolutos y relativos;
- valores singulares o acumulados;
- tendencia central (la tipicidad en el comportamiento de una variable) y dispersión (la atipicidad en el comportamiento de una variable);
- tendencias históricas representadas mediante series de datos;
- asociación entre variables;
- estimaciones de valores poblacionales a partir de información muestral;
- pruebas de hipótesis estadística;
- pruebas de significancia estadística;
- análisis del comportamiento diferencial de ítems en una prueba; etc.

Asimismo, es posible pensar en **herramientas explicativas** que van más allá de la descripción, ya que cuentan con un marco conceptual que permite establecer relaciones causales que doten de sentido a los comportamientos que el análisis descriptivo muestra.

Lo central es tener presente el propósito del análisis y asegurarse de que la evidencia a utilizar y las herramientas analíticas se corresponden con dicho propósito. Esto significa, en último término, que el análisis no termina cuando se confecciona un cuadro o se presentan resultados de procedimientos estadísticos. Estos son productos intermedios que sirven para elaborar el producto que importa: **el relato significativo de lo que la evidencia muestra.**

Por ejemplo, imaginemos una situación en la que se está probando una nueva técnica de generación de información sobre niveles de alfabetismo. Una condición básica para hacer esto de modo robusto consiste en conducir una prueba de campo que permita validar los procedimientos y herramientas y hacer los ajustes que sean necesarios (en los procedimientos, las herramientas, la capacitación del personal de campo, su supervisión, etc.). El análisis de la información producida por la prueba de campo no puede perder de vista este propósito, pues es el que le permite construir un relato relevante.

Así, si en este ejemplo el análisis descriptivo de la información mostrase que existe un conjunto de casos en los que la prueba se administró a personas que no eran parte de la población objetivo, el análisis no termina ahí. Todo lo contrario, este es un punto de partida para explorar un poco más dicha anomalía (por ejemplo: ¿se distribuyen estos casos de modo aleatorio entre entrevistadores y equipos de campo? o, por el contrario, ¿todos los casos corresponden a un solo entrevistador o a un único equipo de campo con el mismo supervisor?).

Una vez identificado el problema (administración de la prueba a individuos fuera de la población objetivo) y explorado el patrón como dicho problema se manifiesta (por ejemplo, se concentra en un solo equipo de campo), es necesario concluir lo que importa en este caso: ¿qué aprendizajes se puede extraer de esta situación a efectos de mejorar la evaluación definitiva? Aquí resulta claro que se requiere que el equipo de campo implicado en este problema y su supervisor revisen las instrucciones para la selección de los individuos y discutan con los responsables generales de la evaluación qué fue lo que llevó al error y cuáles son las acciones necesarias para evitar un problema similar en la evaluación definitiva (por ejemplo, dedicar mayor tiempo a este tema en la capacitación del personal de campo).

De igual modo, si en el mismo esfuerzo de medición de competencias de alfabetismo se encuentra que un conjunto de los ítems que se han piloteado presenta sistemáticamente niveles de dificultad diferentes de los previstos (por ejemplo, nadie puede responderlos correctamente a pesar de que hay un porcentaje de personas que sí pudieron responder correctamente ítems teóricamente más difíciles), entonces es necesario identificar claramente dichos ítems y revisar si se trata de un problema de deficiente construcción o si, en realidad, la dificultad esperada estaba mal definida. Si se trata del primer caso, es necesario enmendar el ítem o descartarlo si es que se cuenta (como debería ser el caso) con un número suficiente de ítems equivalentes que podrían usarse en vez de este.

Un elemento clave del análisis es tener un hilo conductor coherente entre la pregunta que se busca responder, la información que se usa para ello y el tipo de conclusiones que se extrae de su análisis.

Por ejemplo, si se desea conocer cómo se distribuye una variable entre diferentes circunscripciones de un país de modo que se pueda dar cuenta de las disparidades que no son mostradas por las medidas de tendencia central (como una media), se puede

presentar un gráfico en el que se represente dicha medida de tendencia central y alrededor de esta los valores singulares de cada circunscripción. Luego, es posible ver si existe un patrón en el comportamiento de los datos (por ejemplo, las circunscripciones que presentan valores por encima de la media están ubicadas mayormente en el sur del país, mientras que las que se encuentran por debajo de la media son, mayormente, las que se encuentran en el norte). De ser este el caso, debe existir algún fenómeno sustantivo subyacente que explique este patrón no aleatorio.

Por supuesto que para poder identificar lo anterior es preciso tener claridad metodológica sobre lo que se está haciendo. En cierta oportunidad conocí un caso en el que una persona revisaba un gráfico de este tipo y expresaba su desacuerdo con la forma como se describía la situación (algunos casos por encima y otros por debajo de la media), ya que dicha descripción debía subrayar que «lamentablemente había algunos casos por debajo de la media». Esta preocupación es un completo sinsentido, ya que, por definición, una media supone que existan observaciones por encima y por debajo de ella; aspirar a que «todos estén por encima de la media» es un imposible aritmético (como observó alguien más en dicha circunstancia) y solo muestra que la persona que hizo la observación no tenía la menor idea de qué es una media y para qué sirve.

Lo mismo es de aplicación a cualquier otro tipo de herramienta de análisis: para usarla hay que tener claridad sobre lo que ella hace y asegurarse de que sea la herramienta adecuada para el propósito. Lamentablemente, las facilidades que brinda el *software* estadístico hacen posible que gente con pocas competencias use de modo absurdo cosas que no entiende: por ejemplo, graficar variables nominales como si fueran continuas, calcular promedios para variables nominales u ordinales (que fueron codificadas numéricamente para facilitar la captura de datos)²³⁶, etc.

Una vez más, es necesario reforzar que la buena intención no es suficiente para hacer bien las cosas, se requieren niveles mínimos de competencia profesional.

²³⁶ Conocí a una persona, que contaba con un doctorado de una prestigiosa universidad norteamericana, a la que le resultaba «natural» computar la proporción de personas de sexo masculino y femenino en un grupo usando promedios... Para ello codificaba los datos asignando un cero a un grupo (digamos, la población femenina) y un uno a los casos en el otro grupo (digamos, la población masculina) y luego calculaba un promedio de estos ceros y unos. En el caso de variables dicotómicas codificadas de esa forma, dicho promedio equivale **aritméticamente** a la proporción de personas en el grupo codificado como uno (de acuerdo a lo observado por Jacob Bernouilli en el siglo XVII), pero **conceptualmente** son cosas distintas: un promedio de la variable sexo, simplemente, es un sinsentido.

4.3 ALCANCE

Otro elemento de capital importancia consiste en definir con precisión el alcance de cada tarea de análisis singular.

Así como se puede tratar de brindar un relato significativo sobre diversos aspectos de la realidad educativa (por ejemplo, los informes mundiales de seguimiento de la EPT tienen esa intención), también es posible abocarse a temas singulares.

Una vez más, la determinación de cuál de estas opciones se ha de seguir depende directamente del propósito. Un análisis puntual de un tema específico puede ser de un gran valor si se requiere precisar dicho tema en el debate educativo. Así, un texto breve, con información clave y un relato muy puntual bien estructurado que articule la información de forma significativa, puede resultar muy valioso. El tema principal que subyace a esta discusión consiste en tener la capacidad para delimitar el alcance de lo que es objeto de análisis: una de las cosas más inútiles que pueden existir son textos que uno lee y finalmente no hay forma de saber cuál es el punto que se pretende discutir o afirmar.

Por ejemplo, si se quiere dar cuenta de un aspecto específico, como los niveles de conclusión de la educación primaria, se debe seleccionar los indicadores necesarios para dar cuenta de dicho tema y elaborar un relato puntual sobre el mismo. Si el texto empieza a mezclar temas (sin abordarlos en detalle), se pierde el propósito. Por ejemplo, si se discute sobre conclusión y solo sobre ese tema, no es necesario afirmar que concluir la primaria no asegura los aprendizajes esperados. Siendo eso completamente cierto, y seguramente existe suficiente evidencia para decirlo, ese no era el propósito del análisis. Así, es posible pensar en otro documento breve que dé cuenta de los niveles de logro académico de los estudiantes o en un documento diferente que aborde ambas cosas de modo conjunto y cada una en su especificidad²³⁷.

Incluso los informes que buscan mayores niveles de alcance han de acotar lo que observan y es mejor que lo hagan de modo explícito, pues siempre se va a encontrar aspectos que no fueron considerados en el análisis, ya que el enfoque que se da a la tarea define lo que se incluye y lo que se deja de lado.

Al mismo tiempo, existe en el mundo de la educación una tendencia paralizante que parte de una afirmación implícita que seguramente tiene algún grado de verdad: la

²³⁷ Por cierto, en esa dirección apunta el llamado de la Learning Metrics Task Force a contar con un indicador resumen de los temas de acceso y aprendizaje (Crouch *et al.* 2013; Learning Metrics Task Force 2013a) siguiendo lo sugerido por Lewin (2011).

educación es un fenómeno complejo y, por lo mismo, **todo** es importante. Pero como es humanamente imposible abordar **todo**, este tipo de enfoque no permite avanzar. Es mejor tener claridad sobre algunas cosas puntuales y relevantes (el marco de análisis define lo que se considera relevante para una perspectiva de análisis dada) que contar con una afirmación que por una supuesta intención abarcadora no permite precisar nada. En realidad, en muchos casos esta intención de abarcar **todo** no es sino una coartada para no ser preciso y no comprometerse con opciones de análisis o de política.

Ciertamente la realidad es compleja (no solo la educativa), de modo tal que la única forma de abordarla con cierto nivel de efectividad consiste en contar con marcos analíticos sólidos y labores de análisis precisas que se correspondan en alcance y uso de técnicas con lo que el marco define como relevante.

Por supuesto que es también necesario considerar que el análisis estadístico de la información es solo una pieza en un conjunto complejo de evidencias que se requieren para entender las cosas. La tarea de quienes trabajan en los sistemas de información consiste en brindar dicha pieza de la mejor manera posible y esto implica ser precisos y tener clara conciencia de los límites de esta información. Otras fuentes de evidencia pueden ser igualmente valiosas –en algunos casos incluso más valiosas– y nuestra responsabilidad, en adición a hacer nuestra parte bien, es tener apertura respecto de ellas²³⁸.

²³⁸ En este marco, la idea de incluir y promover trabajos etnográficos paralelos a la labor de evaluación de LAMP de la Unesco resulta particularmente ilustrativa. De hecho, dicho programa es la única evaluación de aprendizajes que incluye en su *toolkit* un documento de guías asociadas a la pertinencia de este trabajo etnográfico. Dos textos muy valiosos originados en esta experiencia se encuentran ya publicados (Maddox 2014, 2015).

4.4 EL ANÁLISIS COMPARADO

La comparación entre valores observados para una variable o conjunto de variables en un momento y lugar dados y los observados en otros momentos o lugares puede ser una forma de análisis muy importante y útil, aunque también encierra algunos riesgos.

En cualquier caso, las comparaciones se basan en asumir que la información que se compara es efectivamente comparable. Esto supone que los elementos de información han sido producidos con definiciones y procedimientos equivalentes, de modo que se pueda tener cierto nivel de certeza de que se está hablando de lo mismo. Esto, sin embargo, no es una tarea necesariamente fácil.

El primer elemento clave que define la utilidad del análisis comparado consiste en tener claridad acerca del propósito al que dicho análisis ha de servir.

Como ya se anotó²³⁹, el análisis comparado puede tener varias características: la comparación de diversas poblaciones respecto de fenómenos que se consideran universales; la comparación en el tiempo a efectos de constatar tendencias, así como progresos y retrocesos respecto de aspectos que se consideran importantes; y la comparación entre diversas unidades de información (jurisdicciones, países) cuyos propósitos pueden ser también diversos, como veremos más adelante.

La comparación temporal, para la misma unidad de observación, puede ser la forma de análisis comparado que menos dificultades presente, aunque no esté exenta de las mismas ya que las poblaciones no son estáticas en el tiempo. Por ejemplo, si se observa los niveles de aprendizaje en la población rural de un país dado a lo largo del tiempo, es posible que la tendencia que muestren los datos se encuentre afectada por la dinámica de migraciones internas. Así, si los datos de aprendizajes muestran una caída en los puntajes medios de los estudiantes, caída que acontece de modo simultáneo a un proceso de desruralización donde la población que se mantiene en las áreas rurales es un grupo cada vez menor y más homogéneo en términos de sus condiciones sociales (por ejemplo, los que permanecen son los más pobres y con menos oportunidades de migrar), la caída en los aprendizajes puede ser explicada más por este fenómeno que por un descenso en la «calidad de la educación» que se ofrece en las zonas rurales. (El sistema educativo, claro está, debió ser capaz de responder al cambio, pero ese es otro problema).

Las cosas se tornan aún más complejas cuando se comparan diversas unidades de información, en un momento dado (y, aún más si se hace en una serie de tiempo), respecto de una variable o conjunto de variables de interés.

²³⁹ En el acápite 2.4.8 «Medición de aprendizajes y el caso del alfabetismo adulto» (página 223).

La dificultad mayor consiste en que la comparación suele imbricar una idea de jerarquización (identificar a los que están «mejor» o «peor» a través, por ejemplo, de un *ranking*) y este no es siempre el propósito del análisis comparado. ¿Cuál es la utilidad de la identificación de posiciones relativas usualmente manifestadas mediante el uso de *rankings*?

Los *rankings* pueden tener un valor muy importante cuando, por ejemplo, se necesita seleccionar a un grupo de unidades de observación para un determinado propósito: aquellos estudiantes con puntajes más altos para acceder a la educación terciaria, las escuelas en condiciones de mayor pobreza para dirigir la atención prioritaria hacia ellas, etc.

Pero, en otros casos, la utilidad puede estar más vinculada a la lógica mercantil (informar a los padres acerca de cuáles son, supuestamente, las «mejores» escuelas para que puedan elegir mejor, asumiendo que pueden hacerlo y que la información, efectivamente, muestra las bondades o debilidades de las escuelas y no las ventajas y desventajas derivadas de la condición socioeconómica de las poblaciones a las que atienden) o a la competencia entre jurisdicciones (provincias o países); y esto no necesariamente dice mucho sobre aspectos educativos sustantivos.

Por ejemplo, ¿qué logra un país como Yemen al verse apareciendo en la última posición en un *ranking* de resultados en un estudio como Timms? Probablemente, este era un resultado esperable y su principal impacto sea algún titular desagradable en los medios de comunicación y acusaciones de culpabilidad en el mundo de la política, pero difícilmente un *ranking* permite aprender qué es lo que habría que hacer mejor. De hecho, quedar último no quiere decir gran cosa, pues último no necesariamente quiere decir **mal**, solo quiere decir **peor** que los otros.

En este sentido, el valor fundamental de estudios comparativos como el propio Timms no está en el *ranking* de desempeños sino en los análisis que se desarrollan a efectos de contar con información de niveles de desempeño en diversos contextos y sobre cómo se relacionan diversas variables con el desempeño. Así, es posible por ejemplo observar si las diferencias en los resultados (independientemente de quién esté primero o último) obedecen a variables vinculadas a fenómenos curriculares, a variables vinculadas a los docentes, a la dinámica de la escuela, etc.

El valor más importante de las comparaciones radica en esa identificación de las diversas dinámicas que subyacen a las diferencias en los resultados. Por esta razón, la IEA (que es, como vimos, el consorcio de investigación responsable de estudios como Timms) se preocupa por producir análisis detallados por país que se editan bajo la forma de «enciclopedias» (Mullis, Martin, Minnich, Drucker y Ragan 2012; Mullis, Martin, Minnich, Stanco, Arora, Centurino y Castle 2012) orientados a desentrañar lo que está detrás de los desempeños más que a una mirada superficial a los *rankings*.

Sin embargo, esto no necesariamente justifica extraer conclusiones del tipo: «Veamos qué es lo que están haciendo los que están primeros y hagamos lo mismo para mejorar»,

ya que, en primer lugar, la identificación de asociaciones entre variables no es equivalente a encontrar relaciones causales y, en segundo lugar, lo que da determinados resultados en un contexto no necesariamente los dará en otro.

Así, el análisis comparado encierra algunos riesgos, por un lado, derivados de las connotaciones emocionales y de las agendas o intencionalidades que se manifiestan alrededor de los *rankings*²⁴⁰ y, por otro lado, a propósito de extraer conclusiones sobreesimplificadas acerca del tipo de lecciones que las diferencias entre las unidades de información ilustran.

Lo fundamental a tener en cuenta es que una situación concreta es siempre concreta, justamente porque, usando la expresión de Marx, es la síntesis de diversas determinaciones, la unidad de la diversidad (1983: 14) y, por lo mismo, el análisis comparado debe tomar en cuenta esa complejidad (que jamás es ilustrada en un *ranking* univariable) y es en ello donde radica su mayor riqueza para sugerir pistas y desatar debates que informen los procesos de toma de decisión.

²⁴⁰ Por ejemplo, véase la manera como el gobierno británico define el tema educativo basándose en los resultados de PISA: «[...] lo que realmente importa es cómo nos va en comparación con aquellos con los que competimos a nivel internacional. Eso es lo que definirá nuestro crecimiento económico y el futuro de nuestro país» (Reino Unido-Department of Education 2010: 3; traducción propia).

4.5 ÍNDICES

Un último elemento sobre el que cabe llamar la atención en esta sección tiene que ver con la construcción de índices.

Un índice es básicamente una medida resumen que resulta de combinar diversos elementos de información expresables o reductibles a una escala común. Índices como el de desarrollo humano (UNDP 2009) o el de desarrollo de la EPT (Unesco 2006) son ilustraciones de esto.

Ambos índices conjugan un conjunto de variables en una medida única que se considera un buen resumen de una situación que dicho conjunto de indicadores busca representar (el desarrollo humano, por ejemplo) y donde cada indicador o subconjunto es considerado suficiente para dar cuenta de una de las dimensiones que componen dicha situación. En ese sentido, un índice es una medida resumen de un fenómeno complejo compuesta a partir de información sobre las partes de dicho fenómeno.

Por lo mismo, el valor de los índices reposa en una robusta identificación de las dimensiones del problema, en el valor de las piezas de información que han sido seleccionadas para componerlo (¿dichas piezas representan de modo suficiente y consistente cada una de las dimensiones que se incluyen en el índice?) y en la manera como dichas piezas componen el índice, es decir, en el algoritmo usado para computarlo (¿todas las dimensiones tienen igual ponderación o su ponderación es diferente?, ¿cómo se definió dicha ponderación?).

Así, la construcción de un índice es una tarea que supone mucho cuidado. Por ejemplo cabe preguntarse si combinar los ratios brutos de matrícula de tres niveles educativos, como solía hacer el índice de desarrollo humano, es la mejor manera de representar la situación educativa. De hecho, resulta importante considerar aquí que, dado que estas ratios de matrícula pueden subir (o bajar) tanto por buenas como por malas razones, asumir que valores mayores representan mayores niveles de desarrollo humano resulta, al menos, debatible. Por ejemplo, un país en el que todas las demás variables se mantuvieron constantes entre los años t y $t + n$ y solo las tasas de repetición en educación primaria experimentaron un significativo incremento en dicho período mostrará una «mejora» en el valor del índice.

Preguntas similares pueden plantearse respecto del índice de desarrollo de la EPT: ¿Es la **tasa de supervivencia al quinto grado** una medida suficiente para representar la preocupación por la calidad educativa? ¿Acaso los valores de paridad de género de las ratios brutas de matrícula dicen algo sobre la equidad?

Existen diversos métodos para componer índices. Los más simples, como los dos aquí citados, parten de una selección de indicadores basada en un juicio formulado por aquellos que los componen. Este es el conjunto de indicadores que necesitamos para medir lo que nos importa o, al menos, estos son los indicadores para los que se cuenta con información para un mayor número de unidades de observación (países, en estos casos). Luego se procede a definir un algoritmo de cálculo basado también en el juicio de aquellos que construyen el índice (por ejemplo, asignar ponderaciones iguales).

Las formas más complejas de abordar el tema de la construcción de índices se basan en el modelamiento de datos. En general, estos procedimientos tienen una forma básica: el índice da cuenta de una variable fundamental que no es directamente observable, es decir, que se encuentra como «latente» (como el desarrollo humano) a partir de un conjunto de variables manifiestas u observables. En este esquema, la tarea fundamental consiste en saber si las variables manifiestas son, efectivamente, expresiones de la variable latente, lo que se hace mediante análisis de asociación (todas ellas aparecen asociadas entre sí) y de control de la independencia entre las mismas (la asociación entre ellas desaparece cuando se controla el efecto de la variable latente).

A este universo de modelamientos pertenece el análisis factorial y el de componentes principales (que usa variables manifiestas y latentes continuas), el análisis de rasgos latentes²⁴¹ (que usa variables observadas categóricas para dar cuenta de una variable latente continua), el análisis de clases latentes (que se aplica a variables manifiestas y latentes de tipo categórico), etc.²⁴²

En cualquiera de los casos, la selección de indicadores y los procedimientos usados para definir los algoritmos de cálculo deberían estar plenamente documentados.

Finalmente, existe un problema muy importante con una medida resumen o índice compuesto como los aquí descritos: por su propia naturaleza nunca son suficientes para brindar un mensaje significativo más allá de la representación sintética. ¿Qué es lo que explica un valor alto o bajo en el índice? El valor obtenido puede ser el resultado de diversas combinaciones de valores intermedios posibles, de tal suerte que dos unidades de información con un índice del mismo valor pueden, en realidad, tener situaciones completamente diferentes.

Imaginemos un índice simple compuesto de dos indicadores, el primero mide la dimensión α y el segundo la dimensión β . Imaginemos que tanto el índice como los dos indicadores que lo componen están en una escala de cien puntos y que ambos indicadores tienen la misma ponderación en el cálculo del índice. Un país x puede presentar un valor de 80 puntos para α y un valor de 20 puntos para β , de modo tal que el valor del índice compuesto será: $(80 + 20) / 2$, es decir 50. Ahora imaginemos un

²⁴¹ La teoría de respuesta al ítem usada en sicometría es un modelo de análisis de rasgos latentes.

²⁴² Para una discusión conceptual detallada véase, entre otros: Bollen (2002); Borsboom *et al.* (2003); Collins (2010); Henry (1999); Lazarsfeld y Henry (1968).

país y , donde α tiene un valor de 52 puntos y β un valor de 48 puntos; el índice para este país arrojará un valor de: $(52 + 48) / 2$, es decir, 50. Finalmente, imaginemos un tercer país z , donde los valores de α y β son: 20 y 80 puntos respectivamente; en este caso, el valor calculado para el índice será: $(20 + 80) / 2$; es decir, 50 puntos. Es decir, los tres casos muestran un valor idéntico para el índice aunque presentan patrones de comportamiento muy distintos en los dos indicadores.

4.6 COMENTARIO FINAL

Más allá de las decisiones que se tomen acerca de los alcances que se quiera dar a las labores de análisis, es preciso considerar que hay algunos principios y tareas ineludibles en este terreno.

Estas tareas tiene que ver con: (i) el diálogo permanente con la política y el debate público a efectos de identificar los tópicos más relevantes a abordar; (ii) el diálogo permanente con la investigación social y educativa para contar con las mejores herramientas posibles para el análisis.

Un peligro siempre presente se origina en no escuchar lo que sucede en el entorno y asumir una posición fundacional, de complejo adánico o solipsista. Retomar lo ya hecho, escuchar lo que hacen otros y aprender del diálogo son elementos claves para que las labores de análisis, pero no solo ellas, tengan un valor y representen una contribución.

Finalmente, la rigurosidad académica es un imperativo ineludible. Revisar a fondo los propios supuestos y «ensuciarse las manos» con la evidencia, de modo que se comprenda mejor su alcance y sus limitaciones, son tareas básicas que cualquier análisis que se precie de decente no puede soslayar.

REFLEXIONES FINALES

Empiezo estas reflexiones finales luego de varios meses en los que he escrito las notas que luego derivaron en las páginas precedentes; páginas que fueron apareciendo una tras otra de un modo inadvertido, como se señala en la introducción, pero que han adquirido una extensión que no esperaba cuando en agosto de 2008 elaboré el primer esquema que sirvió como punto de partida para esta empresa.

Este texto, además, ha sido escrito como una iniciativa personal llevada adelante durante mis horas no dedicadas al trabajo²⁴³, en mi residencia y en un número que no podría determinar de salas de espera de aeropuertos, aviones y habitaciones de hotel en cuatro continentes²⁴⁴. Sin embargo, redactarlo no ha sido una carga, todo lo contrario, el proceso ha supuesto aprender un poco más y recapitular muchas experiencias gratas y otras que, si bien no lo fueron, también derivaron en aprendizajes importantes. La composición del borrador final, además, se vio enriquecida por el ambiente de intercambio y riqueza profesional posible en el Centro Bellagio, gracias a los extraordinarios compañeros que tuve durante esas semanas.

Quisiera hacer uso de este espacio final para subrayar algunas ideas clave y afirmar elementos que me resultan centrales a la hora de ponderar lo aquí escrito.

- (i) Más allá de mi responsabilidad individual en la producción de este texto, el mismo es el resultado de un proceso más general mediante el cual diversas personas en América Latina se comprometieron a desarrollar los sistemas

²⁴³ Excepto en el necesario periodo de revisión y composición del borrador final en marzo de 2014, durante la residencia en el Centro Bellagio de la Fundación Rockefeller, periodo posible, entre otras cosas, gracias al apoyo que me ha brindado la Universidad del Pacífico.

²⁴⁴ Recuerdo acá una grata frase de una muy apreciada colega uruguaya: «Jamás imaginé que la estadística educativa me llevaría a conocer el mundo».

de información educativa de sus propios países. En efecto, mi contribución personal la hice mayormente en mi país; luego de ello, solo ha consistido en ser capaz de prestar atención a mis inquietudes, ordenarlas, darles un poco de forma conceptual y sentido de propuesta, así como contribuir a que la Unesco se haya abocado de modo decidido a este proceso regional de desarrollo tanto en el plano metodológico y conceptual como en la acción práctica de respaldo a los países.

- (ii) Los sistemas de información educativa no pueden ser vistos en el aislamiento de los aspectos técnicos que los componen. Estos no tienen sentido sin un propósito mayor, sin un entronque con una perspectiva y con determinadas preocupaciones. **La educación es un derecho humano fundamental** y esa es la preocupación que organiza todo lo aquí presentado, así como el propio hecho de haber puesto esto por escrito. No creo que nuestros países se puedan ver beneficiados mediante iniciativas demasiado centradas en los aspectos utilitarios de la educación y que usualmente son manejadas de un modo que bien podría ser descrito como colonial, donde solo algunos tienen la «sartén por el mango» y a otros nos correspondería ser espectadores, consumidores o simples seres pasivos en un juego que nos es ajeno tanto por contenido como por asimetrías de poder.
- (iii) Avanzar implica poseer convicciones firmes al tiempo que estar abiertos al cambio. Debemos tener mucho cuidado con el temor a mejorar. Si bien reconocer las propias debilidades es un buen punto de partida para mejorar, a veces hacerlo nos cuesta mucho, pues puede implicar enfrentarnos con nuestros propios problemas de autoestima e inseguridades. ¿Cuántos de nosotros estamos listos para acudir donde nuestros ministros y decirles algo que podría sonar a: «Durante los últimos diez años he estado a cargo del sistema de información y he producido religiosamente ratios brutos de matrícula a las que sin mucho sentido he llamado tasas de cobertura bruta pero, en realidad, creo que sería mejor dejar eso de lado y hacer otra cosa que entiendo va a servir más?» No es fácil hacerlo, más fácil es mantener el statu quo y no mejorar; pero eso no es lo éticamente correcto.
- (iv) La transparencia en la gestión de los sistemas de información implica: (a) mantener y fortalecer el carácter público de la información solo prestando atención a limitaciones propias de la reserva estadística (y a la protección de la intimidad de las personas) necesaria para asegurar la solidez de la misma (y los derechos inalienables de las personas); (b) asegurar y facilitar el acceso a la información; (c) documentar exhaustivamente los atributos de la información y de los procedimientos que se encuentran detrás de ella; (d) asegurar que tanto la información como su documentación sean inteligibles; y (e) brindar asistencia a los usuarios.

Dicha transparencia no es un detalle secundario en la operación de los sistemas de información; muy por el contrario, es un atributo central que toca a los fundamentos democráticos que se encuentran a la base de las operaciones del servicio público. Justamente, una diferencia clave entre sistemas de información al servicio de los ciudadanos y un negocio cualquiera está dada por esta transparencia. Cuando al menos uno de los agentes implicados en la generación de información pública retiene para sí elementos claves de dicho proceso de generación de información, tal agente está privatizando (en pro de su propio beneficio) algo que es sufragado por la contribución general de los ciudadanos. Lamentablemente hay muchos ejemplos de esto tanto a escala nacional como internacional²⁴⁵.

Al mismo tiempo, la transparencia en el manejo de los sistemas de información es una herramienta de su mejora continua. Justamente son los usuarios los que pueden, con mayor relevancia, encontrar áreas débiles o vacíos que merecen ser abordados. Por cierto, el diálogo con los usuarios no se resuelve mediante rituales tipo «comité consultivo», sino mediante el trabajo conjunto.

- (v) Parafraseando a Morpheus²⁴⁶, es importante entender que no es lo mismo conocer el camino que recorrerlo y, agregado, tampoco es lo mismo querer hacerlo. Este texto solo puede pretender ser una ayuda para conocer el camino hacia sistemas de información educativa que honren de una mejor manera lo que las personas tienen derecho a esperar de él, pero eso no es gran cosa. Querer recorrer el camino de la mejora y hacerlo son los desafíos que cada equipo nacional ha ido asumiendo estos años y eso es lo que realmente cambia las cosas.
- (vi) Las ideas aquí desarrolladas o recapituladas han sido posibles mediante la combinación de una vocación por la rigurosidad que no se traduce en un apego a la técnica, sino que está puesta al servicio de un foco en los problemas educativos y éticos sustantivos. Aferrarse a las técnicas de lo que hacemos puede llevar a seguir padeciendo, por ejemplo, al tratar de encontrar alguna utilidad a las ratios brutas, al «ajustar» las tasas netas, al desvivirse por rescatar los modelos de cohortes, etc. La creencia básica que sustenta todo el esfuerzo detrás de lo aquí escrito es la necesidad de despojarse de ese aferramiento y centrarse en lo que importa: **los**

²⁴⁵ Me tocó conocer el caso de un ministro y su consultor «estrella» que retuvieron para sí la propiedad de la cartografía digital de un país para usarla en su propio negocio de consultoría tras su paso por la gestión. Lo mismo ocurre con ciertos estudios internacionales que implicaron una inversión millonaria respecto de los cuales una institución contratista —con ventajas tributarias en su país de sede, dado su carácter de organización «sin fines de lucro»— retiene el control de procedimientos estadísticos clave, lo cual, por no ser accesibles, hace imposible replicar de modo independiente los procedimientos seguidos. Por cierto, este último es un ejemplo también de mala práctica técnico-científica, ya que, por definición, la investigación científica debe ser transparente.

²⁴⁶ Capitán de la nave Nabucodonosor en la película *The Matrix* de los hermanos Wachowski.

propósitos. Buscar una solución que sea útil supone el desapego para poder revisar y criticar (lo que hay y lo que uno mismo hace), que son las condiciones para mejorar.

- (vii) Hace un tiempo escuché a alguien decir en un evento internacional: «No hay nada más triste que una tabla estadística». En realidad hay muchas cosas más tristes que dichas tablas, por ejemplo:
- (a) No tener una tabla estadística que se necesita.
 - (b) Tener una tabla estadística con información que no se corresponde con lo que se necesita.
 - (c) Tener una tabla estadística con la información que se necesita pero producida de una manera débil.
 - (d) No darse cuenta de lo anterior y usar irresponsable y/o incompetentemente la información.
 - (e) Tener la tabla estadística con la información necesaria y producida de un modo robusto, pero no saber utilizarla.

Esta lista podría seguir *ad infinitum*, pero si hay algunas cuestiones tristes sobre las que me gustaría llamar la atención, ellas son tanto el hecho de que algunos expertos pretendan serlo sin tener idea de la evidencia que la información brinda y sin ser capaces de exigir tal evidencia a los sistemas de información; así como que tomadores de decisiones procedan sobre la base de sus prejuicios sin prestar la menor atención a la evidencia o que solo lo hagan para tratar de maquillarla o para usos rituales; y, por supuesto, también es muy triste que la producción de información no preste debida atención a los propósitos que la fundan.

- (viii) Entender, o por lo menos aceptar, la complejidad de los asuntos humanos es de capital importancia para evitar soluciones demasiado simples o triviales como para ser útiles o demasiado elaboradas (o enrevesadas) como para tener sentido para cualquiera que no sea el que las pergeñó. Las personas actuamos en el mundo del que somos parte, y al hacerlo lo construimos y reconstruimos continuamente. Cada uno de nosotros, con sus propias motivaciones, intereses, visiones y cuotas de poder material o simbólico, actúa.

Por lo mismo, existe una saludable tendencia actual a entender que la formulación de políticas ha de alejarse de pretensiones fundamentalmente **prescriptivas** (decirle a la gente lo que tiene que hacer) y aproximarse más a posturas **habilitadoras** de la acción de las personas. Por ejemplo, un docente no enseñará mejor porque se le diga que debe hacerlo, sino y fundamentalmente porque tiene **interés, condiciones y puede hacerlo**; en lo cual, el interés se vincula con la estructura de reglas e incentivos que hacen a la carrera docente, las condiciones con los recursos disponibles y el poder hacer con las competencias

profesionales y el soporte material y profesional con el que el docente puede contar en su tarea.

Estas consideraciones son de capital importancia para entender cuál es el ámbito en el que se mueve la acción política que los sistemas de información han de servir.

Termino expresando mi deseo de que las personas con las que recorrí el camino de aprendizajes que este texto condensa puedan reconocerse un poco en él. Esa sería la mejor forma de reconocimiento a sus contribuciones que se me puede ocurrir. Pero, aún más importante que ello, espero sinceramente que quienes están comprometidos con el desarrollo de los sistemas de información educativa, especialmente en la región a la que pertenezco, encuentren en estas páginas elementos que los apoyen en la tarea, ideas que les sugieran pistas para mejorar y un desafío de aprender, de modo que lo aquí escrito solo sea un paso más entre muchos otros. Si este fuera el caso, estoy seguro de que mis sinceros compañeros en esta travesía estarán tan satisfechos como yo por lo que juntos pudimos hacer a lo largo de los últimos años.

ANEXOS

ANEXO 1. CONCEPTOS ESTADÍSTICOS BÁSICOS

Si bien existe un número importante de manuales o introducciones a la estadística, las personas que trabajan en los sistemas de información en los ministerios de Educación no necesariamente acceden a ellos con facilidad. Por esta razón, me pareció oportuno incluir como suplemento un brevísimo –y nada cargado a la técnica– sumario de conceptos básicos de estadística que creo que podrían valorar mis colegas de los ministerios como un punto de entrada a enriquecer posteriormente.

1. VARIABLES Y NIVELES DE MEDICIÓN

Un primer elemento básico a considerar es el hecho de que la estadística opera con variables, las mismas que por su naturaleza tienen un determinado nivel o escala de medición.

Una **variable** es una operacionalización de un concepto, de forma tal que pueda ser observable a través de determinados elementos de información.

Así por ejemplo, el **nivel de escolaridad alcanzado** por una persona es un concepto complejo que pueda abarcar un conjunto muy amplio de aspectos. Sin embargo, a efectos de desarrollar estudios e investigaciones, es usualmente traducido en el número de años de educación o el último grado aprobado. Ambas variables o formas de observación tornan asible el concepto general y al hacerlo permiten observarlo. Ahora bien, como toda observación y conocimiento, estos hacen posible captar algunos aspectos del fenómeno y no **toda** su complejidad¹. De esta forma, resulta muy importante tener en cuenta que la operacionalización debe ser **suficiente** en términos de poder dar cuenta de aspectos que se consideran claves del concepto objeto de interés. En algunos casos, es posible que una operacionalización requiera usar más de un elemento de observación.

Ahora bien, las variables reciben ese nombre justamente por su posibilidad de presentar valores observables distintos. Un concepto que tiene una presencia homogénea en una determinada población² no es una variable sino una constante y, por lo tanto, no será objeto del tratamiento que aquí se presenta.

Por otra parte, esos valores diferentes que puede adoptar una variable reciben el nombre de **categorías o clases de la variable**. Ahora bien, dependiendo de la naturaleza de

¹ La búsqueda de un conocimiento **total** es una quimera, un tema de la ciencia ficción o algo que pertenece legítimamente a la discusión religiosa, pero no al mundo del quehacer científico.

² Entiéndase por **población** a un universo de unidades de observación que puede estar compuesto por personas o por cualquier otro tipo de unidad (centros educativos; secciones; etc.).

la variable, dichas categorías presentan entre sí relaciones que pueden ser de los siguientes tipos³:

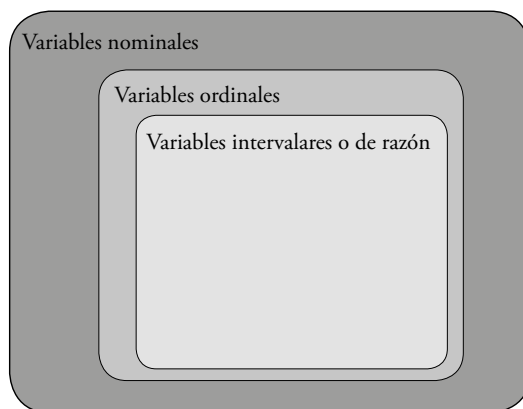
- **Diferencias exclusivamente de calidad.** Es decir, las categorías son diferentes entre sí, pero no existe una relación de orden o de magnitud entre ellas. Por ejemplo, la variable sexo presenta dos categorías entre las cuales solo existe una diferencia cualitativa, mas no de orden o de magnitud. Estas variables son conocidas como **nominales** o categóricas.
- **Diferencias de calidad y orden.** Es decir, las categorías de la variable no son solo diferentes entre sí, sino que existe una gradación entre ellas. Por ejemplo, cuando se habla del nivel socioeconómico de la población usualmente se usan estratos que van desde un valor menor (estrato socioeconómico bajo, pobre, indigente, según sea el caso) hasta uno mayor (estrato socioeconómico alto, rico, pudiente etc.). Sin embargo, se trata de una diferencia de grados no cuantificable, es decir, no es posible decir que, por ejemplo, el estrato alto es cuatro veces más que el estrato bajo. Estas variables son conocidas como **ordinales**.
- **Diferencias de calidad, orden y magnitud.** Es decir, variables en las que sus categorías no solo son diferentes y graduadas, sino que además existe una diferencia de magnitud entre las mismas. Por ejemplo, la edad en años cumplidos. Una persona de treinta años no solo tiene un valor diferente que el de una persona de quince, sino que dicho valor es mayor y además lo es en dos veces. Estas variables pueden, por tanto, ser objeto de múltiples análisis de tipo numérico. Asimismo, este tipo de variables se distinguen entre aquellas en las que es posible constatar un valor natural equivalente a cero, como por ejemplo la edad (variables **de razón**) y otras en las que la escala puede tener una forma arbitraria (donde un punto cero no necesariamente tiene un valor particular), como la temperatura si es medida en grados Celsius o Fahrenheit (variables **de intervalo**).
Estos dos últimos tipos de variable son representados y tratados numéricamente, por lo que es preciso también considerar que sus valores pueden ser **continuos** o **discretos**. Una variable es continua cuando entre dos valores cualquiera de la misma es posible que exista otro, mientras que en el caso de las variables discretas no existen valores intermedios entre sus categorías. Así, el número de años cumplidos por una persona es una variable discreta (solo presenta valores enteros mayores o iguales a cero), mientras que el nivel de ingresos de una familia medido en unidades monetarias es una variable continua.

³ Stevens (1946) fue quien sistematizó estas ideas definiendo cuatro escalas de medición: nominal, ordinal, de intervalo y de razón. Estas escalas se explican a continuación en el texto.

Tener en cuenta el nivel de medición o escala de las variables es un aspecto clave del trabajo estadístico, ya que de este depende el tipo de tratamiento estadístico que se puede hacer con la variable en cuestión. Un ejemplo trivial puede ilustrar esto: es posible obtener un promedio del ingreso monetario de un grupo de personas, pero no lo es del sexo de las mismas. Si bien esto es bastante obvio en el ejemplo mencionado, hay casos en los que ello no necesariamente resulta evidente, especialmente cuando se usan códigos numéricos para representar las categorías de las variables y cuando mediante un *software* estadístico se procesan dichos códigos como si la naturaleza de la variable permitiera un tratamiento numérico.

Nótese, asimismo, que los niveles de medición presentan una relación de inclusión entre ellos. Así, las variables de un nivel superior (por ejemplo, de razón) son también del nivel menor (por ejemplo, ordinales y nominales), por lo que pueden ser tratadas como tales. El ingreso monetario, por ejemplo, siendo una variable continua puede ser tratada como ordinal (un segmento de la población tiene ingresos mayores que otro). Esta observación es importante a la hora de tratar simultáneamente con dos o más variables y las relaciones entre ellas: si las variables son de dos niveles de medición diferentes, es posible tratar dichas variables de modo conjunto usando los procedimientos propios del nivel inferior. Esto se ilustra en el gráfico 1A

Gráfico 1A. Niveles (o escalas) de medición estadística y su jerarquía



2. ELEMENTOS DE ESTADÍSTICA DESCRIPTIVA⁴

La estadística permite, en primer lugar, usar un conjunto de variables para brindar descripciones de lo observado. La descripción puede basarse en el uso de una o más variables y es directamente dependiente del nivel de medición de las mismas.

⁴ Nótese que las fórmulas que se incluyen de aquí en adelante en este anexo solo tienen un propósito ilustrativo (en varios casos, el algoritmo de cálculo expresa con mucha más claridad las características de un

2.1 Descripciones basadas en una sola variable (univariabes)

- (i) **Representación.** Un primer uso de la estadística descriptiva está dado por la presentación y representación de la información. A estos efectos, un recurso inicial consiste en el uso de **distribuciones de frecuencias**.

Una distribución de frecuencias es una tabla que muestra la cantidad de ocurrencias (frecuencia) de cada una de las categorías de la variable en el conjunto observado bajo la forma general de la tabla 1A.

Tabla 1A. Título (debe incluir al menos el nombre de la variable, la población de referencia, el período de referencia y la unidad de medida)

Categorías de la variable	Frecuencias
A	X
B	Y
C	Z
Total	$X + Y + Z$

Fuente: (incluir una referencia explícita a la fuente de la información).

Notas

(incluir las notas explicativas que puedan ser necesarias).

A continuación, en la tabla 1B se presenta una ilustración de distribución de frecuencias.

Como se puede notar, la tabla 1B se refiere a una variable de nivel nominal, por lo que sus categorías (estatal, no estatal subvencionado, no estatal [otras]) no tienen ningún orden en particular y podrían, de hecho, aparecer en cualquiera en la tabla.

Tabla 1B. Centros educativos según modalidad de gestión, País A, 2012

Modalidad de gestión	Cantidad de centros educativos
Estatad	34.835
No estatal subvencionado	2.869
No estatal (otras)	3.279
Total	40.983

Fuente: Censo Escolar 2012.

Nota

Solo incluye los centros educativos de prebásica, básica y media que reportaron información correspondiente al Censo Escolar 2012 (matrícula final). Valores hipotéticos para fines de ilustración.

procedimiento). El *software* estadístico de uso corriente puede hacer todos estos cálculos, de modo que lo más importante aquí es notar la lógica de lo que se hace, la correspondencia entre los procedimientos que se seleccionan y los propósitos asociados a ellos.

Por otra parte, esta distribución de frecuencias presenta solamente las **frecuencias absolutas**, siendo posible hacer una presentación alternativa o complementaria de **frecuencias relativas**, como en la tabla 1C.

Tabla 1C. Centros educativos según modalidad de gestión, País A, 2012

Modalidad de gestión	Centros educativos	
	Cantidad	Porcentaje
Estatal	34.835	85,0
No estatal subvencionado	2.869	7,0
No estatal (otras)	3.279	8,0
Total	40.983	100,0

Fuente: Censo Escolar 2012.

Nota

Solo incluye los centros educativos de prebásica, básica y media que reportaron información correspondiente al Censo Escolar 2012 (matrícula final). Valores hipotéticos para fines de ilustración.

En el caso de variables ordinales o superiores, es además posible –dado que las categorías de la variable **deben** aparecer ordenadas de menos a más o de más a menos– hacer **distribuciones acumuladas** (tanto absolutas como relativas). Esta forma de presentación permite identificar la cantidad o la proporción de observaciones que caen en una categoría dada o menos que ella (o más que ella, dependiendo de cómo se ordene la acumulación).

Así, por ejemplo, es posible tener la tabla 1D.

Tabla 1D. Distribución de matrícula según tamaño de los centros educativos, País A, 2012

Tamaño del centro educativo	Centros educativos			
	Cantidad	Porcentaje	Cantidad acumulada	Porcentaje acumulado
Menos de 50 estudiantes	26.594	64,9	26.594	64,9
Entre 50 y 99 estudiantes	9.987	24,4	36.581	89,3
Más de 100 estudiantes	4.402	10,7	40.983	100,0
Total	40.983	100,0		

Fuente: Censo Escolar 2012.

Nota

Solo incluye los centros educativos de prebásica, básica y media que reportaron información correspondiente al Censo Escolar 2012 (matrícula final). Valores hipotéticos para fines de ilustración.

Nótese que, en el caso de las variables de razón o intervalares, sus categorías pueden ser valores puntuales o intervalos. En el ejemplo de la tabla 10, sería

posible mostrar las cantidades de centros educativos para cada uno de los valores posibles, aunque se optó por usar intervalos ya que lo anterior produciría una tabla muy grande y difícil de leer, además, probablemente no mostraría lo que importa en esta tabla: la concentración de la matrícula en centros educativos de menor tamaño. Al usar intervalos es importante notar que estos pueden, o no, ser de igual tamaño o amplitud y que las referencias a los intervalos se hacen en atención a su nombre o tomando como referencia los **puntos medios** o **marcas de clase**. Esto es particularmente importante cuando no se cuenta con los datos originales y es necesario hacer operaciones con la información (por ejemplo, obtener un promedio del número de alumnos matriculados en los centros educativos)⁵.

Ahora bien, esta forma de representación de la información puede ser convertida en formas gráficas. Así tenemos, por ejemplo, **gráficos de barras** o **histogramas**.

Un gráfico de barras contiene en el eje de las abscisas las categorías de las variables y en el eje de las ordenadas las frecuencias (absolutas o relativas). Es importante notar que las barras deben representar de la mejor manera posible las categorías de las variables, así en el caso de las variables nominales y ordinales es importante notar que las barras deben ir separadas unas de otras, ya que el eje horizontal no representa un continuo de valores, mientras que en las variables de razón o intervalares, dado que las barras representan un rango de valores, el espesor de la barra debe corresponderse con la amplitud del intervalo representado⁶.

Otras formas de representación gráfica son los **gráficos de puntos, de líneas, los pie charts** (o gráficos tipo torta) y las **ojivas** (para el caso de las frecuencias acumuladas ascendentes o descendentes).

- (ii) **Tendencia central y dispersión.** Un segundo elemento clave en la descripción del comportamiento de una variable tiene que ver con la identificación de situaciones típicas y, de modo complementario, con la atipicidad.

En efecto, muchas veces es requerida una descripción sintética de una variable, la que es dada por las situaciones típicas. Un ejemplo usual de esto es, por ejemplo, cuando se habla del promedio de edad de un grupo de estudiantes. Sin embargo, es importante notar que dos situaciones distintas pueden dar lugar a promedios similares: por ejemplo, en una población de dos alumnos que tienen uno diez y

⁵ Por otra parte, cabe anotar que los cálculos, siempre que sea posible, deberían realizarse con los datos originales y no con los datos agregados, pues la agregación resta precisión.

⁶ Si bien ambos diagramas son similares, hay una diferencia muy importante entre ellos: los gráficos de barras se usan para representar las frecuencias de las categorías de variables no continuas (usualmente ordinales), de donde el eje horizontal no representa un continuo y por ello las barras van separadas unas de otras y lo importante es su altura; por su parte, los histogramas se usan para representar variables continuas, de donde el eje horizontal es un continuo y las barras son adyacentes. En los histogramas, las barras pueden ser de diferente ancho (ya que pueden representar segmentos distintos del continuo representado en el eje horizontal) y lo importante es el área que cubre cada barra, ya que es un diagrama que representa una densidad.

el otro veinte años, el promedio será quince, al igual que en una de dos alumnos que tuvieran catorce y dieciséis años. Por esto, entonces, junto a los valores típicos es importante conocer la atipicidad o la dispersión.

A efectos de dar cuenta de la tipicidad (tendencia central) y la dispersión existen diversas medidas que se corresponden con el nivel de medición de las variables. La tabla 1E sintetiza estas medidas.

Tabla 1E. Medidas de tipicidad (tendencia central) y dispersión según nivel de medición

Nivel de medición	Tendencia central	Dispersión
Nominal	Moda o modo	Razón de no modalidad
Ordinal	Mediana	Desviación cuartil y rango intercuartil
Intervalar o de razón	Promedio o media	Desviación típica (estándar) y coeficiente de variabilidad

El **modo** o **moda** es igual a la categoría de la variable con mayor frecuencia. Así por ejemplo, en una sala de clase en la que hay veinte niños y diez niñas, la moda es niño y la **frecuencia modal** veinte.

La **mediana** es la categoría de la variable que divide a la distribución en dos partes iguales con relación al número de casos. Así por ejemplo, si las edades de los niños y niñas del ejemplo anterior estuvieran distribuidas de la manera que se muestra en la tabla 1F, la mediana será seis años; es decir, la mitad de los estudiantes tiene seis años o menos y la otra mitad tiene una edad mayor.

Tabla 1F. Niños y niñas por edad

Edad	Número de niños y niñas			
	Frecuencia		Frecuencia acumulada	
	Absoluta	Porcentual	Absoluta	Porcentual
5	1	3,3	1	3,3
6	14	46,7	15	50,0
7	8	26,7	23	76,7
8	6	20,0	29	96,7
9	1	3,3	30	100,0

El **promedio** o **media** es el resultado de la suma de todos los valores individuales divididos por el número total de casos. Los promedios pueden ser **simples** o **ponderados**. Por ejemplo, la **tasa neta de matrícula promedio** en educación primaria de los países sudamericanos en el año 2007 es el resultado de obtener el valor medio

de los datos nacionales tomados cada uno como una unidad (promedio simple), mientras que la **tasa para la subregión** como agregado será el resultado de tomar los mismos valores individuales pero considerando el peso relativo que tiene cada país en la población total (de los correspondientes grupos de edad) de la región. Así tenemos lo que se muestra en la tabla 1G.

Tabla 1G. Educación primaria: tasa neta de matrícula y población en edad de cursarla, y estudiantes matriculados; por países de América del Sur; 2007

País	Tasa neta de matrícula (%)	Población en edad de cursar	Estudiantes matriculados
Argentina	98,5	4.064.240	4.003.879
Bolivia	93,7	1.396.601	1.308.132
Brasil	92,6	13.888.737	12.858.847
Chile	94,4	1.589.262	1.499.925
Colombia	90,3	4.411.899	3.982.221
Ecuador	96,9	1.721.086	1.667.790
Guyana	94,7	99.104	93.828
Paraguay	92,4	848.490	784.320
Perú	96,8	3.541.003	3.427.524
Surinam	85,6	60.263	51.564
Uruguay	97,5	314.431	306.615
Venezuela	92,2	3.320.433	3.061.473
Total	93,7	35.255.549	33.046.118
Tasa neta promedio (región)	93,8		

Fuente: la primera columna corresponde a datos del año escolar terminado en 2007, tomados del Centro de Datos del UIS-Unesco; los valores en cursiva son del año más próximo disponible (Argentina: 2005; Guyana: 2008; Paraguay: 2006). La segunda columna ha sido tomada de la misma fuente, pero son proyecciones de la División de Población de las Naciones Unidas. Elaboración propia.

El promedio simple de los valores de la tasa neta de matrícula observada en estos doce países sería simplemente el resultado de promediar los valores de la primera columna. Esto da como resultado un valor de 93,8%.

El promedio ponderado se calcula de una de las siguientes dos maneras:

- (a) se multiplica el valor de la primera columna por la segunda; (b) se suma los resultados obtenidos; y (c) se divide dicho valor por la población total.

- (a) si no se cuenta con el dato de matrícula, se recalcula⁷ dicho valor para cada país convirtiendo las tasas de matrícula en una fracción de uno y multiplicando dicho valor por la segunda columna (el resultado se muestra en la tercera columna); (b) se suma los resultados obtenidos; y (c) se divide dicho valor por la población total.

Ambos procedimientos son equivalentes y arrojan el mismo resultado: una tasa subregional de 93,7%, distinta de la obtenida mediante el cálculo de la media simple. Si bien en este caso la diferencia entre la media simple y la media ponderada es pequeña, esto obedece a la poca variabilidad en los valores de las tasas observadas. Así, en situaciones en las que las tasas tengan mayor variabilidad, estas diferencias serán mayores y ambos valores pueden dar imágenes muy distintas de la situación, como se puede observar en el mismo ejemplo pero aplicado al caso de la educación secundaria (tabla 1H).

Tabla 1H. Educación secundaria: tasa neta de matrícula y población en edad de cursarla, y estudiantes matriculados; por países de América del Sur; 2007

País	Tasa neta de matrícula	Población en edad de cursar	Estudiantes matriculados
Argentina	<i>79,4</i>	4.100.713	3.254.055
Bolivia	69,9	1.285.297	898.825
Brasil	77,0	23.405.131	18.011.964
Chile	85,3	1.778.925	1.516.700
Colombia	70,3	5.256.800	3.696.868
Ecuador	59,2	1.639.988	971.313
Paraguay	<i>57,7</i>	809.692	467.596
Perú	75,9	2.931.012	2.224.031
Surinam	<i>64,6</i>	64.019	41.335
Uruguay	67,7	320.475	217.082
Venezuela	68,4	2.739.706	1.874.212
Total	74,8	44.331.758	33.173.981
Tasa neta promedio (región)	70,5		

Fuente: la primera columna corresponde a datos del año escolar terminado en 2007, tomados del Centro de Datos del UIS-Unesco; los valores en cursiva son del año más próximo disponible (Argentina: 2006; Paraguay: 2006; Surinam: 2005). La segunda columna ha sido tomada del mismo lugar, pero son proyecciones de la División de Población de las Naciones Unidas. Se excluyó a Guyana por no tener datos. Elaboración propia.

⁷ Se debe recalcular, ya que la tasa en cuestión fue calculada mediante el procedimiento inverso: dividir la matrícula por la población.

En este caso, la media simple arroja un valor de 70,5%, mientras que la media ponderada es 74,8%. La importante diferencia es explicada por el efecto de Brasil en el cálculo, ya que tiene un valor mayor a la media y su población en edad de cursar la educación secundaria representa un 54,3% del total para los países considerados.

Las tres medidas de tendencia central que hemos mencionado (modo o moda, mediana, promedio o media) sirven para la identificación de las situaciones típicas y, como ya se mencionó, el análisis requiere considerar también la dispersión existente.

Así, la medida de dispersión que acompaña a la moda es la **razón de no modalidad**. Esta medida da cuenta de la proporción de observaciones que no corresponden a la categoría modal. En términos formales, la razón de no modalidad se calcula de la siguiente manera:

$$RNM = \frac{N - F_{modal}}{N} * 100$$

Donde:

RNM: razón de no modalidad

N: número total de observaciones

F_{modal}: frecuencia modal

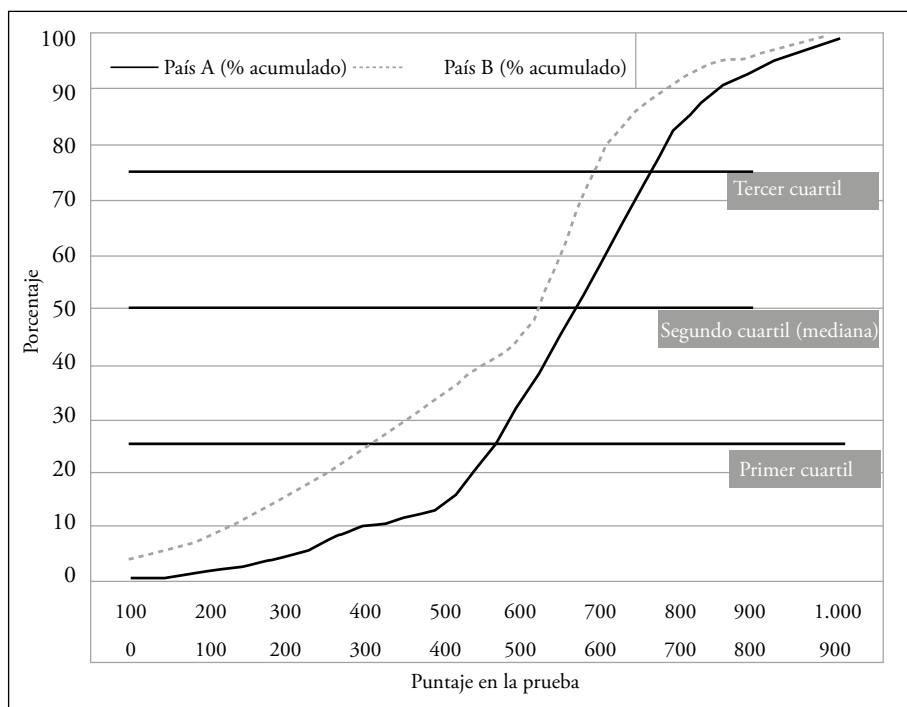
Así, en el ejemplo mencionado anteriormente donde se tiene 20 niños y 10 niñas en una sala de clases, la moda es niño (con una frecuencia modal de 20) y la razón de no modalidad se calcula como los casos que no se encuentran en la categoría modal (en este caso: 30 - 20 = 10) sobre el total de observaciones (30). Así, el valor calculado será 10 dividido por 30 y expresado como porcentaje: 33,3%. Este valor se interpreta como que un tercio (33,3%) de los casos no corresponden a la categoría modal. Cuando la variable es dicotómica, como sexo, la medida de la razón de no modalidad es simple, pero en el caso de una variable politómica (con varias categorías) calcularlo puede no resultar tan evidente.

En las variables de nivel ordinal, dado que sus categorías se presentan de modo ordenado (mayor a menor o menor a mayor) la forma de dar cuenta de la tendencia central tiene que ver con la acumulación de observaciones hasta determinado valor. Así, es posible dividir una distribución de frecuencias acumuladas en diversos grupos de igual tamaño o **cuantiles**. Típicamente, estos pueden ser **cuartiles** (cuatro grupos con 25% de la población cada uno); **quintiles** (cinco grupos con 20% de la población cada uno); **deciles** (diez grupos con 10% de la población cada uno); o **percentiles** (cien grupos con 1% de la población cada uno).

La mediana, al dividir a la población en dos grupos de igual tamaño (50% por encima y 50% por debajo de la mediana), es equivalente al cuartil 2, al decil 5 y al percentil 50.

Del mismo modo que la mediana muestra el valor central, la distancia existente entre los cuartiles primero y tercero muestra entre qué categorías de la variable se ubica el 50% central de los casos. Esta medida de dispersión es llamada **rango intercuartil** y complementa a la mediana. Los **gráficos de cajas** o las **ojivas de frecuencias acumuladas** son usualmente el mejor recurso para graficar esto, como en el ejemplo del gráfico 1B.

Gráfico 1B. Distribución de frecuencias porcentuales acumuladas de estudiantes de cuarto grado según puntajes en las pruebas de rendimiento académico de lenguaje, País A y País B, año x



Nota

Los valores hipotéticos para efectos de ilustración. Los puntajes de los estudiantes (en el eje de las abscisas) se han agrupado en intervalos cerrados al inicio y abiertos en el límite superior; es decir, el primer intervalo va de 0 al valor inmediatamente anterior a 100, el segundo de 100 al valor inmediatamente anterior a 200 y así sucesivamente. Se asume que ambos países aplicaron una prueba comparable cuyos resultados se expresan en una misma escala.

El gráfico 1B muestra que, si bien los valores medianos son relativamente próximos (sobre la escala de 1.000 puntos usada para dar puntajes a los estudiantes y presentada en el gráfico en diez intervalos de igual tamaño), las dispersiones son muy distintas. Así, en el País B se observa una mayor dispersión (la distancia entre

el Q1 y el Q3, o rango intercuartil, es mayor que en el País A) y esa dispersión es resultado de una mayor acumulación de casos en la parte inferior de la escala. En el caso de las variables de naturaleza numérica (de razón o intervalares), es posible calcular, como medida de tendencia central, la **media** o **promedio** y, como medida de dispersión, la **desviación estándar** o **desviación típica**. Como ya se anotó, la media es el resultado de agregar todos los valores individuales divididos por el número total de casos. Frente a esto, si se quiere conocer la dispersión, basta con mirar la distancia de cada caso individual respecto de dicho valor medio. Sin embargo, esto presenta un problema: dado que existen casos que se distancian del promedio tanto hacia arriba como hacia abajo y que el promedio busca encontrar el punto que minimice esas distancias, la suma algebraica de dichas distancias (es decir, considerando el signo) da un valor de cero. En virtud de esto, la medida de dispersión más usual da cuenta de la distancia típica, para lo cual se usa la raíz cuadrada de la sumatoria del cuadrado de las distancias individuales dividida por el número de observaciones. Es decir:

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{n}}$$

Donde:

- σ : desviación estándar o desviación típica
- x : cada valor individual observado
- μ : promedio
- n : número total de observaciones

Nótese que el valor de la desviación estándar estará expresado en las unidades de medida de la propia variable, por lo que, a efectos de comparar dos distribuciones diferentes (con variables expresadas en diferentes unidades) y determinar cuál de ellas presenta mayor grado de dispersión, es necesario transformar dichos valores en unidades comparables. Para esto, el valor de la desviación estándar es dividido por el valor del promedio, con lo que se obtiene un **coeficiente de variabilidad**. Este expresa el valor de la desviación estándar como una fracción del valor de la media (usualmente expresada como porcentaje). Es decir:

$$CV = \left(\frac{\sigma}{\mu} \right) * 100$$

Donde:

- CV : coeficiente de variabilidad
- σ : desviación estándar o desviación típica
- μ : promedio

Las medidas aquí presentadas dan cuenta de la agrupación y de la dispersión de los casos alrededor de una tendencia central. Esta agrupación puede ser mayor o menor, por lo que una distribución puede ser más concentrada o más dispersa

según sea el caso. Asimismo, independientemente de los niveles de concentración, una distribución puede tener mayor dispersión hacia uno de los lados del valor central (lo que solo se aplica a variables ordinales o superiores), lo que denota una situación de **asimetría**. Así, es posible tener distribuciones con diversas formas, tales como:

- **Distribuciones asimétricas**, donde los casos tienden a estar agrupados hacia un extremo y distribuidos hacia el otro a través de un conjunto amplio de categorías de la variable. Aquí, el sesgo de la variable, por tanto, puede ser positivo o negativo (dependiendo de hacia dónde se concentren o dispersen los valores) y es medido con un **coeficiente de sesgo**.
- **Distribuciones bimodales o multimodales**, donde existe más de una categoría de la variable en la que se concentran los valores mayores, de modo que se tiene más de una moda.

Por otra parte, las distribuciones pueden tener distintos grados de concentración. Una curva muy usada en estadística es la **distribución normal estándar**, también conocida como «campana de Gauss», en honor al científico Carl Friedrich Gauss. La curva normal es una curva con propiedades conocidas que están a la base de muchas operaciones con probabilidades. Más adelante (ver el gráfico 1C en la página 349) se incluye esta curva y su uso básico con relación a la estadística inferencial.

Ahora bien, dicha curva muestra un alto nivel de simetría y, en comparación a esta forma, una distribución cualquiera puede presentar una concentración central mayor, de modo tal que la gráfica aparecerá con una cúspide más marcada (a estas formas de distribución se les llama distribuciones **leptocúrticas**); de modo análogo, otras distribuciones tenderán a mostrar una cúspide menor o una forma más achatada (a estas se les llama distribuciones **platocúrticas**); finalmente, existen distribuciones con niveles de concentración similares a los de la curva normal (estas son conocidas como **mesocúrticas**). Esta propiedad de las distribuciones es medida con un coeficiente llamado *curtosis*, de donde les viene el nombre.

2.2 DESCRIPCIONES BASADAS EN MÁS DE UNA VARIABLE (MULTIVARIABLES)

La descripción de una sola variable a la vez no siempre resulta suficiente para brindar una imagen particularmente útil, en este caso, para la formulación de políticas. Así, por ejemplo, puede ser necesario no solo conocer la matrícula por nivel educativo sino también según los tipos de gestión de las instituciones educativas, su ubicación en área urbana o rural, etc.

Esto lleva a la necesidad de contar con herramientas de presentación y análisis de la información usando más de una variable a la vez.

- (i) **Análisis bivariable.** Lo primero que se puede hacer es presentar una variable **según** las propiedades de otra. Esto da lugar a tabulaciones cruzadas, o tablas de dos entradas, en las que las categorías de las variables consideradas dan cuenta de las columnas y filas de la tabla. Una tabla de dos variables es aquella que organiza los datos de una variable según otra. Así, la primera variable es colocada del mismo modo que en el caso de las distribuciones de frecuencia univariadas, mientras que la segunda es ubicada en la parte superior de la tabla formando las columnas de la misma.

Del mismo modo, los valores de una tabla así construida pueden permitir componer un gráfico, donde cada punto muestra los dos valores que cada observación tiene, uno para cada una de las variables observadas.

Ahora bien, la sola presentación gráfica puede sugerir la existencia de una asociación entre las variables, es decir, la existencia de un patrón de comportamiento conjunto que no es resultado del azar sino de alguna relación entre las variables. La asociación deriva en la existencia de **covariación** entre las variables: de cambios en los valores de una que van acompañados de cambios en los valores de la otra. En el caso de las variables ordinales o intervalares y de razón, estos cambios pueden, o no, darse en una misma dirección. Así, cuando incrementos en una variable van acompañados de incrementos en los valores de la otra, se habla de una asociación o relación **positiva** o **directa**; mientras que, cuando los incrementos en el valor de una variable van acompañados por disminuciones en el valor de la otra, se habla de una relación o asociación **negativa** o **inversa**.

A efectos de medir la asociación entre variables es posible usar un conjunto de coeficientes de asociación. Algunos de los coeficientes más usados se presentan en la tabla II.

Tabla II. Coeficientes de asociación en la medición de variables

Nivel de medición de la variable	Coficiente de asociación
Nominal	Q de Yule ϕ (phi) V de Cramer
Ordinal	γ (gamma) T (tau) de Kendall ρ (rho) de Spearman
Razón o intervalar	r^2

Los coeficientes de asociación arrojan valores entre 0 y 1, donde 0 indica la no existencia de asociación y 1 (-1 y 1 para el caso de las variables de nivel ordinal o superior; el signo permite denotar si la relación es directa o inversa) la existencia de una relación «perfecta» o de una intensidad absolutamente determinante.

Los principales *software* estadísticos permiten calcular estos y otros coeficientes. Sin embargo, el *software* no tiene la capacidad para discriminar qué coeficiente debe usarse según la naturaleza de la variable y, por lo mismo, es preciso tener particular cuidado en este respecto. Asimismo, para las variables ordinales, es preciso tener en cuenta que sus categorías (o la codificación usada) presenten un ordenamiento sistemático, ya que si una variable está ordenada en un sentido (de menor a mayor) y la otra en el otro (de mayor a menor) el signo del coeficiente resultará invertido.

Asimismo, en el caso de contar con variables de dos niveles de medición diferentes, el coeficiente a usar corresponde al nivel de medición menor, dada la relación que existe entre los diversos niveles de medición, que fuera explicada e ilustrada anteriormente.

A continuación se muestra la manera como se calculan los coeficientes mencionados.

A nivel nominal

- Q de Yule

Este coeficiente es usado para medir la asociación entre dos variables nominales dicotómicas, siempre y cuando no exista ninguna combinación de las categorías que carezca de observaciones (en la tabla 1J los valores *a*, *b*, *c* y *d* deben ser diferentes a cero).

La mejor manera de mostrar esto es con un ejemplo algorítmico como el de la tabla 1J.

Tabla 1J. Ejemplo de cálculo del coeficiente Q de Yule

		Variable 2	
		Categoría 1	Categoría 2
Variable 1	Categoría 1	<i>a</i>	<i>b</i>
	Categoría 2	<i>c</i>	<i>d</i>

El coeficiente se calcula de la siguiente manera:

$$Q = \frac{ad - bc}{ad + bc}$$

- ϕ (phi)

También es usado para medir la asociación entre dos variables nominales dicotómicas. Así, es similar al anterior, pero puede calcularse aunque exista alguna de las combinaciones de las categorías de las variables con valor de 0. El coeficiente se calcula de la siguiente manera:

$$\phi = \frac{ad - bc}{\sqrt{(a + c) * (b + d) * (a + b) * (c + d)}}$$

- V de Cramer

Se usa asimismo para medir la asociación entre dos variables nominales, pero estas no tienen que ser necesariamente dicotómicas, es decir, pueden tener cualquier número de categorías.

Este coeficiente se calcula de la siguiente manera:

$$V = \sqrt{\frac{\chi^2}{\min[(f - 1)(c - 1)]]}$$

Donde χ^2 es el valor de la función correspondiente (chi cuadrado, representado por χ^2) calculado para la tabla observada⁸ y $\min [(f - 1) (c - 1)]$ es el valor mínimo que resulte de comparar el número de filas menos 1 ($f - 1$) y el número de columnas menos 1 ($c - 1$) en la tabla de dos entradas que muestra la distribución de los casos entre las categorías de ambas variables.

A nivel ordinal

Dado que las variables de nivel ordinal presentan una gradación entre sus categorías, el cálculo de varios de los coeficientes requiere que las tablas de doble entrada que se compongan muestren las categorías de ambas variables ordenadas en una misma dirección. En el caso de operar con casos singulares no agregados en tablas, es decir, cuando se usa algún *software* estadístico, el ordenamiento resulta de la codificación de las categorías de las variables.

⁸ El valor de χ^2 es calculado comparando los valores observados en la tabla de doble entrada que muestra los valores observados para cada combinación de categorías de ambas variables y una distribución teórica de no relación entre variables. Se calcula mediante la siguiente fórmula:

$$\chi^2 = \sum \frac{(f_o - f_t)^2}{f_t}$$

donde f_o es igual a la frecuencia observada y f_t a la frecuencia teórica.

Así, si en las filas las categorías están ordenadas de menor a mayor (la fila superior para la categoría menor y la inferior para la superior), las columnas deben estar ordenadas de la misma manera (la columna más a la izquierda con la categoría menor y la columna más a la derecha con la categoría mayor).

La tabla 1K muestra dicho ordenamiento y será usada como referencia para el cálculo de los coeficientes.

Tabla 1K. Tabla de doble entrada con variables de nivel ordinal

		Variable 2		
		Categoría 1 (-)	Categoría 2	Categoría 3 (+)
Variable 1	Categoría 1 (-)	<i>a</i>	<i>b</i>	<i>c</i>
	Categoría 2	<i>d</i>	<i>e</i>	<i>f</i>
	Categoría 3 (+)	<i>g</i>	<i>h</i>	<i>i</i>

En una tabla como la 1K, se conoce como **frecuencias de acuerdo** (o **pares concordantes**) al producto de cada casillero por la suma de los casilleros que se encuentran por debajo y a la derecha de este. Así, por ejemplo, en el caso de la celda *a*, se trata de $a * (e + f + h + i)$ y en el caso de la celda *b* se trata de $b * (f + i)$. Del mismo modo, se conoce como **frecuencias de inversión** (o **pares discordantes**) al producto de cada casillero por la suma de los que se encuentren por debajo y a la izquierda de este. Así, por ejemplo, en el caso de la celda *c*, se trata de $c * (d + e + g + h)$ y en el caso de la celda *f* de $f * (g + h)$.

- **γ (gamma) de Goodman y Kruskal**

Se calcula de la siguiente manera:

$$\gamma = \frac{fa - fi}{fa + fi}$$

donde fa es la sumatoria de las frecuencias de acuerdo para todas las celdas y fi , la sumatoria de las frecuencias de inversión para todas las celdas.

- **T (tau) de Kendall**

Este coeficiente tiene tres versiones: T_a es la forma general; T_b la forma específica para tablas cuadradas (igual número de categorías en columnas y filas); y T_c la forma específica para tablas rectangulares (diferente número de categorías en columnas y filas).

El coeficiente se calcula de las siguientes maneras:

Donde:

$f\hat{a}$: sumatoria de las frecuencias de acuerdo para todas las celdas

$f\hat{i}$: sumatoria de las frecuencias de inversión para todas las celdas

$n_0 = n(n-1)/2$

$n_1 = \sum t_i(t_i - 1)/2$

$n_2 = \sum u_i(u_i - 1)/2$

t_i : número de pares empatados para la primera variable

u_i : número de pares empatados para la segunda variable

m : menor número de pares no empatados

$$T_a = \frac{f\hat{a} - f\hat{i}}{\frac{1}{2}n*(n-1)}$$

$$T_b = \frac{f\hat{a} - f\hat{i}}{\sqrt{(n_o - n_1)*(n_o - n_2)}}$$

$$T_c = \frac{2m(f\hat{a} - f\hat{i})}{n^2(m-1)}$$

- **ρ (rho) de Spearman**

Se utiliza para comparar dos conjuntos de datos ordenados en *rankings* de forma tal que en estos se comparan las posiciones relativas de cada observación. Se calcula de la siguiente manera:

$$\rho = 1 - \left(\frac{6 * (\sum di^2)}{N^3 - N} \right)$$

donde N es el número de observaciones y di la diferencia posicional de cada observación en los dos *rankings* que se comparan.

A nivel intervalar o de razón

- r^2

En el caso de las variables de razón o intervalares es posible calcular una función que representa la relación lineal entre las variables y medir el ajuste entre dicha función y los valores observados.

A esa función se le conoce como la **recta de mínimos cuadrados**, ya que la sumatoria del cuadrado de todas las diferencias entre los valores observados y los valores de la función es la mínima posible, de manera tal que la recta resultante de la función es la que presenta el **mejor ajuste** posible.

La función tiene la forma $y = a + bx$, donde el valor de la variable y es igual a una constante que representa el valor de y cuando x es cero, a la que se suma el valor de x multiplicado por la inclinación o pendiente de la recta (b).

A mayor ajuste entre la función calculada y los valores observados, mayor el valor del coeficiente r^2 que da cuenta de la variación en una de las variables que es explicada por la otra.

El valor del coeficiente se calcula de la siguiente manera:

$$r^2 = \left(\frac{N \sum (x * y) - (\sum x * \sum y)}{\sqrt{N \sum x^2 - (\sum x)^2} * \sqrt{N \sum y^2 - (\sum y)^2}} \right)^2$$

La relación entre dos variables de nivel intervalar o de razón puede, sin embargo, seguir también un patrón no lineal. Por lo tanto, en estos casos la relación puede ser mejor representada con una función no lineal (por ejemplo logarítmica, exponencial o polinómica), para la cual también es posible calcular el nivel de ajuste de los valores observados a los valores de la función calculada mostrando la intensidad de la relación entre las dos variables con un coeficiente r^2 .⁹

Es importante anotar que la existencia de asociación entre variables no debe entenderse inmediatamente como que hay una relación de **causalidad** entre las mismas. A efectos de establecer una relación de este tipo, es necesario, en primer lugar, contar con un marco conceptual explicativo que permita dotar de sentido causal a la asociación identificada. Por otra parte, la causalidad supone entender que una de las variables es **dependiente** de la otra, es decir, el establecimiento de una relación en la que cambios en la variable **independiente** se traducen en cambios en la variable **dependiente**. Esto se traduce en el hecho de representar matemáticamente los valores de la variable dependiente como una función de la variable independiente.

Asimismo, la constatación empírica de valores que sugieren la existencia de una asociación entre variables debe ser verificada, ya que eventualmente la covariación entre dos variables es en realidad el resultado de la interacción de ambas con una tercera, con lo que la relación identificada inicialmente sería espuria. Esto lleva a la introducción de mayores variables en el análisis y al uso de la técnica conocida como **control estadístico**, es decir, a determinar el grado de covariación entre dos variables **controlando** el efecto de una tercera (o de un conjunto mayor de variables).

Un ejemplo típico de esto en el terreno educativo está dado por las diferencias de puntajes en pruebas de medición de logro académico entre centros

⁹ Los programas estadísticos y las hojas de cálculo permiten calcular y graficar este tipo de funciones (a las que en algunos casos definen como **líneas de tendencia**) y los coeficientes correspondientes.

educativos estatales y no estatales. En efecto, la observación de los resultados según el tipo de gestión de los programas suele mostrar un patrón: puntajes más altos en los centros educativos no estatales. Esto podría sugerir que los centros educativos no estatales brindan un mejor servicio que los estatales, asegurando que sus estudiantes logren mejores resultados. Sin embargo, si se introduce el nivel socioeconómico de los estudiantes como una tercera variable de análisis, dichas diferencias cambian y, en algunos casos, desaparecen o se invierten. Así, es posible que el puntaje obtenido por los estudiantes esté, en realidad, más asociado a su condición socioeconómica que al tipo de gestión del centro educativo al que asisten.

(ii) **Análisis multivariable.** Como ya se mencionó, un primer elemento a considerar está dado por la introducción de variables de control en el análisis bivariable. En adición a esto, existe un conjunto amplio de técnicas vinculadas al manejo simultáneo de un conjunto también amplio de variables.

Estas técnicas atienden diversos objetivos, entre ellos: determinar asociaciones múltiples, clasificar o reducir el número de variables, etc.

Un primer caso de análisis multivariado está dado por situaciones en las que es posible identificar covariaciones entre más de dos variables. Es decir, en el caso de variables intervalares o de razón, es posible construir una **matriz de correlación múltiple** en la que se identifiquen las relaciones entre las diferentes variables dando una imagen más compleja de la covariación.

Un segundo caso de análisis multivariable está vinculado a la identificación de patrones en la información que devienen en la existencia de conglomerados o *clusters*. En efecto, en una población dada, se puede identificar que las relaciones entre las variables tienen comportamientos específicos para diferentes grupos poblacionales; así se puede caracterizar **agrupamientos de los casos** (o conglomerados de observaciones) con relación a estos patrones. Del mismo modo, se puede identificar **conglomerados de variables** que se comportan de manera tal que revelan pautas comunes para cada uno de esos grupos. En ambos casos (conglomerados de observaciones y conglomerados de variables), la intención es identificar una pauta latente que deriva en agrupamientos (de casos o de variables).

Este tipo de análisis es muy pertinente con relación a aspectos como los siguientes:

- Al estimar la omisión: la identificación de *clusters* de casos permite observar, para cada agrupamiento, ciertas características propias para ajustar mejor el modelo a usar para imputar valores.

- Al observarse significativos agrupamientos en las variables: podría simplificarse los procesos de acopio de datos o de análisis, considerando un número menor de variables (aquellas que mejor representen la dinámica de los conglomerados).
- Al diseñar estrategias de muestreo: dado que los conglomerados de casos son relativamente homogéneos en su interior respecto del comportamiento de las variables usadas para identificar el agrupamiento, por lo tanto, las estrategias de muestreo pueden tomar esta información a efectos de optimizar la asignación de casos a cada grupo.

Lo anterior también sugiere otra forma de análisis, que consiste en identificar en un conjunto grande de variables si en realidad hay grupos de ellas que podrían estar midiendo una realidad común subyacente. Es decir, existe la posibilidad de que las variables en un conjunto estén midiendo manifestaciones singulares de un (o varios) factor(es) comunes latentes. La identificación de estos elementos comunes subyacentes es el propósito de los diversos modelos de análisis de variables latentes.

3. ELEMENTOS DE ESTADÍSTICA INFERENCIAL

Así como es posible usar herramientas estadísticas para describir poblaciones, hay un área de la estadística que trata con asuntos derivados de operar con información muestral (lo que es en muchas ocasiones preferible a usar información de todas las unidades observables –es decir, información censal–). Si bien es cierto que las muestras pueden, o no, representar de modo suficientemente bueno la realidad general que se desea observar, siempre lo harán con un margen de error que hay que tomar en consideración.

Si, por ejemplo, a partir de información censal se cuenta con evidencia acerca de que el promedio de años de educación logrados por la población de quince años y más del País A es de 10,3 años, mientras que en el País B es de 7,5, entonces resulta sencillo afirmar que dicho promedio es menor en el segundo caso. Sin embargo, si esa misma información hubiese sido generada mediante una encuesta por muestreo, hacer la misma afirmación **con rigurosidad** ya no sería tan sencillo.

Existe un fenómeno llamado **variabilidad de muestreo** que alude al hecho de que muestras probabilísticas distintas de una misma población arrojan resultados que no son idénticos así estén diseñadas de la misma manera y el levantamiento de información haya seguido los mismos procedimientos. Por esta razón, resulta equívoco asumir, por muy buena que sea la información muestral, que los valores poblacionales son idénticos a los observados en muestras. De hecho, en los casos mencionados en el ejemplo del párrafo precedente, si se tratase de información muestral, lo que es propio asumir es que el promedio de años de educación logrados por la población del País A es $10,3 \pm \alpha$ años

y en el País B $7,5 \pm \beta$ años, donde α y β representan los intervalos en los que es posible estimar que, dados ciertos parámetros y con cierto nivel de confianza, se encuentran los valores poblacionales correspondientes. Si dichos márgenes llevaran a que los intervalos calculados para ambas poblaciones tengan un área de superposición, resultaría poco riguroso asumir que el valor en un caso es superior al del otro.

Es decir, una vez que se trabaja con muestras se introduce en la operación factores nuevos a tomar en cuenta: **márgenes de error** e **intervalos de confianza**, así como la llamada **significancia estadística**.

El presente acápite habrá de reseñar los principales procedimientos de la estadística inferencial que requieren ser tomados en consideración por los sistemas de información educativa: muestreo, estimación de parámetros, prueba de hipótesis y prueba de significancia¹⁰. Sin embargo, cabe notar que los principales estadísticos permiten calcular estos procedimientos, por lo que el propósito de este acápite es explicar su lógica y servir como referencia que ayude a manejar lo que el *software* estadístico produce.

3.1 Muestreo

Una muestra, por definición, es una fracción de una población dada. Para graficar esta situación, usualmente las muestras son representadas por la letra n , mientras que las poblaciones lo son con la letra N .

Ahora bien, la simplísima definición anterior no implica que todas las muestras sean iguales, ya que se puede obtener observaciones de una fracción de la población de muy distintas maneras. Por lo mismo, una primera distinción a hacer entre muestras es la que corresponde a su potencial para formular inferencias estadísticas.

Existen muestras que permiten estimar con procedimientos estadísticos los valores poblacionales y otras que no lo permiten. Las últimas pueden ser de mucha utilidad en estudios exploratorios y en pruebas de campo (para validar instrumentos y procedimientos, por ejemplo) y son usualmente obtenidas a partir de juicios expertos que permiten definir cuáles son las observaciones que mejor se ajustan al propósito de la investigación.

Por ejemplo, si se está planificando una encuesta nacional y se necesita validar instrumentos y procedimientos, se puede definir que interesa hacerlo en zonas urbanas y rurales, contando con población diferente según lengua materna (con un número similar de observaciones en cada caso independientemente de cómo se distribuya la población entre estos agrupamientos a nivel nacional) y que se desea priorizar zonas de más fácil

¹⁰ Nótese que aquí, como en todo este anexo, se brinda solo una reseña destinada a favorecer un mejor entendimiento de estos aspectos por parte de personas que carecen de formación profesional específica en el terreno de la estadística. Así, no se pretende ofrecer una guía práctica para operar con estas técnicas. El diseño muestral, por ejemplo, es una tarea muy delicada que requiere conocimientos especializados y un significativo nivel de experiencia para ser ejecutada de modo robusto.

acceso para reducir costos u otros criterios a considerar. La información de esta validación es muy importante para los fines propuestos, pero **no** permite hacer inferencias estadísticas, es decir, estimar los valores poblacionales de las variables de estudio a partir de la información muestral¹¹.

Esto quiere decir que, junto a este tipo de muestras, existen otras que sí permiten hacer inferencias estadísticas. Lo que distingue a unas de otras en términos de su potencial para la inferencia se deriva de ciertos elementos conceptuales básicos de la estadística inferencial que es preciso explicar para que los nombres de estos dos tipos de muestreo sean claros.

Estos elementos conceptuales se derivan de la teoría de las probabilidades, la misma que permite afirmar lo siguiente: si se toma un conjunto de elementos al azar de una población dada para estudiar una variable particular, la probabilidad de obtener un conjunto con características x (definidas por la variable de estudio) tenderá a ser igual a la proporción de conjuntos de elementos con dichas características en la población total.

Esto quedará más claro si se ilustra de la siguiente manera: si se quiere medir la variable sexo en una población en la que la proporción de hombres y mujeres es igual (50% de cada grupo), la probabilidad de seleccionar una mujer (o un hombre) si se extrae un caso aleatoriamente, es de 50%. Si, en cambio, se selecciona dos casos: se presentan las siguientes situaciones posibles: (i) una mujer y un hombre; (ii) dos mujeres; (iii) dos hombres. Dado que la selección es aleatoria, existen muchas muestras posibles que cuenten con dos observaciones y, por lo mismo, se puede obtener cualquiera de estos tres resultados si solo toma dos casos. Sin embargo, en la medida en que el tamaño de la muestra se incrementa a tres, cuatro, cinco, ... cien, ... doscientos casos, las observaciones en la muestra (siempre y cuando sean aleatorias) tenderán a tener una distribución similar a la de la población, ya que la probabilidad de obtener muestras con 50% de mujeres y 50% de hombres se incrementará. **A mayor número de observaciones tomadas al azar, mayor probabilidad de acercarse más al valor poblacional.**

Así, el estudio de las probabilidades permite concluir que muestras aleatorias de un tamaño dado tienen una probabilidad conocida de presentar la distribución de una característica dada, con cierta similitud a lo que sucede en la población.

Así, el muestreo que permite hacer inferencias estadísticas es conocido como **muestreo probabilístico** y tiene como elemento central de su diseño la **aleatoriedad**. Mientras tanto, el otro tipo de muestreo (el que no permite hacer inferencias estadísticas) es

¹¹ Lamentablemente he visto en más de una ocasión a «expertos» interesados en usar información generada de esta manera para hacer afirmaciones sobre la población. Esto puede obedecer tanto a falta de conocimientos como a falta de responsabilidad y ética profesional. Del mismo modo, he conocido gente que está dispuesta a realizar estudios de cierta complejidad **sin** hacer el menor esfuerzo por validar los instrumentos y procedimientos mediante muestras no destinadas a la inferencia, como las que aquí se mencionan.

conocido como **no probabilístico** y tiene como característica central de su diseño la no aleatoriedad.

3.1.1 El tamaño de la muestra

Debe tomarse en cuenta que el tamaño de la muestra es, por tanto, el primer elemento a considerar, ya que permite determinar una probabilidad de que dicha muestra represente las características a estudiar de la población con cierto nivel de certeza. Esto es siempre una probabilidad, ya que al tomarse las observaciones de modo aleatorio siempre es posible que, por puro azar, se tome casos atípicos.

En el ejemplo anterior, también siempre es posible que, como resultado del azar, se tome una muestra compuesta únicamente por mujeres o únicamente por hombres; sin embargo, la probabilidad de que esto sea así disminuye en la medida en que el tamaño de la muestra aumenta. Esto opera exactamente de manera análoga a arrojar un dado: si se arroja un dado una vez se puede obtener cualquier valor, pero si se lo arroja, por ejemplo, 60.000 veces, el resultado será muy próximo a 10.000 veces cada uno de los valores del dado (a menos que esté alterado).

Así, el **tamaño de una muestra probabilística** se calcula considerando cuatro criterios:

- El **nivel de confianza** que se requiere para las estimaciones. Es decir, la probabilidad de que la muestra sea efectivamente una que represente a la población en términos de la variable de estudio. La intención es contar con resultados que uno pueda considerar una buena aproximación a los valores poblacionales con una probabilidad dada; por ejemplo, con un 95% de confianza. Esto significa que existe un 95% de probabilidades de que la muestra obtenida dé cuenta de los valores típicos de la población, es decir, hay un 5% de probabilidad de que no sea así.

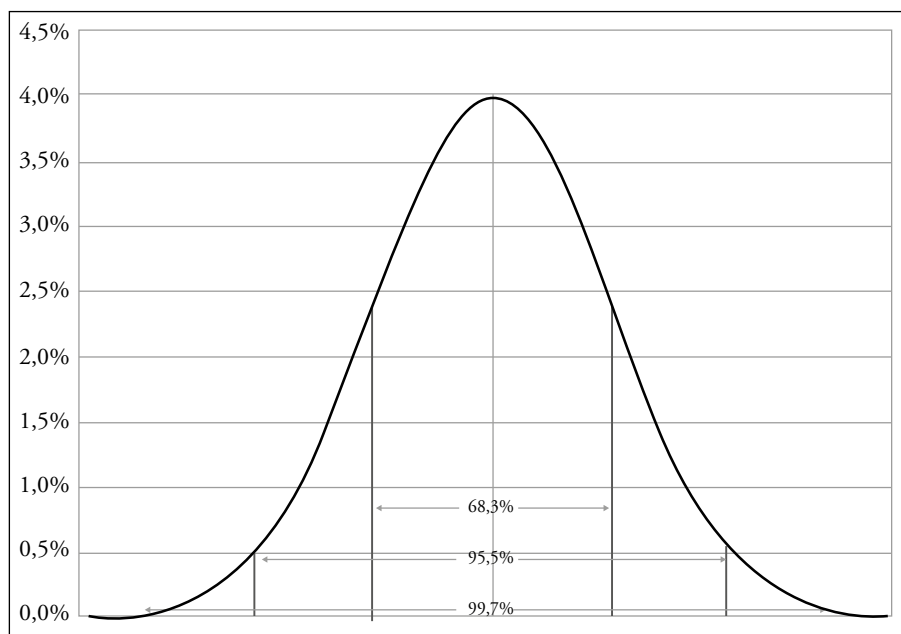
En general, se tiende a asumir que para poblaciones de un tamaño suficiente (más de treinta observaciones) las características de estudio se distribuyen de modo «normal», entendiéndose por esto una distribución que sigue un patrón conocido y representado por una curva normal o de Gauss que, al ser una función conocida, permite contar con probabilidades conocidas¹².

La curva normal se expresa como una distribución con media de 0 y desviación estándar de 1 y en ella los puntajes «normalizados» de una variable, es decir, expresados en términos de dicha curva teórica, son conocidos como puntajes z . Así, un valor z de 1,5, significa que dicho valor es mayor a la media en 1,5 veces la desviación estándar.

¹² Nótese también que existen otras distribuciones de probabilidades especificadas, así como el hecho de que algunas variables no se distribuyen normalmente.

De la misma manera, un rango de valores que se encuentra, por ejemplo, entre $-1,96z$ y $+1,96z$, (o sea, valores de 1,96 desviaciones estándar alrededor de la media) tiene una probabilidad de ocurrencia equivalente al área bajo la curva normal entre dichos valores, el mismo que es igual al 95% (ver el gráfico 1C)¹³.

Gráfico 1C. Áreas bajo la curva normal



El gráfico 1C ilustra los valores de las áreas bajo la curva para rangos que van del promedio a ± 1 , 2 y 3 desviaciones estándar. Estos valores indican que, para una variable cualquiera para la que se puede asumir que se distribuye de modo **normal** en una población dada, encontrar un valor entre: ± 1 desviación estándar alrededor de la media tiene una probabilidad de 68,3%; ± 2 desviaciones estándar alrededor de la media tiene una probabilidad de 95,5%; y ± 3 desviaciones estándar alrededor de la media tiene una probabilidad de 99,7%.

- La **heterogeneidad** de la población. Este es un factor muy importante, ya que si una población es muy homogénea respecto de la característica a estudiar, una muestra menor puede ser suficientemente buena para los propósitos que se tenga; mientras que si la población es heterogénea se requiere una muestra mayor, que justamente tenga la capacidad de representar dicha diversidad.

¹³ Las tablas con los puntajes z y las áreas bajo la curva normal suelen estar presentes en la mayor parte de libros de estadística inferencial. Asimismo, diversos productos de *software* estadístico, así como las funciones estadísticas de las hojas de cálculo electrónicas, permiten hacer estos cálculos.

- El **margen de error** con el que se hará las estimaciones de los valores poblacionales. Es decir, en qué medida esperamos que los valores obtenidos a nivel de la muestra (a los que se conoce como **estadísticos muestrales**) se distancien como máximo de los valores poblacionales a estimar (los que son conocidos como **parámetros poblacionales**). Por ejemplo, si se está estimando el promedio de edad de una población escolar en la muestra, cuán distinto podemos permitir que el valor obtenido en la muestra sea del valor a estimar para la población.
- El **tamaño de la población**. Como se desprende de los fundamentos básicos de las probabilidades, este no tiene una relación lineal con el tamaño de la muestra. De hecho, a partir de tamaños de población suficientemente grandes (100.000 observaciones) el tamaño de la población tiene una influencia marginal, por lo que deja de incluirse en el cálculo.

El tamaño de una muestra se suele calcular de la siguiente manera:

$$n = \frac{z^2 * p * (1 - p)}{E^2}$$

o

$$n = \frac{z^2 * \sigma^2}{E^2}$$

Donde:

n : tamaño de la muestra

z : nivel de confianza

p : heterogeneidad de la población

σ : desviación estándar poblacional o un estimador de esta

E : error muestral esperado.

Es importante destacar que la discusión anterior deja en claro que los tamaños de muestras probabilísticas no solo dependen de la magnitud de población y que su relación con esta no es lineal, sino también que no es pertinente usar una proporción determinada, del tipo x por ciento de la población, como manera de fijar el tamaño de tal muestra.

Sin embargo, el tamaño de la muestra no es el único elemento que importa. Para empezar, el impacto de incrementos en su tamaño es mayor mientras menor sea el tamaño de la muestra, aunque, en cierto punto –computable dadas las características de la población– dichos impactos empiezan a perder intensidad, de modo que las ganancias en precisión explicadas por cada caso adicional tienden a ser cada vez menores. Es decir, por ejemplo, para una encuesta nacional que desea conocer una variable como el sexo entre la población escolar, tomar cien casos daría determinados niveles de certeza que son muy inferiores a los que se podría obtener con trescientos casos más. Sin embargo, lo que se ganaría por tener setecientos casos en vez de cuatrocientos es mucho menos de lo que se gana pasando de cien a cuatrocientos (incluso si el incremento en el número de observaciones es similar).

Las curvas de probabilidades (como la ya mencionada curva normal) que se usan para definir el comportamiento de una variable en una población dada permiten conocer el impacto del tamaño de las muestras en la potencia de las estimaciones y, por lo mismo,

definir tamaños mínimos de muestras para niveles de precisión dados, así como el potencial impacto en la precisión de los resultados a partir de incrementos en dichos tamaños.

3.1.2 Formas de obtener una muestra

La forma como se obtienen las observaciones también tiene impactos en la confiabilidad de las estimaciones. Por lo tanto, el tamaño de la muestra es una condición necesaria pero no suficiente para hacer buenas estimaciones.

Ahora bien, en el terreno del muestreo probabilístico hay diversas formas de obtener muestras. La más simple es la conocida como **muestreo simple al azar** (SRS, por sus siglas en inglés). El SRS es la base sobre la cual una parte muy significativa del cuerpo conceptual del muestreo ha sido desarrollada, ya que representa la forma aleatoria *par excellence*, aunque no es el tipo de muestreo más usado.

Supongamos que se desea obtener una muestra de estudiantes de instituciones educativas de una zona geográfica o jurisdicción administrativa determinada. El procedimiento más simple (en términos conceptuales pero no necesariamente operacionales) consistiría en contar con un listado completo de dichos estudiantes y luego, por sorteo, seleccionar a un número de estos que se corresponda con el tamaño de muestra deseado.

El listado de los estudiantes, en este caso, constituye el **marco muestral** sin el cual no sería posible desarrollar el muestreo. Este marco permite determinar el tamaño de la población y la probabilidad de selección de cada caso en la muestra y esto es fundamental, ya que la probabilidad de selección requerida debe ser igual (en el caso del SRS) en cada caso y conocida. Si la probabilidad de selección no es conocida, no hay forma de controlar el sesgo que esto puede implicar. Por ejemplo, si el listado contiene duplicaciones (estudiantes que se encuentran matriculados en más de un programa educativo), esto llevará a que dichas personas tengan doble probabilidad de selección, lo que puede ser un problema mayor dependiendo de los objetivos de la muestra.

Un segundo tipo de muestreo, basado en el anterior, es el llamado **muestreo sistemático**. En este caso el procedimiento de selección varía, ya que en vez de sortear cada caso a la vez, solo se realiza un sorteo de partida y a partir de ahí se seleccionan casos según su posición en el listado y el tamaño de muestra a obtener. Supongamos que se trata de una población de 1.854 estudiantes y se ha definido que se requiere una muestra de 125 estudiantes. Este tamaño de muestra significa que se tomará a un estudiante de cada 14,8 en la población. Esto es lo que se conoce como **fracción de muestreo**. Así, el muestreo sistemático procede de la siguiente manera: (i) se computa un número aleatorio entre 1 y 14,8, este número define el primer caso del listado que debe seleccionarse para la muestra; (ii) a partir de ese caso (supongamos que el número obtenido fue 5,2 y por lo tanto se tomó al quinto elemento de la lista), se toma al que está 14,8 lugares más adelante ($14,8 + 5,2$), es decir, al que ocupa el puesto 20 en la lista y así sucesivamente. Debe tomarse en cuenta que si el listado se encuentra ordenado de determinada manera, esto puede tener un efecto sobre la probabilidad de selección de los casos.

Un tercer tipo es el **muestreo estratificado**. En este caso, las observaciones en la muestra son calculadas tomando en cuenta algún criterio de desagregación de la población que se considera relevante para propósitos analíticos. Por ejemplo, la muestra puede verse estratificada por jurisdicciones (departamentos o provincias), por áreas geográficas (urbanas y rurales), etc. Si se desea hacer buenas estimaciones para cada uno de los estratos, esto es prácticamente equivalente a tomar una muestra de cada uno. Estas muestras pueden estratificar o repartir los casos entre los diversos estratos de modo proporcional o no proporcional a su tamaño. Por ejemplo, si los estratos son urbano y rural y la población bajo estudio se distribuye según el patrón 85/15, la cantidad de casos en la muestra que corresponde a áreas rurales podría ser igual a 15% del total (en el caso de un muestreo proporcional), pero también podría ser mayor si por alguna razón se quiere sobrerrepresentar a esta población en la muestra a efectos de contar con mayor precisión para las estimaciones correspondientes a este subgrupo. Esta sobrerrepresentación no tiene por qué traducirse en problemas a la hora de estimar la información poblacional, ya que se usa ponderadores que corrigen esa sobrerrepresentación. Cuando los casos no tienen igual probabilidad de ser seleccionados (como en el caso de que se sobrerrepresente a un grupo), el uso de ponderadores es esencial.

Existen **otras formas de muestreo** que no es necesario reseñar acá; sin embargo, cabe detenerse por un momento en procedimientos que son muy usuales y deben ser tratados por especialistas en muestreo que permitan ayudar, primero, a diseñar formas adecuadas de muestreo y, luego, a estimar los márgenes de error de las estimaciones a hacerse. Esta discusión es importante porque usualmente el muestreo no sigue las pautas teóricas del SRS ya que, por ejemplo, no se cuenta con marcos muestrales equivalentes a los esperados por la teoría (por ejemplo, puede ser que no se disponga de un listado de estudiantes, pero sí de instituciones y/o programas educativos). Esto lleva a que el muestreo muy usualmente sea desarrollado por etapas.

En efecto, existen procedimientos usuales de muestreo de individuos que se aproximan a estos a través en diversas fases, por ejemplo:

- Uno de tales procedimientos consiste en, primero, definir un conjunto de unidades geográficas mayores (departamentos o provincias) y tomar algunas de estas; luego, dentro de cada unidad geográfica definir segmentos (usualmente los construidos para efectos censales) y seleccionar un grupo de estos; posteriormente, dentro de cada segmento, seleccionar un número de viviendas; dentro de estas, un hogar; y, finalmente, dentro de cada hogar, un número dado de individuos.
- En otro procedimiento, para el caso de estudiantes, primero se define un conjunto de unidades geográficas mayores (departamentos o provincias) y se toma algunas de estas; luego, dentro de cada unidad geográfica se toma unidades menores (por ejemplo, correspondientes a los ámbitos de acción de las autoridades educativas locales) y

se selecciona un grupo de estas; posteriormente, dentro de cada unidad menor se selecciona un número de grupos (secciones) y, finalmente, un grupo de estudiantes.

Es importante considerar que cada una de las etapas del muestreo influye en los atributos finales de la información (ya que todos los pasos del proceso tienen un impacto sobre la probabilidad de selección de cada caso) y, por lo mismo, es preciso que el procedimiento a seguir sea siempre discutido con y refrendado por un experto en muestreo. Es decir, cada etapa introduce fuentes de error propias y lo importante es minimizar y conocer el tamaño de esos errores, de modo que las estimaciones que se hagan se vean adecuadamente informadas por el proceso de diseño muestral. Nótese también que los errores de muestreo tienen la virtud de ser objeto de medición, mientras que otro tipo de errores (no vinculados al muestreo) pueden pasar inadvertidos y afectar de una manera no conocida los resultados que se obtengan.

3.1.3 El cálculo de los errores de estimación

Es importante precisar que la forma como se calcula los errores de estimación varía según se trate de muestreo simple o complejo (polietápico, por ejemplo) y esto es algo que se debe manejar con cautela, pues los *software* estadísticos suelen asumir por defecto que se está trabajando con muestras simples aleatorias, de forma que, si no se indica al *software* que este no es el caso, la estimación de los errores resultará equívoca. Cada *software* contiene formas propias de introducir información sobre el diseño muestral que requieren información específica al respecto así como instrucciones concretas diferentes de las que se usan por defecto.

Del mismo modo, cuando se combina información de muestras desarrolladas con diseños muestrales diferentes (por ejemplo, de estudios internacionales) no resulta posible ingresar el diseño, ya que no hay uno solo. En estos casos se suele proceder de modo distinto, mediante técnicas de remuestreo que usan la información muestral para simular datos de varias muestras; con dicha simulación, se recomputa los ponderadores a efectos de mejorar la estimación de los errores. Por esta razón, por ejemplo, en la base de datos de PISA se encuentra un ponderador original además de ochenta «ponderadores replicados» que son resultado de una técnica de remuestreo. En este caso, es preciso indicar al *software* que tome en cuenta dichos ponderadores, ya que, de otra forma, el cálculo de los errores de estimación será equívoco.

El diseño de una encuesta por muestreo es una tarea muy delicada que, además, debe verse acompañada por controles operacionales que minimicen los errores no atribuibles al muestreo. Para esto, debe tratarse con mucho cuidado aspectos como:

- la selección del personal a cargo de las labores de campo, incluyendo la supervisión;
- el diseño de los instrumentos de generación de información;

- las secuencias de aplicación cuando se utiliza distintos instrumentos;
- el diseño operacional de rotación de instrumentos cuando se utiliza bloques a rotar; y
- el plan operativo de las actividades de campo.

Aunque parezca extremadamente complejo, la generación de información mediante encuestas por muestreo es muchas veces preferible a los eventos censales, dado que el menor tamaño de observaciones permite un mayor control de los factores que potencialmente afectan negativamente el proceso, así como combinar baterías de instrumentos y/o administrar instrumentos más extensos que los posibles en eventos censales.

3.2 Estimación de parámetros (medias y proporciones)

Una vez que se ha obtenido los valores observados a nivel muestral (**estadísticos**), es preciso seguir un conjunto de procedimientos que permitan estimar los valores poblacionales (**parámetros**).

En otras palabras, asumir que el valor poblacional es idéntico al valor muestral, aunque parezca más simple y resulte más fácil de comunicar, no es un procedimiento riguroso, ya que no muestra de modo transparente que las estimaciones se hacen siempre dentro de determinados límites de confiabilidad y error.

El muestreo probabilístico parte de tomar en cuenta que existe una **variabilidad de muestreo**, lo cual es equivalente a decir que en una muestra cualquiera compuesta por casos seleccionados al azar ese mismo azar hubiese podido llevar a, por ejemplo, seleccionar casos distintos que llevarían a valores muestrales también diferentes.

La estimación de parámetros como procedimiento de la estadística inferencial, rigurosamente hablando, se hace formulando un **intervalo de confianza**. Así, en sentido estricto, la estimación de un promedio debería incluir la información que se menciona en el siguiente ejemplo: el promedio de años de educación de la población de quince y más años, **estimado** a partir de **información muestral**, se encuentra entre 9,2 y 10,3 años, con un 95% de **confianza**. Esto también se puede representar de modo más simple dando el valor del estimador muestral y del error estándar de dicha estimación (a partir del cual se calcula el intervalo de confianza y que se computa de forma distinta de acuerdo al tipo de muestreo conducido). Toda esta información debería estar presente cuando se usen los datos que se desea comunicar, aunque no necesariamente de esa forma. Por ejemplo, el hecho de tratarse de una encuesta podría mencionarse en la nota sobre la fuente y el nivel de confianza puede incluirse también en una nota explicativa. Por su parte, el tamaño del intervalo de confianza de la estimación puede indicarse de las siguientes maneras: (i) como un intervalo, tal y como aparece en la frase precedente; (ii) como un intervalo puesto junto al valor puntual; (iii) como una nota al pie; o (iv) utilizando un recurso gráfico.

Para estimar parámetros poblacionales se sigue el siguiente procedimiento:

Donde:

μ : media poblacional

\bar{x} : media muestral

z : nivel de confianza

E : error estándar que se estima, en el caso de SRS,¹⁴

Para estimar medias:

E

mediante: $E = \frac{1}{\sqrt{n}}$ (disminución estándar muestra)
(n : tamaño de la muestra)

$$\mu = \bar{x} \pm z^*$$

Para estimar proporciones:

P_u : proporción poblacional

p : proporción muestral

z : nivel de confianza

E : error estándar de la proporción que se estima, en el caso de SRS, mediante

$$p_u = p \pm (z * E)$$

$$\sigma_p = \left(\frac{p^*(1-p)}{n} \right)$$

donde n : tamaño de la muestra

3.3 Pruebas de hipótesis (diferencias de medias y proporciones)

Un conjunto adicional de recursos que brinda la estadística inferencial se refiere a la prueba de hipótesis.

La prueba de hipótesis estadística es también, como el muestreo aleatorio, un procedimiento probabilístico y consta de dos momentos:

- (i) en el primero se plantea una hipótesis que se va a someter a prueba, a la que se llama **hipótesis cero** o **hipótesis nula**. La prueba consiste en asumir que lo afirmado por esta hipótesis corresponde a la realidad poblacional y luego comparar ese valor con el efectivamente observado en la muestra. Si el valor observado se distancia de modo **significativo** de lo afirmado por la **hipótesis nula**, entonces esta puede ser rechazada. La significancia de la distancia entre el valor muestral y lo planteado por la **hipótesis nula** es definida por la probabilidad de obtener el valor observado

¹⁴ En el caso de otro tipo de muestreo, que es lo más usual, los procedimientos de cálculo del error estándar son diversos. Quepa aquí subrayar que no debe usarse el procedimiento propio en SRS.

en caso de que tal hipótesis sea cierta; si dicha probabilidad es muy baja, entonces hay motivos para descartar la **hipótesis nula**;

- (ii) si se ha rechazado la **hipótesis nula**, el segundo momento consiste en aceptar una **hipótesis alternativa**.

La prueba de hipótesis, por tanto, implica siempre dos riesgos: (i) rechazar una hipótesis que debería ser aceptada; y (ii) aceptar una hipótesis que debería ser rechazada.

¿Cuál es el error que uno está más dispuesto a aceptar? La respuesta a esta pregunta define la forma como se conduce la prueba de hipótesis, es decir, cuánto se **exige** de la diferencia entre el valor muestral y el de la **hipótesis nula**. Por ejemplo, si se está haciendo pruebas acerca de la potencial toxicidad de un medicamento, siempre es preferible cometer el error de rechazar uno que podría ser inocuo, que el de aceptar uno que podría ser tóxico.

3.3.1 Pasos en la prueba de hipótesis

Dado lo anterior, la prueba de hipótesis sigue los siguientes pasos típicos: (i) formular una hipótesis a contrastar; (ii) tomar el valor muestral y calcular un valor en términos de su probabilidad de ocurrencia dada la hipótesis nula; (iii) aceptar o rechazar la hipótesis nula.

El elemento crítico es el segundo paso aquí mencionado, por lo que cabe ilustrarlo con mayor detalle:

- La hipótesis nula define dos elementos de la prueba: el valor que se acepta como característico de la población y el sentido de la prueba. Este segundo elemento refiere a si se desea trabajar con la probabilidad de que el valor muestral sea diferente¹⁵, menor o mayor¹⁶ que el de la hipótesis nula. Por lo mismo, una prueba de hipótesis suele ser tanto bilateral, o de ambas colas de la curva¹⁷, como unilateral, o de una sola cola de la curva¹⁸.
- Una vez que se define la hipótesis nula, se usa la función o curva de probabilidades (como la curva normal u otra que corresponda según la prueba que se esté realizando) para calcular la probabilidad de ocurrencia de distintos valores. En una prueba bilateral, por ejemplo, la probabilidad de obtener un valor superior a tres desviaciones estándar por encima del valor típico asumido es de 0,13%, es decir, muy pequeña.

¹⁵ Es decir, puede ser mayor o menor y por lo mismo se usa probabilidades correspondientes a la distribución hacia ambos lados del valor típico asumido.

¹⁶ En estos casos solo se usan probabilidades correspondientes a las mitades de la distribución que presentan valores menores o mayores al asumido como típico.

¹⁷ Si se desea verificar diferencias sin importar el sentido de las mismas.

¹⁸ Si se desea verificar diferencias en un sentido particular: mayor o menor que.

- Luego de haber calculado la probabilidad de ocurrencia del valor muestral observado en términos de la curva, se verifica si este es muy probable dada la hipótesis nula. Así, si el valor calculado es $z = 3,1$, vemos que se trata de un valor mayor al típico asumido en más de tres unidades de desviación estándar y sabemos, dado que la curva normal es una curva conocida, que la probabilidad de haber obtenido este valor (asumiendo que la hipótesis nula es válida) es menor al 0,13%.
- En un caso como el anterior, es posible concluir que estamos en presencia de un valor muestral muy poco probable si la hipótesis nula es válida; por lo tanto, es posible concluir que la hipótesis nula debe ser rechazada con una confianza mayor al 99% ($100 - 0,13 = 99,87$).

3.3.2 Usos de la prueba de hipótesis

La prueba de hipótesis estadística es usada en situaciones como las siguientes:

- Cuando se cuenta con un valor poblacional conocido y se desea comparar con uno obtenido por muestreo para: (i) verificar si ha habido un cambio en una misma población entre dos momentos en el tiempo (la información poblacional y la información censal corresponden a la misma población pero en diferentes momentos en el tiempo); y (ii) comparar dos poblaciones distintas donde para una se cuenta con información poblacional y para la otra con información muestral. En estos casos se utiliza las siguientes fórmulas para el cálculo del valor de la prueba siempre que se trate de SRS (consultar nota 14 de este anexo):

Donde:

- z : valor de la prueba expresado como valor en la distribución normal
- \bar{x} : media muestral
- μ : media poblacional
- σ : desviación típica poblacional
- n : tamaño de la muestra

En el caso de **diferencias entre media poblacional y media muestral:**

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

En el caso de **diferencias entre proporción poblacional y proporción muestral:**

$$z = \frac{p - Pu}{\sqrt{P * Q / n}}$$

- p : proporción observada en la muestra
- Pu : proporción observada en la población
- $P * Q$: varianza poblacional
- n : tamaño de la muestra

- Cuando se cuenta con información muestral de dos poblaciones o de la misma población pero para dos momentos distintos en el tiempo. En el primer caso, la prueba de hipótesis permite determinar si las diferencias en los valores muestrales son suficientes como para, por ejemplo, descartar que las poblaciones sean similares. En el segundo caso, la prueba de hipótesis permite determinar si las diferencias observadas en dos momentos distintos son suficientes como para rechazar la hipótesis que sugiere que no ha habido cambios. En estos casos se utiliza las siguientes fórmulas para el cálculo del valor de la prueba, siempre que se trate de SRS (consultar la nota 14 de este anexo):

Donde:

V : valor calculado de la prueba (z si se usa una curva normal, t si se usa una curva t de Student)

\bar{x}_1 : media en la muestra 1

\bar{x}_2 : media en la muestra 2

En el caso de **diferencias entre medias**:

$$V = \frac{\bar{x}_1 - \bar{x}_2}{\sigma_{dif}}$$

σ_{dif} : con muestras de igual tamaño:

$$\sigma_{dif} = \sqrt{\sigma_{x1}^2 + \sigma_{x2}^2}$$

σ_{dif} : con muestras de diferente tamaño:

$$\sigma_{dif} = \sqrt{\frac{(n_1 * \sigma_1^2) + (n_2 * \sigma_2^2)}{n_1 + n_2 - 2}}$$

En el caso de **diferencias entre proporciones**:

$$z = \frac{p_1 - p_2}{\sigma_{dif}}$$

z : valor calculado de la prueba

p_1 : valor de la proporción en la muestra 1

p_2 : valor de la proporción en la muestra 2

$$\sigma_{dif} = \sqrt{(Pu * (1 - Pu)) * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Pu :

$$Pu = \frac{(n_1 * p_1) + (n_2 * p_2)}{n_1 + n_2}$$

Como ya se anotó, en todos los casos el valor calculado mediante la prueba permitirá tomar una decisión sobre la hipótesis de prueba (cero o nula), la misma que se basa en determinar la probabilidad de obtener el valor calculado. Es decir, asumiendo que la hipótesis cero es válida, se calcula la probabilidad de obtener el valor efectivamente

observado; si esta probabilidad es lo suficientemente pequeña, se procederá a rechazar la hipótesis nula con un determinado nivel de confianza estadística.

La confianza estadística es usualmente expresada como porcentaje (se rechaza la hipótesis nula con un 90, 95, 99% de confianza) o como su complemento, expresado como fracción de 1 (se rechaza la hipótesis nula con un 0,1, 0,05, 0,01 de confianza).

3.4 Pruebas de significancia (asociación bivariable)

Las pruebas de significancia para la asociación entre dos variables tienen como propósito determinar si es posible asumir que una asociación observada a nivel muestral también se verifica a nivel poblacional.

En el acápite sobre estadística descriptiva se presentaron algunos coeficientes que se utilizan para medir la asociación entre dos variables. Aquí, a continuación, se presenta las pruebas que se utilizan para verificar la significancia estadística de dichos coeficientes.

Las pruebas de significancia más usadas a nivel nominal son las siguientes:

- En el caso del coeficiente Q de Yule, la prueba consiste en calcular el intervalo de confianza en el que se encontrará el valor poblacional de dicho coeficiente. Si dicho intervalo excluye el valor cero, entonces es posible asumir que la relación también existe a nivel poblacional.

Así, se tiene que $Q = q \pm (z * \sigma_q)$, donde Q es el valor poblacional estimado, q el valor observado en la muestra, z el nivel de confianza de la estimación expresado como valores de la curva normal y σ_q el error estándar de q . Este último valor se calcula de la siguiente manera:

$$\sigma_q = \sqrt{\frac{(1-q)^2 * \left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}\right)}{4}}$$

Donde los valores a , b , c , y d corresponden a las celdas de la tabla dos por dos que se usó para el cálculo del coeficiente (véase el acápite sobre análisis descriptivo bivariable presentado en la página 338).

- En el caso del coeficiente φ , la prueba consiste en calcular un valor crítico en términos de la distribución de χ^2 . Este valor luego es comparado con el correspondiente al nivel de confianza esperado y se toma una decisión en función a dicha comparación.

Este valor se calcula de la siguiente manera: $\chi^2 = \varphi^2 * n$, donde n es el tamaño de la muestra y φ el valor del coeficiente observado en ella.

- En el caso del coeficiente V de Cramer, se usa un procedimiento similar, de modo que se calcula también un valor crítico de la distribución de χ^2 .

Este valor se calcula de la siguiente manera: $\chi^2 = v^2 * n$, donde n es el tamaño de la muestra y v el valor del coeficiente observado en la muestra.

Las pruebas de significancia más usadas a nivel ordinal son las siguientes:

- En el caso del coeficiente γ , la prueba consiste en calcular un valor crítico en términos de la distribución normal (z). Este valor luego es comparado con el valor correspondiente al nivel de confianza esperado y se toma una decisión en función a dicha comparación.

El valor z se calcula mediante la fórmula:

$$z = \sqrt{\frac{|\sum fac - \sum finv|}{n * (1 - \gamma^2)}}$$

Donde n es el tamaño de la muestra, γ el valor del coeficiente observado en ella, *fac* las «frecuencias de acuerdo» y *finv* las «frecuencias de inversión» (estas se explican en el acápite sobre análisis descriptivo bivariable antes presentado, ver la página 338).

- En el caso del coeficiente T de Kendall, la prueba consiste en calcular un valor crítico en términos de la distribución normal (z). Este valor luego es comparado con el valor correspondiente al nivel de confianza esperado y se toma una decisión en función a dicha comparación.

Este valor se calcula únicamente cuando se cuenta con un número de observaciones en la muestra mayor a diez casos y mediante la fórmula:

$$z = \frac{T}{\sqrt{\frac{2 * (2n + 5)}{9n * (n - 1)}}}$$

Donde n es el tamaño de la muestra y T el valor del coeficiente observado en ella.

- En el caso del coeficiente ρ de Spearman, la prueba consiste en calcular un valor crítico en términos de la distribución t de Student. Este valor luego es comparado con el valor correspondiente al nivel de confianza esperado y se toma una decisión en función a dicha comparación.

Este valor se calcula únicamente cuando se cuenta con un número de observaciones en la muestra mayor a treinta casos y mediante la fórmula:

$$t = \rho * \frac{\sqrt{n - 2}}{\sqrt{1 - \rho^2}}$$

Donde n es el tamaño de la muestra y ρ el valor del coeficiente observado en ella.

La prueba de significancia más usada a nivel intervalar o de razón es la siguiente:

- Coeficiente r de Pearson, donde la prueba consiste en calcular un valor crítico en términos de la distribución t de Student. Este valor luego es comparado con el correspondiente al nivel de confianza esperado y se toma una decisión en función a dicha comparación.

El valor se calcula mediante la fórmula:

$$t = \frac{r * \sqrt{n-2}}{\sqrt{1-r^2}}$$

Donde n es el tamaño de la muestra y r el valor del coeficiente observado en ella.

4. ESCALAS ACTITUDINALES

La medición de actitudes es un terreno de gran importancia para la educación, aunque relativamente poco explorado en los esfuerzos de generación regular de información de los sistemas de información educativa. En efecto, la medición de actitudes aparece fundamentalmente en estudios e investigaciones puntuales y, en algunos casos, como parte de los estudios destinados a la medición de logros de aprendizaje. Es particularmente importante detenerse en este tema a efectos de contrarrestar la tendencia a ver los resultados educativos solo en términos de habilidades con una gran carga instrumental.

Si las competencias que se espera que las personas logren mediante su paso por el sistema educativo son la combinación de conocimientos, habilidades y actitudes, entonces la medición de estas últimas tiene una importancia comparable a la de los dos primeros.

Por ejemplo, si siguiendo a Delors (Unesco 1996) se entiende que la educación debe propender a asegurar aprendizajes vinculados al saber, al hacer, a la identidad y a la convivencia; entonces la medición de actitudes tiene un lugar central que va más allá de la búsqueda de variables asociadas a los logros cognitivos.

Sin embargo, la medición de actitudes plantea dos problemas de singular magnitud: (i) por una parte, la formación de las actitudes está profundamente imbricada en las distintas experiencias de la vida de las personas, por lo que es muy difícil encontrar un vínculo nítido entre actitudes y experiencias educativas; (ii) por otra parte, la medición de actitudes suele requerir complejas baterías de instrumentos cuya elaboración y administración son muy complejas y necesitan trabajos especiales. Esto, sin embargo, no debe ser óbice para conducir un esfuerzo de generación de información de suma importancia para conocer el estado de la educación en un país.

La medición de actitudes usualmente se realiza mediante la construcción de escalas que dan cuenta de una actitud singular¹⁹. A estos efectos, se presenta a las personas un conjunto de ítems en un cuestionario que expresan opiniones respecto de un tema y se

¹⁹ Para una exposición sistemática y detallada sobre la construcción de escalas actitudinales, véase Padua y Ahman (1979). Este acápite se basa principalmente en la presentación hecha por estos autores.

registran las respuestas de las personas en términos del acuerdo o desacuerdo con dichas opiniones.

4.1 Escala de Likert

El caso más conocido de las escalas de medición de actitudes está dado por las de tipo Likert (creadas por Rensis Likert). Una escala Likert busca describir las predisposiciones de individuos o grupos de personas respecto de algún tema y opera presentando a las personas un conjunto de afirmaciones (ítems) respecto de las cuales se les solicita que expresen su grado de acuerdo o desacuerdo. El instrumento con el que se administra los ítems para construir una escala Likert tiene usualmente una forma como la de la tabla 1L.

Tabla 1L. Tabla para la escala tipo Likert

	Muy de acuerdo	De acuerdo	Ni de acuerdo ni en desacuerdo	En desacuerdo	Muy en desacuerdo
Ítem 1					
Ítem 2					

Por lo general, los ítems que se ofrece a las personas deben: (i) referirse al presente; (ii) no dar demasiada información, ya que se busca una reacción actitudinal, no un razonamiento; (iii) ser breves, claros y precisos (evitando la ambigüedad); (iv) expresar juicios u opiniones; (v) componer un conjunto que contenga tanto afirmaciones positivas como negativas, si bien cada ítem debe tener una sola opinión, el conjunto debe ser balanceado.

Adicionalmente, es preciso que se esté seguro de que los ítems se correspondan con la actitud que se quiere medir y estén formulados usando un vocabulario apropiado (comprensible) de acuerdo a la población objetivo.

Finalmente, los ítems deben tener capacidad discriminatoria para poder generar información: si todas las personas los contestan de la misma manera, los ítems no van a permitir caracterizar a la población, ya que la escala busca establecer un ordenamiento entre ellas (actitudes positivas, negativas o indiferentes respecto de algo en una escala ordinal).

Antes de discutir el tema de la capacidad discriminatoria, es necesario mencionar cómo se codifica las respuestas. En principio, es posible asignar a cada respuesta un valor numérico (- 2, - 1, 0, 1, 2; o 1, 2, 3, 4 y 5, respectivamente); sin embargo, hay que notar

que estas son escalas ordinales y los valores numéricos, si bien útiles a efectos de procesar los datos, no deben llevar a que los resultados sean tratados como en el caso de variables continuas (de nivel intervalar o de razón, según se hayan codificado). Es decir, el recurso numérico para codificar y valorar las respuestas no debe velar o alterar la naturaleza de lo que se está midiendo.

A efectos de que la medición funcione con propiedad, los ítems deben ser piloteados con una muestra de **jueces**, es decir, con un grupo de personas que ayudan a calificar la capacidad discriminatoria de los ítems. Los jueces han de responder al instrumento y, por lo mismo, es importante que sean personas con atributos similares a los de la población a la que finalmente se le administrará el instrumento. Se debe asegurar esta similitud con relación a variables que se considere relevantes respecto a la actitud a medir. Por ejemplo, si se trata de una actitud que es posiblemente afectada por el nivel de escolaridad de las personas, los jueces deberán presentar características educativas similares a las de la población objetivo: si son estudiantes de educación universitaria, los jueces deberían serlo también.

Una vez administrado el instrumento a los jueces, se procede al análisis de la capacidad discriminatoria de los ítems. Esta capacidad discriminatoria puede ser medida de diversas maneras:

(i) Capacidad discriminatoria de cada ítem

- Se hace un *ranking* de los puntajes obtenidos por cada persona en la muestra y se selecciona a los grupos extremos: al 25% superior e inferior en términos de su puntaje agregado.
- Se calcula el valor promedio obtenido por cada ítem para cada grupo.
- Se hace una prueba de las diferencias entre estos dos valores y se determina (mediante una prueba *t* de Student) qué ítems presentan diferencias estadísticamente significativas entre los dos grupos.

(ii) Relación ítem-prueba

- Se calcula el nivel de asociación (mediante un coeficiente de correlación biserial) entre los puntajes de cada ítem y los puntajes totales obtenidos por cada individuo y se selecciona los ítems con mayor asociación.

(iii) Prueba de la mediana

- Se hace un *ranking* de los puntajes obtenidos por cada persona en la muestra y se selecciona a los grupos extremos: al 25% superior e inferior en términos de su puntaje agregado.
- Se calcula la mediana de cada ítem para todos los casos seleccionados en el paso anterior (la mitad de las observaciones)
- Se dicotomiza los valores observados (por encima y por debajo de la mediana) y se construye para cada ítem un cuadro de doble entrada como la tabla 1M.

Tabla 1M. Tabla de doble entrada en base a mediana

	25% con puntajes más altos	25% con puntajes más bajos
Número de puntajes por debajo de la mediana combinada	<i>a</i>	<i>b</i>
Número de puntajes por encima de la mediana combinada	<i>c</i>	<i>d</i>

- Dependiendo del número de casos, se aplica una prueba de χ^2 (si hay al menos cuarenta observaciones) o de Fisher.

El valor de χ^2 se calcula mediante la siguiente fórmula:

$$\chi^2 = \frac{n * \left[(ad - bc) - \frac{n}{2} \right] * 2}{(a + b) * (c + d) * (a + c) * (b + d)}$$

Por su parte, para la prueba de Fischer se usa la siguiente fórmula:

$$p = \frac{(a + b)! * (c + d)! * (a + c)! * (b + d)!}{n! * a! * b! * c! * d!}$$

- Posteriormente, se selecciona los ítems que muestran diferencias significativas según la prueba que se determine adecuada.

De estos tres procedimientos, el último es el que se corresponde mejor con la naturaleza ordinal de las escalas tipo Likert, ya que los dos anteriores utilizan el artificio numérico usado para codificar las respuestas como si la variable fuera también de naturaleza numérica.

Realizado el análisis de los ítems, es posible proceder a confeccionar el instrumento final que ha de ser administrado a la población o muestra correspondiente.

Finalmente, es necesario subrayar que este tipo de escalas es de confección relativamente fácil, pero el procedimiento para ello no permite verificar si efectivamente los ítems están midiendo de modo singular una sola dimensión. Por lo mismo, el trabajo con jueces es de capital importancia así como discutir con detenimiento los resultados que se observen. En todo caso, la realización de pruebas más complejas de dimensionalidad es altamente recomendable.

4.2 Escala de Thurstone

Un segundo tipo de escalas fue elaborado por Louis Leon Thurstone y lleva su nombre. Esta escala tiene la pretensión de establecer «intervalos aparentemente iguales», lo que permite operar los valores como correspondientes a un nivel de medición intervalar.

La escala se construye partiendo por definir un conjunto muy grande de ítems que serán también sometidos a una muestra de jueces. En este caso, los jueces son expertos en el tema y no necesariamente tienen características comunes con la población.

Así, la operación con ellos se centra en el estímulo y no en la persona. Es decir, en vez de pedir a los jueces que se ubiquen ellos mismos en la escala (como en el caso de la escala tipo Likert), se les pide que ubiquen los ítems en ella. Así, cada ítem expresa un juicio u opinión sobre un determinado tema que puede ser categorizado por cada juez como representativo de un punto en la escala (si se inclina más hacia un extremo, por ejemplo, el ítem expresa simpatía respecto de algo; y si se inclina más hacia el otro extremo, por ejemplo, el ítem expresa antipatía respecto de algo; el ítem también puede ser indiferente o neutro respecto de la escala)²⁰.

Como en el caso anterior, los ítems deben: (i) referirse al presente; (ii) no dar demasiada información, ya que se busca una reacción actitudinal, no un razonamiento; (iii) ser breves, claros y precisos (evitando la ambigüedad); (iv) expresar juicios u opiniones; (v) componer un conjunto que contenga tanto afirmaciones positivas y negativas como neutrales (cada ítem debe tener una sola opinión, pero el conjunto debe ser balanceado: un tercio de cada tipo).

Los ítems serán evaluados por los jueces y, a partir de ello, se les asignará un valor en una escala de 11 puntos, donde 1 corresponde a la actitud extrema hacia un lado del continuo; 11, a la actitud extrema hacia el otro lado del continuo; y 6 a una actitud neutral. Así, el instrumento que se utilizará con los jueces tendrá una forma como la de la tabla 1N.

Tabla 1N. Tabla para la escala tipo Thurstone

	Puntaje (1 a 11)
Ítem 1	
Ítem 2	

Construir una escala de Thurstone presupone que los jueces cuenten con una explicación detallada previa, en la que se les informa que la escala a usar para asignar los

²⁰ Ejemplos típicos de la escala de Thurstone son las escalas que miden opiniones políticas (liberal/conservadora; radical/conservadora; de izquierda/de derecha; etc.).

puntajes tiene la forma de la tabla 1Ñ (usando como ejemplo una escala sobre actitudes en política).

Tabla 1Ñ. Tabla base para la escala tipo Thurstone

Puntajes posibles										
1	2	3	4	5	6	7	8	9	10	11
-----	-----	---	--	-	- +	+	++	+++	++++	++++
Conservadurismo					Neutra- lidad	Liberalismo				

Nótese que en la tabla 1Ñ las categorías podrían invertirse en la escala (en el ejemplo, «Conservadurismo» podría ir a la derecha y «Liberalismo» a la izquierda). Esto muestra que, en realidad, lo que se está midiendo sigue siendo de naturaleza ordinal.

Por lo tanto, a la hora de analizar el comportamiento de los ítems con la muestra de jueces es conveniente utilizar medidas que se correspondan con una escala ordinal, calculando medianas y rango intercuartil.

La selección de los ítems se realiza considerando que se requiere al menos dos para cada valor de la escala. Para ello se identifica la mediana de cada uno de ellos. Si se cuenta con un número amplio de ítems aplicados a los jueces, es posible identificar para cada valor en la escala un número importante de ítems y luego seleccionar aquellos que presenten menor dispersión (rango intercuartil).

Una vez que se han seleccionado los ítems (al menos dos por valor en la escala, es decir, al menos veintidós ítems para medir una sola actitud), se compone el instrumento a administrar a la población o muestra.

Este instrumento tendrá la forma de la tabla 1O.

Tabla 1O. Instrumento a administrar para aplicar la tabla Thurstone

Ítem 1	De acuerdo
	En desacuerdo
Ítem 2	De acuerdo
	En desacuerdo

Esta forma tiene un modo tal que cada persona debe reaccionar a cada ítem indicando si está de acuerdo o no con lo que este afirma. No se pide un grado de acuerdo, pues la intensidad de la actitud será determinada por el valor que se asigna en la escala a cada ítem. Cada respuesta del tipo «acuerdo» recibirá un valor equivalente a la mediana para el ítem que se calculó con la muestra de jueces.

Esta escala permite tener seguridad respecto a que el instrumento ha de cubrir todo el espectro de la escala (pues los ítems fueron creados para ese efecto). Sin embargo, también puede presentar problemas vinculados a no asegurar la unidimensionalidad de lo que se está midiendo.

4.3 Escala de Guttman

Finalmente, otro tipo de escala actitudinal es el creado por Louis Guttman y, como en los dos casos anteriores, también es conocida con el nombre de su creador. La escala de Guttman fue creada a efectos de subsanar el problema relativo a asegurar la unidimensionalidad en la medición de las actitudes (es decir, a tener certeza de que se está midiendo una sola cosa) mediante la búsqueda de coherencia en las respuestas de las personas²¹.

La escala se empieza a construir de un modo muy similar al que vimos en el caso de la tipo Likert: un conjunto de ítems se administra a un grupo de jueces que tienen características similares a las de la población a estudiar y, finalmente, se selecciona los ítems de acuerdo a su capacidad discriminatoria.

A partir de ahí, el procedimiento varía, ya que ahora se hace necesario un análisis de escalamiento de las respuestas vinculado al tema de la coherencia. Para ello se construye un escalograma como el que se muestra en la tabla 1P.

Tabla 1P. Escalograma inicial para la escala tipo Guttman

Ranking	Juez ID	Puntaje	Ítem 4	Ítem 1	Ítem 6	Ítem 3	Ítem 5	Ítem 2
			2	4	5	8	9	10
1	J04	6	A	A	A	A	A	A
2	J03	6	A	A	A	A	A	A
3	J05	5	N	A	A	A	A	A
4	J10	5	N	A	A	A	A	A
5	J08	4	N	N	A	A	A	A
6	J06	3	N	N	N	A	A	A
7	J09	3	N	A	N	A	A	A
8	J01	2	N	N	N	A	A	A
9	J02	2	N	N	N	N	A	A
10	J07	1	N	N	N	N	N	A

El escalograma de la tabla 1P tiene algunas características que es necesario identificar:

- Solo hemos incluido diez jueces y seis ítems por tratarse de un ejemplo; en la realidad estos números son mucho mayores.

²¹ Esto, por su parte, también puede ser discutido, ya que: ¿por qué tendríamos que asumir que las personas somos coherentes, especialmente cuando de actitudes se trata?

- Los jueces están ordenados por su puntaje (el número de ítems con el que expresaron acuerdo). Su identificación se da solo para destacar que el orden obedece al puntaje y no a ningún otro criterio. De hecho, habría que excluir la identificación, a menos que se quiera explorar potenciales anomalías.
- Se ha presentado un patrón casi «perfecto», donde todas las respuestas de acuerdo (A), excepto una destacada en negrita, se encuentran en un lado de la tabla y todas las respuestas de no acuerdo (N) se encuentran en el otro lado (la zona sombreada). Se debe señalar que en la práctica se suelen presentar «anomalías», es decir, acuerdos, que aparecen en medio de los no acuerdos (zona sombreada) y viceversa.

En este caso, se tiene un «error» sobre un total de sesenta respuestas posibles (diez jueces y seis ítems dicotómicos). Esa información permite calcular un coeficiente de concordancia de la siguiente manera:

$$c = 1 - \frac{\text{errores}}{\text{observaciones}} = 1 - \frac{1}{60} = 0,983$$

Es decir, se obtiene un valor alto de concordancia o replicabilidad. Ahora bien, dado que situaciones como esta no son muy esperables en la realidad, el problema de la selección de ítems se transforma en uno de búsqueda de la mejor combinación posible que asegure el coeficiente de concordancia más elevado posible.

Esto suena relativamente fácil, aunque pueda ser tedioso. Sin embargo, el procedimiento se complica cuando cada ítem es categorizado en más de dos variables (grados de acuerdo o desacuerdo y no la dicotomía acuerdo/no acuerdo). Sin embargo, si bien las técnicas a usar pueden ser más complejas, la lógica es la misma: buscar la combinación empírica que asegure la mayor concordancia, ya que así se maximiza la unidimensionalidad de la medida, lo cual es el propósito que justifica el esfuerzo.

La versión final del instrumento a utilizar estará compuesta por los ítems seleccionados y tendrá una forma similar a la que se usa para una escala Likert, como la tabla 1Q.

Tabla 1Q. Escalograma final para la escala tipo Guttman

	Totalmente de acuerdo	De acuerdo	Ni de acuerdo ni en desacuerdo	En desacuerdo	Totalmente en desacuerdo
Ítem 1					
Ítem 2					

Todo el trabajo realizado a efectos de asegurar el escalamiento unidimensional de los ítems permite definir a la escala de Guttman como una que ordena tanto a las personas como a los ítems usados. Es decir, mientras las escalas de Likert se centran en el sujeto y las tipo Thurstone en el instrumento, las tipo Guttman toman el comportamiento de ambos (personas e instrumentos) a efectos de determinar la respuesta de cada persona.

* * *

El trabajo aquí descrito con relación a la construcción de pruebas actitudinales es muy similar al que se desarrolla para la construcción de pruebas de logro académico. En estos casos también es preciso construir ítems que midan de modo unidimensional un determinado constructo y, cuando se utiliza la «teoría de respuesta al ítem», se termina identificando el lugar tanto de las personas como de los ítems en una escala común que da cuenta de la intensidad o dificultad del constructo medido.

Finalmente, cabe reiterar que este anexo solo pretende ser una brevísima introducción que, fundamentalmente, enfatiza que los métodos por utilizar han de corresponder a la naturaleza de lo que se está observando: (i) métodos propios al nivel de medición; (ii) métodos que corresponden al tipo de muestreo usado;)iii) tener claro el significado de los análisis realizados; etc.

ANEXO 2. CRONOLOGÍA DEL DESARROLLO INTERNACIONAL DE LA INFORMACIÓN EDUCATIVA

- 1853 El Primer Congreso Estadístico Internacional (celebrado en Bruselas [Bélgica] en septiembre de este año) aprueba recomendaciones sobre diversas áreas, incluyendo las estadísticas educativas (Levi 1854: 12)¹.
- 1932 La Oficina Internacional de Educación (IBE) establecida en Ginebra [Suiza] en 1929 publica *The Organization and Public Instruction in 53 Countries* sobre la base de información estadística que empieza a acopiar.
- 1947 Primera reunión del Comité de Estadísticas Educativas, cuya creación fue recomendada por la Comisión Preparatoria para la Unesco a partir de la propuesta formulada por el gobierno de los Estados Unidos en mayo de 1946.
- 1950 Establecimiento del Servicio de Estadísticas de la Unesco y conformación de un Comité de Expertos para la estandarización de las estadísticas escolares.
- 1952-1962 La Unesco publica anualmente *Basic Facts and Figures*.
- 1955-1972 La Unesco publica los cinco volúmenes de *World Survey of Education*.
- 1958 La Décima Conferencia General de la Unesco adopta la «Recomendación sobre la normalización internacional de las estadísticas relativas a la educación».
- Se establece en el Instituto de Educación de la Unesco en Hamburgo (Alemania; hoy UIL, Instituto de la Unesco para el Aprendizaje a lo Largo de Toda la Vida) un grupo de expertos que reflexiona sobre la necesidad de desarrollar evaluaciones educativas. Este grupo devendrá en la IEA.
- 1959-1962 La IEA conduce su primer estudio piloto en doce países (se administraron pruebas a estudiantes de trece años en cinco áreas: matemáticas, comprensión lectora, geografía, ciencia y habilidades no verbales).
- 1963 Publicación del primer anuario estadístico de la Unesco.

¹ Este congreso tuvo como propósito **introducir unidad en los documentos estadísticos de todos los países** y fue organizado por la Comisión Central de Estadísticas de Bélgica. Los participantes fueron: Austria, Baden, Bavaria, Bélgica, Dinamarca, Egipto, Francia, Frankfurt, Gran Bretaña, Hamburgo, Hanover, Hesse-Casel y el Gran Ducado Hesel, Lubeck, Holanda, Noruega, Portugal, Prusia, Cerdeña, Sajonia, España, Suiza, Toscana, las dos Sicilias, Estados Unidos y Wurtemberg. Los principales temas se vincularon a población (con la recomendación de hacer censos cada diez años), territorio, emigración, agricultura, industria, comercio, navegación, presupuestos económicos (recursos y gastos de las familias pobres), pobreza e indigencia, educación y criminalidad. Los tópicos muestran claramente la agenda de preocupaciones propias de Europa a mediados del siglo XIX, del mismo modo que la composición de países señala la estructura política vigente (por ejemplo, la unificación alemana y la italiana se concretarían solo cerca de dos décadas después).

- 1964 La IEA conduce el primer Estudio Internacional sobre Matemáticas (FIMS) en doce países.
La legislación sobre derechos civiles de los Estados Unidos, requiere que se conduzca un estudio sobre equidad en las oportunidades educativas. Este estudio deriva en el Informe Coleman publicado en 1966.
- 1970-1971 La IEA conduce su estudio en seis materias (matemáticas, comprensión lectora, ciencias, francés como segunda lengua, inglés como segunda lengua y educación cívica).
- 1972 Publicación de *A Study of Wastage at School*, que incluye el modelo de cohortes reconstruidas (Unesco 1972).
- 1975 Aprobación por parte de la Conferencia Internacional de Educación (Ginebra) de la primera versión de la CINE.
- 1978 La Vigésima Conferencia General de la Unesco adopta la CINE y la «Recomendación revisada sobre la normalización internacional de las estadísticas relativas a la educación».
- 1980 El Banco Mundial publica su «Education Policy Paper», señalando que es necesaria «[una nueva] estrategia para formular e implementar políticas educativas mediante la compilación de mejor información y la mejora de los procedimientos institucionales para acopiar, procesar, almacenar y recuperar datos» (World Bank 1980: 56-57)².
- 1980-1981 La IEA conduce su segundo Estudio Internacional sobre Matemáticas (SIMS) en veinte países.
- 1983-1984 La IEA conduce su segundo Estudio Internacional sobre Ciencias (SISS) en veinticuatro países.
- 1984 Creación del SIRI en la Orealc de la Unesco Santiago (Chile); produce *La situación educativa de América Latina y el Caribe* en 1988, 1990, 1992, 1996, 2001 y 2007.
- 1985 La IEA conduce su estudio sobre composición escrita.
Se conduce en los Estados Unidos el Young Adult Literaty Survey.
- 1986 La IEA inicia su estudio longitudinal sobre educación preprimaria, completado en 2003.
- 1987 Conferencia de la OECD sobre indicadores educativos. Creación de INES (Indicators of Education Systems) en Washington (Estados Unidos).
- 1988 La Academia Nacional de Ciencias de los Estados Unidos crea un Board on Internacional Comparative Studies in Education (Bicse).

² Traducción del autor; el texto original dice: «[a new] strategy to formulate and implement education policy by gathering better information and by improving the institutional procedures for collecting, processing, analyzing, storing and retrieving data.»

- 1989 La IEA conduce su primer Estudio sobre Computadoras en la Educación (Comped).
- 1990 Primer informe de Bicse señalando la importancia de las comparaciones internacionales y de conducir estudios sobre logros de aprendizaje de los estudiantes.
- 1990-1991 La IEA conduce su Estudio en Alfabetismo Lector (RLS).
- 1991 Primera edición de la publicación bienal 1991 *World Education Report* de la Unesco.
Creación del Programa de Análisis de los Sistemas Educativos (Pasec) en los países de la Conferencia de Ministros de Educación de los Países Africanos que usan el Idioma Francés (Confemen).
- 1992 La IEA conduce su segundo Comped.
Se conduce en los Estados Unidos el Young Adult Literaty Survey.
- 1993 Inicio de los primeros estudios de Pasec.
- 1994 IALS es conducido en nueve países; en 1996 cuatro países y un territorio lo administran; ocho países y un territorio adicional lo conducen en 1998. Este estudio se desarrolló sobre la base de una encuesta nacional elaborada en los Estados Unidos en la década de 1980.
Creación del Llece, coordinado por la Unesco Santiago. Conduce estudios regionales comparados sobre logro académico en 1997 y 2006.
- 1995 Informe de Bicse (Guthrie y Hansen 1995) sobre el estado de las estadísticas educativas de la Unesco.
Artículo de Puryear (1995) sobre las estadísticas educativas internacionales.
La IEA conduce el Timms en 45 países. A partir de esta versión del estudio se estabiliza su diseño y se cuenta con una base para estudios comparables en el tiempo.
Quince países establecen el Consorcio de África Meridional y Oriental para la Supervisión de la Calidad de la Educación (Sacmeq) con el apoyo del Instituto Internacional de Planificación Educativa (IIEP) Unesco e inician una primera ronda de estudios.
- 1995-1996 La IEA conduce su Estudio sobre Segundas Lenguas en Educación (LES [Language Education Study]).
- 1996 El Llece conduce su primer estudio.
- 1996-1997 La IEA conduce su segundo Estudio sobre Educación Cívica (Cived).
- 1997 Revisión de la CINE.
Inicio del proyecto conjunto Unesco-OECD World Education Indicators Programme (WEI).
- 1997-1998 La IEA conduce su Segundo Estudio sobre Tecnologías de la Información en Educación (Sites).

- 1999 La Conferencia General de la Unesco cierra la División de Estadísticas y crea el UIS.
Se discontinúa la publicación del anuario estadístico (la última edición se hizo en 1998) con el propósito de adoptar formas más efectivas de diseminación de la información (mediante un centro de datos en línea y publicaciones con mayor valor agregado).
Publicación del artículo de Stephen Heyneman titulado «The Sad Story of Unesco's Education Statistics» («La triste historia de las estadísticas educativas de la Unesco»; Heyneman 1999).
La IEA conduce una nueva ronda de Timss.
- 2000 Primera ronda de PISA de la OECD. Este estudio se conduce cada tres años.
- 2001 El UIS se instala en Montreal (Canadá).
La IEA conduce su estudio Pirls.
- 2002 Se inicia la publicación anual del informe de seguimiento de la EPT (Global Monitoring Report, GMR) a cargo de un equipo independiente residente en la Unesco.
- 2003 El UIS inicia la publicación anual del *Global Education Digest*.
El UIS comienza su presencia en el terreno mediante la designación de oficiales en las oficinas regionales de la Unesco. Fusión con el SIRI en Latinoamérica y el Caribe; creación de AIMS (Assessment, Information Systems, Monitoring and Statistics) en Bangkok (Tailandia); proyecto de fortalecimiento de capacidades en África.
El UIS inicia LAMP.
La IEA conduce una nueva ronda de Timss.
El estudio ALL, seguidor de IALS, es conducido en seis países.
La OECD lleva a cabo la segunda ronda de PISA.
- 2005 Sacmeq conduce su segunda ronda de estudios.
- 2006 Primera evaluación externa del UIS; muy positiva, excepto con relación a la experiencia de «fortalecimiento de capacidades» en África y los problemas de viabilidad de LAMP.
La IEA conduce una nueva ronda del estudio Pirls; se establece que se realizará en ciclos de cinco años. También conduce una nueva ronda de Sites.
La Unesco Santiago lleva a cabo el segundo estudio del Llece.
ALL es conducido en siete países.
La OECD realiza la tercera ronda de PISA.

- 2007 La Conferencia General de la Unesco solicita que el UIS lidere un esfuerzo para evaluar la necesidad de una nueva revisión de la CINE para el año 2011.
La IEA lleva a cabo una nueva ronda de Timss.
Sacmeq realiza su tercera ronda de estudios.
- 2008 La IEA conduce TEDS-M.
- 2008-2009 La IEA lleva a cabo su nuevo Estudio Internacional sobre Educación Cívica y Ciudadana (Civics).
- 2009 La OECD conduce una nueva ronda de PISA. Al enfocarse nuevamente en lectura, esta ronda permite ver tendencias (2000-2009).
- 2010 Se inicia el primer ciclo de evaluaciones del Programa LAMP del UIS.
- 2011 La IEA conduce simultáneamente Timss y Pirls; esto permite evaluar la relación entre desempeños lectores y desempeños en matemáticas y ciencias (Martin y Mullis 2013).
La OECD conduce Piac.
- 2012 Nueva ronda de PISA, esta vez enfocada en matemáticas (como en 2003), lo que permite ver tendencias (2003-2009).
El UIS y la Institución Brookings establecen un acuerdo para actuar de modo conjunto como Secretariado de la Learning Metrics Task Force. Se organiza la primera reunión virtual y la primera reunión presencial de la misma. Se inician las labores de dos de los tres grupos de trabajo previstos.
- 2013 Se publican los primeros resultados de LAMP y de Piac.
La Unesco Santiago conduce el tercer estudio del Llece.
En el marco de las discusiones sobre la agenda internacional de desarrollo post-2015, el United Nations Development Group publica su informe en el que llama a una «revolución de información» (*data revolution*) (United Nations Development Group 2013).
La Learning Metrics Task Force publica sus recomendaciones acerca de qué deben aprender los niños (2013c), un marco global para la medición de la educación a escala global y las recomendaciones generales resultantes de sus reuniones, de los grupos de trabajo y de los procesos de consulta celebrados en los cinco continentes (Learning Metrics Task Force 2013b).

REFERENCIAS

- ACEDO, Clementina; Don ADAMS y Simona POPA (eds.)
2012 *Quality and Qualities. Tensions in Education Reforms*. Rotterdam y Boston: ensePublishers.
Fecha de consulta: 15/3/2014. <<http://dx.doi.org/10.1007/978-94-6091-951-0>>.
- ADAMS, Don
1993 *Defining Educational Quality*. Pittsburgh: University of Pittsburgh. Fecha de consulta: 15/3/2014. <http://pdf.usaid.gov/pdf_docs/PNACA245.pdf>.
- AKYEAMPONG, Kwame y David STEPHENS
2000 «*On the Threshold: The Identity of Student Teachers in Ghana*». Brighton: Centre for International Education-University of Sussex.
- ALLSOP, Judith y Linda MULCAHY
1998 «Maintaining Professional Identity: Doctors' Responses to Complaints». En: *Sociology of Health & Illness*, vol. 20, N° 6, pp. 802-824.
- ARCHER, Margaret
2013 *Social Origins of Educational Systems* [Kindle]. Londres y Nueva York: Routledge.
2000 *Being Human: The Problem of Agency*. Cambridge, UK y Nueva York: Cambridge University Press.
- ARCHER, Margaret; Roy BHASKAR; Andrew COLLIER; Tony LAWSON y Alan NORRIE (eds.)
2013 *Critical Realism: Essential Readings*. Londres: Routledge.
- ARCHER, Margaret y William OUTHWAITE
2004 *Defending Objectivity: Essays in Honour of Andrew Collier*. Londres: Routledge.
- ARGENTINA-MINISTERIO DE EDUCACIÓN, CIENCIA Y TECNOLOGÍA
s. f. *Sistema nacional de indicadores educativos*. Buenos Aires: Ministerio de Educación.
- ÁVALOS, Beatrice
2004 «Teacher Regulatory Forces and Accountability Policies in Chile: From Public Servants to Accountable Professionals». En: *Research Papers in Education*, vol. 19, N° 1, pp. 67-85.
- BECKER, Gary
1993 *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Chicago: The University of Chicago Press.
- BENAVOT, Aaron
1992 «Curricular Content, Educational Expansion, and Economic Growth». En: *Comparative Education Review*, vol. 36, N° 2, pp. 150-174.

BHASKAR, Roy

1998 *The Possibility of Naturalism: A Philosophical Critique of the Contemporary Human Sciences*. Londres: Routledge.

1975 *A Realist Theory of Science*. York: Books.

BLANCO, Rosa

2005 *Versión preliminar del módulo innovación educativa y calidad de la educación*. Santiago de Chile: Unesco.

BOLLEN, Kenneth A.

2002 «Latent Variables in Psychology and the Social Sciences». En: *Annual Review of Psychology*, N° 53, pp. 605-634.

BORSBOOM, Denny; Gideon MELLENBERGH y Japp VAN HERDEN

2003 «The Theoretical Status of Latent Variables». En: *Psychological Review*, vol. 110, N° 2, pp. 203-219.

BOURDIEU, Pierre

2000 *Esquisse d'une théorie de la pratique: précédé de trois études d'ethnologie kabyle*. París: Seuil.

BRUNS, Barbara; Alain MINGAT y Ramah RAKOTOMALALA

2003 *Achieving Universal Primary Education by 2015. A Chance for Every Child*. Washington, DC: The World Bank.

BUCHERT, Lene

1995 «The Concept of Education for All: What has Happened after Jomtien?» En: *International Review of Education*, vol. 41, N° 6, pp. 537-549.

CARNOY, Martin; Jeffery H. MARSHALL y Amber GOVE

2007 *Cuba's Academic Advantage: Why Students in Cuba do Better in School*. Stanford, Calif.: Stanford University Press.

CEPAL

2011 *Propuesta de código de buenas prácticas de las estadísticas en América Latina y el Caribe y plan de implementación para 2012-2013*. Bávaro: Cepal.

CHILE-MINISTERIO DE EDUCACIÓN

2007 *Indicadores de la educación en Chile 2006*. Santiago de Chile: Ministerio de Educación.

COCKCROFT, Wilfred. H.

1982 *Mathematics Counts. Report of the Committee of Inquiry into the Teaching of Mathematics in Schools under the Chairmanship of Dr. W. H. Cockcroft*. Londres: H. M. S. O. Fecha de consulta: 15/3/2014. <<http://www.educationengland.org.uk/documents/cockcroft/cockcroft1982.html>>.

COLEMAN, James S.; Ernest Q. CAMPBELL; Carol J. HOBSON; James MCPARTLAND; Alexander M. MOOD; Frederic D. WEINFELD y Robert YORK

1966 *Equality of Educational Opportunity*. Washington, DC: US Government Printing Office.

COLLINS, Linda

2010 *Latent Class and Latent Transition Analysis: With Applications in the Social Behavioral, and Health Sciences*. Hoboken N.J.: Wiley.

COOPER, Barry y Máiréad DUNNE

2000 *Assessing Children's Mathematical Knowledge. Social Class, Sex and Problem-solving*. Buckingham, Philadelphia: Open University Press.

1998 «Anyone for Tennis? Social Class Differences in Children's Responses to National Curriculum Mathematics Testing». En: *The Sociological Review*, vol. 46, N° 1, pp. 115-148.

CROUCH, Luis; Jenny PERLMAN ROBINSON y Lauren GREUBEL

2013 «Crafting an Education Goal in the Post-2015 Development Framework: Having Our Cake and Eating It Too». Brookings, 17 de mayo. Fecha de consulta: 15/3/2014. <<http://www.brookings.edu/blogs/education-plus-development/posts/2013/05/17-crafting-education-goal-post-2015-robinson>>.

DARVILLE, Richard

1999 «Knowledges of Adult Literacy: Surveying for Competitiveness». En: *International Journal of Educational Development*, N° 19, pp. 273-285.

DAY, Christopher; Alison KINGTON; Gordon STOBART y Pam SAMMONS

2006 «The Personal and Professional Selves of Teachers: Stable and Unstable Identities». En: *British Educational Research Journal*, vol. 32, N° 4, pp. 601-616.

DE WAAL, Frans

2013 *The Bonobo and the Atheist: In Search of Humanism Among the Primates* [Kindle]. Nueva York y Londres: W. W. Norton & Company.

DEMEUSE, Marc; Marcel CRAHAY y Christian MONSEUR

2001 «Efficiency and Equity». En: HUTMACHER, W.; D. COCHRANE y N. BOTTANI (eds.). *In Pursuit of Equity in Education. Using International Indicators to Compare Equity Policies*. Nueva York, Boston y Dordrecht: Kluwer Academic Publishers, pp. 65-91.

DUNNE, Máiréad; John PRYOR y Paul YATES

2005 *Becoming a Researcher: A Research Companion for the Social Sciences*. Maidenhead: Open University Press.

EASTERLY, William

2013 *The Tyranny of Experts: Economists, Dictators, and the Forgotten Rights of the Poor* [Kindle]. Nueva York: Basic Books.

EINSTEIN, Albert

1934 «On the Method of Theoretical Physics». En: *Philosophy of Science*, vol. 1, N° 2, pp. 163-169.

ESCOCIA-HM INSPECTORATE OF EDUCATION

- 2007 *How Good is our School? The Journey to Excellence. Part 3*. Edimburgo: HMIE. Fecha de consulta: 15/3/2014. <http://www.educationscotland.gov.uk/Images/HowgoodisourschoolJtEpart3_tcm4-684258.pdf>.

EUROPEAN COMMISSION

- 2006 *Classification of Learning Activities. Manual*. Luxemburgo: EC.
 1999 *Evaluating Socio-economic Programmes. Vol. 2: Selection and Use of Indicators for Monitoring and Evaluation*. Luxemburgo: Office for Official Publications of the European Communities.

GIDDENS, Anthony

- 1984 *The Constitution of Society*. Berkeley: University of California Press.

GLASER, Robert

- 1963 «Instructional Technology and the Measurement of Learning Outcomes: Some Questions». En: *American Psychologist*, vol. 18, N° 8, pp. 519-521.

GOOD HUMANITARIAN DONORSHIP

- 2003 «23 Principles and Good Practice of Humanitarian Donorship». GHD. Fecha de consulta: 15/3/2014. <<http://www.ghdinitiative.org/ghd/gns/principles-good-practice-of-ghd/principles-good-practice-ghd.html>>.

GUADALUPE, César

- 2015 «How Feasible is it to Develop a Culturally-sensitive Large-scale Standardised Assessment of Literacy Skills?» En: HAMILTON, M.; B. MADDOX y C. ADDEY (eds.). *Literacy as Numbers: Researching the Politics and Practices of International Literacy Assessment Regimes*. Cambridge: Cambridge University Press, pp. 111-128.
 2013 «Estadísticas e indicadores educativos: reflexiones generales». En: KISILEVSKY, M. y E. ROCA (eds.). *Indicadores, metas y políticas educativas*. Madrid: OEI, pp. 47-60. Fecha de consulta: 15/3/2014. <<http://www.oei.es/metas2021/EVAL2.pdf>>.
 2010 «Unidad 5. Estadísticas e indicadores educativos». Materiales del curso experto universitario de indicadores y estadísticas educativas. Universidad Nacional de Educación a Distancia de España. Madrid: OEI.
 2009 *Teachers in Peruvian Public Primary Schools: Views on Teaching as a Profession in a Challenging Context*. Brighton: University of Sussex. Tesis doctoral. Fecha de consulta: 15/3/2014. <<http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.506814>>.
 2006 «Understanding Teachers' Professional Action: Biography and History in Teachers' Agency» [manuscrito]. Brighton: University of Sussex.
 2002 *Indicadores de cobertura, eficiencia y flujo escolar. Necesidades de política, problemas metodológicos y una propuesta*. Santiago de Chile: Unesco. Fecha de consulta: 15/3/2014. <<http://unesdoc.unesco.org/images/0016/001617/161771s.pdf>>.
 2001 *Evolución de los principales indicadores de la educación primaria y secundaria escolarizada de menores (1993-1998)*. Lima: Meccep-Ministerio de Educación. Fecha de consulta:

15/3/2014. <<http://www2.minedu.gob.pe/umc/admin/images/publicaciones/mecep/doc08.pdf>>.

GUADALUPE, César y Manuel CARDOSO

2011 «Measuring the Continuum of Literacy Skills among Adults: Educational Testing and the LAMP Experience». En: *International Review of Education*, vol 57, N° 1-2, pp. 199-217.

GUADALUPE, César; Juana HUILLCAMISA; Liliana MIRANDA; María Luisa QUINTANA; José RODRÍGUEZ; Néstor SANTILLÁN; Juan Pablo SILVA y Gloria ZAMBRANO

2002 *La educación peruana a inicios del nuevo siglo*. Lima: Mecep-Ministerio de Educación. Fecha de consulta: 15/3/2014. <<http://umc.minedu.gob.pe/?p=381>>.

GUADALUPE, César y Paula LOUZANO

2003 *Medición de la conclusión universal de la educación primaria en América Latina*. Santiago de Chile: Unesco. Fecha de consulta: 15/3/2014. <<http://unesdoc.unesco.org/images/0015/001593/159305s.pdf>>.

GUADALUPE, César y Daniel TACCARI

2004 *Conclusión universal de la educación primaria. ¿Cómo evaluar el progreso hacia esta meta?* Santiago de Chile: Unesco. Fecha de consulta: 15/3/2014. <<http://unesdoc.unesco.org/images/0015/001592/159288s.pdf>>.

GUADALUPE, César y Alejandra VILLANUEVA

2013 «PISA 2009/2000 en América Latina: una relectura de los cambios en el desempeño lector y su relación con las condiciones sociales». En: *Apuntes*, vol. XL, N° 72, pp. 157-192. Fecha de consulta: 15/3/2014. <<http://revistas.up.edu.pe/index.php/apuntes/article/view/22/17>>.

GUTHRIE, James y Janet HANSEN (eds.)

1995 *World-wide Education Statistics: Enhancing Unesco's Role*. Washington, DC: National Academy Press.

HABERMAS, Jürgen

1984 *The Theory of Communicative Action*, vol. I. Boston: Beacon Press.

HAMBLETON, Ronald K.

2005 «Issues, Designs and Technical Guidelines for Adapting Tests Into Multiple Languages and Cultures». En: HAMBLETON, R. K.; P. MERENDA y C. SPIELBERGER (eds.). *Adapting Psychological and Educational Tests for Cross-Cultural Assessment*. Nueva Jersey: Lawrence Erlbaum, pp. 3-38.

HAMILTON, Mary

2012 *Literacy and the Politics of Representation* [Kindle]. Oxon: Routledge.

2001 «Privileged Literacies: Policy, Institutional Process and the Life of the IALS». En: *Language and Education*, N° 15, pp. 178-196.

- HAMILTON, Mary y David BARTON
 2000 «The International Adult Literacy Survey: What Does It Really Measure?» En: *International Review of Education*, N° 46, pp. 377-389.
- HANUSHEK, Eric A. y Richard R. PACE
 1995 «Who Chooses to Teach (and Why)?». En: *Economics of Education Review*, vol. 14, N° 2, pp. 101-117.
- HANUSHEK, Eric A.; Steven G. RIVKIN; Richard ROTHSTEIN y Michael PODGURSKY
 2004 «How to Improve the Supply of High-Quality Teachers». En: *Brookings Papers on Education Policy*, N° 7, pp. 7-44.
- HECKMAN, James
 2008 «Schools, Skills, and Synapses». En: *Economic Inquiry*, vol. 46, N° 3, pp. 289-324.
 2000 «Policies to Foster Human Capital». En: *Research in Economics*, vol. 54, N° 1, pp. 3-56.
- HECKMAN, James y Tim KAUTZ
 2012 «Hard Evidence on Soft Skills». En: *Labour Economics*, vol. 19, N° 4, pp. 451-464.
- HECKMAN, James; Rodrigo PINTO y Peter SAVELYEV
 2013 «Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes». En: *American Economic Review*, vol. 103, N° 6, pp. 2052-2086.
- HECKMAN, James; Jora STIXRUD y Sergio URZÚA
 2006 «The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior». En: *Journal of Labor Economics*, vol. 24, N° 3, pp. 411-482.
- HENRY, Neil W.
 1999 «Latent Structure Analysis at Fifty». En: *Annual Meeting of the American Statistical Association*, Baltimore, 11 de agosto, pp. 587-592. Fecha de consulta: 15/3/2014. <https://www.amstat.org/SECTIONS/SRMS/PROCEEDINGS/PAPERS/1999_102.PDF>.
- HEYNEMAN, Stephen. P.
 1999 «The Sad Story of Unesco's Education Statistics». En: *International Journal of Educational Development*, N° 19, pp. 65-74.
- HODAS, Steven
 1993 «Is Water an Input to a Fish? Problems with the Production-Function Model in Education». En: *Education Policy Analysis Archives*, vol. 1, N° 12, pp. 1-7.
- HUFF, Darrell
 1991 *How to Lie with Statistics* [Kindle]. Londres: Penguin Books.
- HUSÉN, Torsten
 1996 «Lessons from the IEA Studies». En: *International Journal of Educational Research*, vol. 25, N° 3, pp. 207-218.

IMF, INTERNATIONAL MONETARY FUND

2012 *Data Quality Assessment Framework. Generic Framework*. Washington, DC: IMF. Fecha de consulta: 15/3/2014. <http://dsbb.imf.org/images/pdfs/dqrs_Genframework.pdf>.

INTERNATIONAL CONSULTATIVE FORUM ON EDUCATION FOR ALL

1998 *Education for All. The Year 2000 Assessment. Technical Guidelines*. París: Unesco. Fecha de consulta: 15/3/2014. <<http://unesdoc.unesco.org/images/0011/001137/113746eo.pdf>>.

ISO, INTERNATIONAL ORGANIZATION FOR STANDARDIZATION

2009 «ISO9000:2005.QualityManagementSystems.FundamentalsandVocabulary». ISO. Fecha de consulta: 11/6/2015. <http://www.iso.org/iso/catalogue_detail?csnumber=42180>.

KENNEDY, Mary M.

2004 «Reform Ideals and Teachers' Practical Intentions». En: *Education Policy Analysis Archives*, vol. 12, N° 13. Fecha de consulta: 15/3/2014. <<http://epaa.asu.edu/ojs/article/view/168>>.

KISILEVSKY, Marta y Enrique ROCA (eds.)

2013 *Indicadores, metas y políticas educativas*. Madrid: OEI.

LALIBERTÉ, Lucie; Werner GRÜNEWALD y Laurent PROBST

2004 *Data Quality: A Comparison of IMF's Data Quality Assessment Framework (DQAF) and Eurostat's Quality Definition*. Washington, DC: IMF.

LAZARSELD, Paul y Neil HENRY

1968 *Latent Structure Analysis*. Nueva York: Houghton Mifflin.

LEARNING METRICS TASK FORCE

2013a *Toward Universal Learning: A Global Framework for Measuring Learning*. Washington, DC y Montreal: The Brookings Institution / Unesco. Fecha de consulta: 15/3/2014. <http://www.brookings.edu/-/media/Research/Files/Reports/2013/07/global-framework-measuring-learning/LMTFReport2_final.pdf?la=en>.

2013b *Toward Universal Learning: Recommendations from the Learning Metrics Task Force*. Washington, DC y Montreal: The Brookings Institution / Unesco. Fecha de consulta: 15/3/2014. <<http://www.brookings.edu/-/media/research/files/reports/2013/09/learning-metrics-task-force-universal-learning/lmtf-recommendations-report-final-web.pdf>>.

2013c *Toward Universal Learning: What Every Child Should Learn*. Washington, DC y Montreal: The Brookings Institution / Unesco. Fecha de consulta: 15/3/2014. <<http://www.brookings.edu/-/media/Research/Files/Reports/2013/02/learning-metrics/LMTFRpt1TowardUnivrsLLearning.pdf?la=en>>.

LEVI, Leone

1854 «Resume of the Statistical Congress, held at Brussels, September 11th, 1853, for the Purpose of Introducing Unity in the Statistical Documents of all Countries». En: *Quarterly Journal of the Statistical Society*, vol. 17, pp. 1-14. Fecha de consulta: 15/3/2014. <<http://archive.org/details/jstor-2338350>>.

LEWIN, Keith

2011 *Making Rights Realities. Researching Educational Access, Transitions and Equity*. Falmer: University of Sussex / DFID.

LEWIN, Keith y Janet STUART

2003 *Researching Teacher Education: New Perspectives on Practice Performance and Policy : Multi-Site Teacher Education Research Project (Muster): Synthesis Report*. Londres: Department for International Development.

LOVELESS, Tom

2013a «International Tests are not All the Same». Brookings, 9 de enero. Fecha de consulta: 18/10/2013. <<http://www.brookings.edu/blogs/brown-center-chalkboard/posts/2013/01/09-timss-pisa-loveless>>.

2013b «PISA's China Problem». Brookings, 9 de octubre. Fecha de consulta: 18/10/2013. <<http://www.brookings.edu/blogs/brown-center-chalkboard/posts/2013/10/09-pisa-china-problem-loveless>>.

LURIA, Alexander

1976 *Cognitive Development, its Cultural and Social Foundations*. Cambridge, Mass.: Harvard University Press.

MADDOX, Bryan

2015 «Inside the Assessment Machine: The Life and Times of a Test Item». En: HAMILTON, Mary; Bryan MADDOX y Camilla ADDEY (eds.). *Literacy as Numbers: Researching the Politics and Practices of International Literacy Assessment Regimes*. Cambridge: Cambridge University Press.

2014 «Globalising Assessment: An Ethnography of Literacy Assessment, Camels and Fast Food in the Mongolian Gobi». En: *Comparative Education*, vol. 50, N° 2, pp. 474-489.

MAGIDSON, Jay y Jeroen VERMUNT

2001 «Latent Class Factor and Cluster Models, Bi-plots and Tri-plots and Related Graphical Displays». En: *Sociological Methodology*, N° 31, pp. 223-264.

MARTIN, Michael e Ina V. S. MULLIS (eds.)

2013 *Timss and Pirls 2011: Relationships Among Reading, Mathematics, and Science Achievement at the Fourth Grade. Implications for Early Learning*. Chestnut Hill, MA: Timss & Pirls International Study Center-Boston College.

MARX, Karl

1983 *Grundrisse der Kritik der politischen Ökonomie*. En: *Marx-Engels Werke*, vol. 42. Berlín: Dietz Verlag.

MAURETTE, Marie-Thérèse

1948 *Educational Techniques for Peace. Do they Exist?* París: Unesco.

MCCOWAN, Tristan

2013 *Education as a Human Right: Principles for a Universal Entitlement to Learning*. Londres: Bloomsbury Academic.

MCMAHON, Walter W.

1998 «Conceptual Framework for the Analysis of the Social Benefits of Lifelong Learnings». En: *Education Economics*, vol. 6, N° 3, pp. 309-346.

1997 «Recent Advances in Measuring the Social and Individual Benefits of Education». En: *International Journal of Educational Research*, vol. 27, N° 6, pp. 449-531.

MCMAHON, Walter W. y BOEDONIO (eds.).

1992 *Education and the Economy. The External Efficiency of Education*. Tallahassee y Jakarta: Florida State University Press / Government of Indonesia.

MILL, John Stuart

1869 *On Liberty*, 4ª ed. Londres: Longmans.

MONTT, Pedro

2005 «Impacto demográfico y educación». Presentación en: Seminario Las Transformaciones Demográficas en Chile y sus Consecuencias para las Políticas Públicas. Cepal, 28 de setiembre, Santiago de Chile. Fecha de consulta: 5/3/2014. <<http://www.cepal.org/celade/noticias/paginas/7/22677/Pmontt.pdf>>.

MULLIS, Ina V. S.; Michael MARTIN; C. MINNICH; Kathleen T. DRUCKER y Moira RAGAN (eds.)

2012 *Pirls 2011 Encyclopedia*, 2 vol. Chestnut Hill, MA: Timss & Pirls International Study Center-Boston College.

MULLIS, Ina V. S.; Michael MARTIN; Chad MINNICH; Gabrielle STANCO; Alka ARORA; Victoria CENTURINO y Courtney CASTLE (eds.)

2012 *Timss 2011 Encyclopedia*, 2 vol. Chestnut Hill, MA: Timss & Pirls International Study Center-Boston College.

MUÑOZ IZQUIERDO, Carlos; María de los Ángeles NÚÑEZ GORNÉS y H. SÁNCHEZ PÉREZ

2004 *Educación y desarrollo socioeconómico en América Latina y el Caribe: desarrollo de una propuesta para la construcción de indicadores de los efectos de la educación formal en la economía y la sociedad*. México, DF: Universidad Iberoamericana.

MURILLO, Javier

2005 «Una panorámica de la carrera docente en América Latina». En: *Revista Prelac*, N° 1, pp. 52-59.

MURNANE, Richard J.

1991 *Who will Teach? Policies that Matter*. Cambridge, Mass.: Harvard University Press.

NACIONES UNIDAS

1989 *Convención sobre los derechos del niño*. Nueva York: Naciones Unidas.

- 1948 *Declaración universal de los derechos humanos*. Nueva York: Naciones Unidas.
- NAVARRO, Juan Carlos (ed.)
- 2002 *¿Quiénes son los maestros? Carreras e incentivos docentes en América Latina*. Washington, DC: Banco Interamericano de Desarrollo.
- NIETZSCHE, Friedrich
- 1880 *Der Wanderer und sein Schatten*. Projekt Gutenberg-DE. Fecha de consulta: 15/3/2014. <<http://gutenberg.spiegel.de/buch/der-wanderer-und-sein-schatten-3251/22>>.
- NUTLEY, Sandra; Huw DAVIES e Isabel WALTER
- 2002 *Evidence Based Policy and Practice: Cross Sector Lessons from the UK*. Londres: ESRC UK Centre for Evidence Based Policy and Practice.
- OCKHAM, William of
- 1340 *Summa Logicae*. Fecha de consulta: 30/7/2015. <http://www.logicmuseum.com/wiki/Authors/Ockham/Summa_Logicae>.
- OECD
- 2014 *PISA 2012 Results: What Students Know and Can Do*, vol. I (ed. revisada). París: OECD.
- 2013a *Education at a Glance: 2013 OECD Indicators*. París: OECD.
- 2013b *OECD Skills Outlook 2013. First Results from the Survey of Adult Skills*. París: OECD.
- 2012a *Education at a Glance 2012: OECD Indicators*. París: OECD.
- 2012b *PISA 2009 Technical Report*. París: OECD.
- 2010 *PISA 2009 Results: Learning Trends*. París: OECD.
- 2006a *PISA 2006. Marco de evaluación*. París: OECD.
- 2006b *Starting Strong II: Early Childhood Education and Care*. París: OECD.
- 2005 *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*. París: OECD.
- 2004 *OECD Handbook for Internationally Comparative Statistics. Concepts, Standards, Definitions and Classifications*. París: OECD.
- 2001 *Starting Strong: Early Childhood Education and Care*. París: OECD.
- OECD y CANADÁ-STATISTICS CANADA
- 2005 *Learning a Living: First Results of the Adult Literacy and Life Skills Survey*. París y Ottawa: OECD / Statistics Canada.
- 2000 *Literacy in the Information Age: Final Report of the International Adult Literacy Survey*. París y Ottawa: OECD / Statistics Canada.
- OECD y UNESCO-UIS
- 2003 *Literacy Skills for the World of Tomorrow. Further Results from PISA 2000*. París y Montreal: OECD / Unesco-UIS.
- OEI
- 2010 *Metas educativas 2021: la educación que queremos para la generación de los bicentenarios: documento final. Síntesis*. Madrid: OEI.

OSTLER, Nicholas

2006 *Empires of the Word: A Language History of the World*. Nueva York: Harper Perennial.

PADUA, Jorge e Ingvar AHMAN

1979 «Escalas para la medición de actitudes». En: PADUA, J.; I. AHMAN; H. APEZECHEA y C. BORSOTTI (eds.). *Técnicas de investigación aplicadas a las ciencias sociales*. México, DF: El Colegio de México / Fondo de Cultura Económica.

PAWSON, Ray y Nick TILLEY

1997 *Realistic Evaluation*. Londres y Thousand Oaks: Sage.

PEANO, Serge

1993 *The Financing of Education Systems*. París: Unesco.

PEREYRA, Miguel A.; Kotthoff HANS-GEORG y Robert COWEN (eds.)

2011 *PISA under Examination. Changing Knowledge, Changing Tests, and Changing Schools*. Rotterdam y Boston: Sense Publishers.

PERFETTI, Charles A. y Laurence RIEBEN

1991 *Learning to Read: Basic Research and its Implications*. Hillsdale, NJ: L. Erlbaum Associates.

PERÚ-INEI, INSTITUTO NACIONAL DE ESTADÍSTICA E INFORMÁTICA

2012 *Código de buenas prácticas estadísticas*. Lima: INEI.

PERÚ-MINISTERIO DE EDUCACIÓN

2001 *Boletín UMC 9*. Lima: Ministerio de Educación.

PIGOZZI, Mary Joy

2006 «What is “Quality of Education” (A Unesco Perspective)». En: ROSS, K. e I. JÜRGENS GENEVOIS (eds.). *Cross-national Studies of the Quality of Education: Planning their Design and Managing their Impact*. París: Unesco / IIEP, pp. 39-50.

PILOS, Spyridon

2001 «Report of the Eurostat Task Force on Measuring Lifelong Learning [TF/MLLL]». Eurostat Working Papers Population and Social Conditions 3/2001/E/N°4. Luxemburgo: Eurostat.

POST-2015 EDUCATION INDICATORS TECHNICAL ADVISORY GROUP OF THE EFA STEERING COMMITTEE

2014 «Towards Indicators for a Post-2015 Education Framework». Documento preliminar. Fecha de consulta: 10/12/2014. <<http://www.uis.unesco.org/Education/Documents/towards-indicators-for-post-2015-education-framework.pdf>>.

POSTLETHWAITE, T. Neville

2004 *Monitoring Educational Achievement*. París: Unesco / IIEP.

PREAL

2006 *Cantidad sin calidad. Un informe del progreso educativo en América Latina*. Washington, DC: Preal.

- PRIE
 2003 *La experiencia del Proyecto Regional de Indicadores Educativos 2000-2003*. Santiago de Chile: Unesco / Chile-Ministerio de Educación.
 2002 *Panorama educativo de las Américas*. Santiago de Chile: Unesco-Orealc / Chile-Ministerio de Educación. Fecha de consulta: 15/3/2014. <<http://www.prie.oas.org/espanol/documentos/Panorama%20educativo%20sp.pdf>>.
- PSACHAROPOULOS, George y M. WOODHALL
 1985 *Education for Development: An Analysis of Investment Choices*. Nueva York: World Bank / Oxford University Press.
- PURYEAR, Jeffrey M.
 1995 «International Education Statistics and Research: Status and Problems». En: *International Journal of Educational Development*, vol. 15, N° 1, pp. 79-91.
- RAE, REAL ACADEMIA ESPAÑOLA
 2001 *Diccionario de la Real Academia Española*, 22ª ed. Madrid: RAE. Fecha de consulta: 15/3/2014. <<http://www.rae.es>>.
- RAO, Jon N. K.
 2003 *Small Area Estimation*. Hoboken NJ: John Wiley & Sons, Inc.
- REINO UNIDO-DEPARTMENT OF EDUCATION
 2010 *The Importance of Teaching*. Londres: UK-Department of Education. Fecha de consulta: 15/3/2014. <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/175429/CM-7980.pdf>.
- REPÚBLICA BOLIVARIANA DE VENEZUELA
 2013 *Cumpliendo las metas del milenio. 2012*. Caracas: Instituto Nacional de Estadística. Fecha de consulta: 15/3/2014. <http://www.undp.org/content/dam/undp/library/MDG/english/MDG%20Country%20Reports/Venezuela/CUMPLIENDO_LAS_METAS_DEL_MILENIO_23-09-13.pdf>.
- REPÚBLICA BOLIVARIANA DE VENEZUELA-GABINETE SOCIAL
 2004 *Cumpliendo las metas del milenio*. Caracas: CDB Publicaciones. Fecha de consulta: 15/3/2014. <http://www.undp.org/content/dam/undp/library/MDG/english/MDG%20Country%20Reports/Venezuela/Venezuela_MDGReport_2004_SP.pdf>.
- ROBALINO, Magaly
 2005 «¿Actor o protagonista? Dilemas y responsabilidades sociales de la profesión docente». En *Revista Prelac*, N° 1, pp. 6-23.
- ROBEYNS, Ingrid
 2006 «Three Models of Education: Rights, Capabilities and Human Capital». En: *Theory and Research in Education*, vol. 4, N° 1, pp. 69-84.
- ROBINSON, Ken
 2011 *Out of our Minds: Learning to be Creative* [Kindle]. Oxford: Capstone.

RTI INTERNATIONAL

2009 *Early Grade Assessment Toolkit*. RTP, North Carolina: RTI / Usaid / World Bank.

RUTKOWSKI, David

2008 «Towards an Understanding of Educational Indicators». En: *Policy Futures in Education*, vol. 6, N° 4, pp. 470-481.

SAINT-EXUPÉRY, Antoine de

1939 *Terre des hommes* [ebook]. La Bibliothèque Electronique du Québec.

SANTIAGO, Paulo

2002 «Teacher Demand and Supply». Documento de trabajo. París: OECD. Fecha de consulta: 15/3/2014. <<http://www.oecd-ilibrary.org/content/workingpaper/232506301033>>.

SCHIEFELBEIN, Ernesto y María Clara GROSSI

1984 «Informe estadístico sobre la repetición escolar en América Latina». En: *Métodos estadísticos para mejorar la estimación de la repetición escolar: dos estudios metodológicos*. París: Unesco, pp. 1-95.

SCHMELKES, Sylvia

1996 «La evaluación de los centros escolares». México: OEI. Fecha de consulta: 15/3/2014. <<http://www.oei.es/calidad2/sylvia.htm>>.

SEN, Amartya

2005 *The Argumentative Indian: Writings on Indian History, Culture, and Identity*. Nueva York: Farrar, Straus and Giroux.

1995 *Inequality Reexamined*. Nueva York: Harvard University Press.

ST. CLAIR, Ralf

2012 «The Limits of Levels: Understanding the International Adult Literacy Surveys (IALS)». En: *International Review of Education*, N° 58, pp. 759-776.

STEVENS, Stanley S.

1946 «On the Theory of Scales of Measurement». En: *Science*, vol. 103, N° 2684, pp. 677-680.

STREET, Brian

1998 «New Literacies in Theory and Practice: What are the Implications for Language in Education?» En: *Linguistics and Education*, N°10, pp. 1-24.

TED

2013 «Ken Robinson: cómo escapar del valle de la muerte de la educación» [video]. Ted.com. Fecha de consulta: 15/3/2014. <http://www.ted.com/talks/lang/es/ken_robinson_how_to_escape_education_s_death_valley.html>.

2006 «Ken Robinson dice que las escuelas matan la creatividad» [video]. Ted.com. Fecha de consulta: 15/3/2014. <http://www.ted.com/talks/lang/es/ken_robinson_says_schools_kill_creativity.html>.

TIANA, Alejandro; Niels EGELUND y Morin THYGESEN

2004 «National Interpretations of the International Indicators». En: *European Educational Research Journal*, vol. 3, N° 1, pp. 213-245.

TORANZOS, Lilia

1996 «Evaluación y calidad». En: *Revista Iberoamericana de Educación*, N° 10, pp. 63-78.
Fecha de consulta: 15/3/2014. <<http://www.rieoei.org/oeivirt/rie10a03.pdf>>.

UNDP, PROGRAMA DE LAS NACIONES UNIDAS PARA EL DESARROLLO

2009 *Human Development Report 2009: Overcoming Barriers. Human Mobility and Development*. Nueva York y Basingstoke: UNDP / Palgrave Macmillan.

UNESCO

2009 *Aprovechar el poder y el potencial del aprendizaje y la educación de adultos para un futuro viable. Marco de acción de Belén*. Belén: Unesco.

2006 *Strong Foundations: Early Childhood Care and Education. Education for All, Global Monitoring Report 2007*. París: Unesco.

2004a *Education for All: The Quality Imperative. Education for All, Global Monitoring Report 2005*. París: Unesco.

2004b *Informe de seguimiento de la EPT en el mundo. Panorama regional*. París: Unesco.

2000 *Marco de acción de Dakar*. París: Unesco.

1999a *Clasificación Internacional Normalizada de Educación*. París: Unesco.

1999b *World Education Indicators [CD]*. París: Unesco.

1996 *La educación encierra un tesoro. Informe a la Unesco de la Comisión Internacional sobre la Educación para el siglo XXI*. París: Unesco.

1976 *International Standard Classification of Education*. París: Unesco.

1972 *A Study of Wastage at School*. París y Ginebra: Unesco.

1961 *Manual of Education Statistics*. París: Unesco.

1960 *Convención contra la discriminación en la educación*. París: Unesco.

UNESCO SANTIAGO

2013 *Situación educativa de América Latina y el Caribe: Hacia la educación de calidad para todos al 2015*. Santiago de Chile: Unesco.

2010 *Factores asociados al logro cognitivo de los estudiantes de América Latina y el Caribe*. Santiago de Chile: Unesco.

2008 *Situación educativa de América Latina y el Caribe: garantizando la educación de calidad para todos*. Santiago de Chile: Unesco. Fecha de consulta: 15/3/2014. <<http://unesdoc.unesco.org/images/0015/001528/152894s.pdf>>.

2007 *Educación de calidad para todos: un asunto de derechos humanos*. Santiago de Chile: Unesco. Fecha de consulta: 15/3/2014. <<http://unesdoc.unesco.org/images/0015/001502/150272s.pdf>>.

2004 *La conclusión universal de la educación primaria en América Latina: ¿Estamos realmente tan cerca?* Santiago de Chile: Unesco. Fecha de consulta: 15/3/2014. <<http://unesdoc.unesco.org/images/0013/001373/137330s.pdf>>.

- 2002a *Estudio cualitativo de escuelas con resultados destacables en siete países latinoamericanos*. Santiago de Chile: Unesco.
- 2002b *Proyecto regional de educación de América Latina y el Caribe*. Santiago de Chile: Unesco.
- 2001 *Situación educativa de América Latina y el Caribe 1980-2000*. Santiago de Chile: Unesco.
- 1998 *Primer estudio internacional comparativo sobre lenguaje, matemática y factores asociados en tercero y cuarto grado*. Santiago de Chile: Unesco.
- 1996 *Situación educativa de América Latina y el Caribe 1980-1994*. Santiago de Chile: Unesco.
- 1992 *Situación educativa de América Latina y el Caribe 1980-1989*. Santiago de Chile: Unesco.
- 1990 *Situación educativa de América Latina y el Caribe 1980-1987*. Santiago de Chile: Unesco.
- 1988 *Situación educativa de América Latina y el Caribe*. Santiago de Chile: Unesco.

UNESCO SANTIAGO y MÉXICO-SECRETARÍA DE EDUCACIÓN PÚBLICA

- 2008 *Balance de las misiones de diagnóstico de los sistemas de información educativa. Resumen ejecutivo*. Santiago de Chile y México DF: Unesco.

UNESCO-UIS

- 2014 «Post-2015 Education Indicators Technical Advisory Group of the EFA Steering Committee». En: *Towards Indicators for a Post-2015 Education Framework*. Montreal: Unesco-UIS.
- 2013a *Clasificación Internacional Normalizada de Educación. CINE 2011*. Montreal: Unesco-UIS.
- 2013b *International Standard Classification of Education. ISCED 2011*. Montreal: Unesco-UIS.
- 2010a «Assuring Data Quality at the Unesco Institute for Statistics». Presentación en: CCSA Conference on Data Quality for International Organisations. Helsinki, 6-7 mayo. Fecha de consulta: 15/3/2014. <<http://unstats.un.org/unsd/acsub/2010docs-CDQIO/Ses2-UNESCO.pdf>>.
- 2010b *Education Indicators. Technical Guidelines*. Montreal: Unesco.
- 2009a *Indicadores de la educación. Especificaciones técnicas*. Montreal: Unesco. Fecha de consulta: 15/3/2014. <<http://www.uis.unesco.org/Library/Documents/eiguide09-es.pdf>>.
- 2009b *The Next Generation of Literacy Statistics. Implementing the Literacy Assessment and Monitoring Programme (LAMP)*. Montreal: Unesco-UIS. Fecha de consulta: 15/3/2014. <<http://www.uis.unesco.org/Library/Documents/Tech1-eng.pdf>>.

UNESCO-UIS y OECD

- 2000 *Investing in Education: Analysis of the 1999 World Education Indicators*. París y Montreal: OECD / Unesco-UIS.

UNESCO-UIS y WORLD BANK

- 2003 *A Framework for Assessing the Quality of Education Statistics*. Washington, DC y Montreal: World Bank / Unesco-UIS. Fecha de consulta: 15/3/2014. <<http://unstats.un.org/unsd/dnss/docs-nqaf/WB-UNESCO-DQAF%20for%20education%20statistics.pdf>>.

UNICEF

- 2000 *Defining Quality in Education*. Nueva York: Unicef. Fecha de consulta: 15/3/2014. <<http://www.unicef.org/education/files/QualityEducation.PDF>>.

UNICEF y UNESCO

2007 *A Human-Rights Approach to Education for All*. Nueva York: Unicef / Unesco.

UNITED NATIONS DEVELOPMENT GROUP

2013 *A Million Voices: The World we Want. A Sustainable Future with Dignity for All*. Nueva York: United Nations Development Group.

UNSD, UNITED NATIONS STATISTIC DIVISION

2013 *Fundamental Principles of Official Statistics*. Nueva York: UNSD. Fecha de consulta: 15/3/2014. <<http://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx>>.

2008 *Principles and Recommendations for Population and Housing Censuses. Rev 2*. Nueva York: UN.

2005 *Principles Governing International Statistical Activities*. Nueva York: UNSD. Fecha de consulta: 15/3/2014. <https://unstats.un.org/unsd/methods/statorg/Principles_stat_activities/principles_stat_activities.htm>.

2004 *Manual de organización estadística: el funcionamiento y organización de una oficina de estadística*, 3ª ed. Nueva York: UNSD.

1997 *Principles and Recommendations for Population and Housing Censuses. Rev 1*. Nueva York: UN.

VAILLANT, Denise

2005 «Reformas educativas y rol de docentes». En: *Revista Prelac*, N° 1, pp. 38-51.

VEGAS, Emiliana y Chelsea COFFIN

2015 «Cuando el gasto en la educación importa: un análisis empírico de información internacional reciente». Documento de trabajo N° 574. Washington, DC: BID. Fecha de consulta: 5/3/2015. <http://publications.iadb.org/bitstream/handle/11319/6803/Cuando_el_gasto_en_la_educacion_importa_un_analisis_empirico_de_informacion_internacional_reciente.pdf?sequence=1>.

VINCENT, David

2009 «The End of Literacy: The Growth and Measurement of British Public Education Since the Early 19th Century». Documento de trabajo BWPI N° 67. Manchester: Brooks World Poverty Institute.

VON DAVIER, Mathias; Eugenio GONZALEZ y Robert J. MISLEVY

2009 «What are Plausible Values and Why are they Useful?». En: *IERI Monograph Series. Issues and Methodologies in Large-Scale Assessments*, vol 2. Wellington: IERI, pp. 9-36.

VYGOTSKY, Lev S.

1978 *Mind in Society: The Development of Higher Psychological Processes* (M. Cole, S. Scribner, E. Souberman y V. John-Steiner, eds.). Cambridge: Harvard University Press.

WAGNER, Daniel A.

2011 *Smaller, Quicker, Cheaper. Improving Learning Assessments for Developing Countries*. París: Unesco / IIEP.

2003 «Smaller, Quicker, Cheaper: Alternative Strategies for Literacy Assessment in the UN Literacy Decade». En: *International Journal of Educational Research*, vol. 39, N° 3, pp. 293-309.

WOLFF, Edward N.

2000 «Human Capital Investment and Economic Growth: Exploring the Cross-country Evidence». En: *Structural Change and Economic Dynamics*, vol. 11, N° 4, pp. 433-472.

WOLFF, Lawrence; Ernesto SCHIEFELBEIN y Paulina SCHIEFELBEIN

2002 *Primary Education in Latin America: The Unfinished Agenda*. Washington, DC: BID.

WORLD BANK

1980 «Education». Sector Policy Paper. Washington DC: World Bank.

WORLD BANK e INDEPENDENT EVALUATION GROUP

2006 *From Schooling Access to Learning Outcomes. An Unfinished Agenda. An Evaluation of World Bank Support to Primary Education*. Washington, DC: World Bank.

WORLD CONFERENCE ON EDUCATION FOR ALL

1990 *Meeting Basic Learning Needs: A Vision for the 1990s*. Nueva York: UNDP / Unesco / Unicef / World Bank.

BASES DE DATOS

Escale-Minedu <<http://escale.minedu.gob.pe>>

Siteal, Unesco-OEI <<http://www.siteal.iipe-oei.org>>

Unesco-UIS <<http://www.uis.unesco.org>>

