

USO Y ABUSO DE LAS

ESTADÍSTICAS

La distorsión en la percepción pública
de los problemas sociales y políticos

JOEL BEST



Título en inglés

*MORE DAMNED LIES AND STATISTICS:
How Numbers Confuse Public Issues*

Este libro es traducción de
MORE DAMNED LIES AND STATISTICS
How Numbers Confuse Public Issues

© Joel Best, 2004

© The Regents of the University of California, 2004
Published by arrangement with University of California Press

USO Y ABUSO DE LAS ESTADÍSTICAS
La Distorsión en la Percepción Pública de los Problemas Sociales y Políticos
© Editorial Cuatro Vientos, Santiago de Chile, 2009

Derechos reservados para todos los países de habla hispana.

Registro de Propiedad Intelectual N° 184.933
ISBN N° 978-956-242-105-8

Traducción: Daniela Huneeus
Revisión: Paulina Correa
Edición general: Francisco Huneeus
Diseño de portada: Mario Albarracín
Diagramación: Héctor Peña

Ninguna parte de esta publicación, incluido el diseño de la cubierta, puede ser reproducida, almacenada o transmitida en manera alguna ni por ningún medio, ya sea eléctrico, químico, mecánico, óptico, de grabación o de fotocopia, sin permiso previo, por escrito, del editor.

Editorial Cuatro Vientos
Maturana 19, CP 8340608, Santiago de Chile
Teléfonos: (56 2) 672 9226, (56 2) 695 4477
Fax: (56 2) 673 2153
email: editorial@cuatrovientos.cl
www.cuatrovientos.cl

USO Y ABUSO DE LAS ESTADÍSTICAS
*La Distorsión en la Percepción Pública de
los Problemas Sociales y Políticos*

JOEL BEST

Traducción

Daniela Huneeus



EDITORIAL CUATRO VIENTOS
WWW.CUATROVIENTOS.CL

CONTENIDOS

AGRADECIMIENTOS.....	xi
----------------------	----

Prefacio

LA GENTE TAMBIÉN CUENTA.....	xiii
------------------------------	------

Capítulo 1

CIFRAS QUE BRILLAN POR SU AUSENCIA.....	1
El poder de los ejemplos	4
Lo que no se puede calcular	8
Lo que no se contabiliza.....	14
Números olvidados.....	19
Cifras legendarias.....	20
¿Qué es lo que falta?.....	25

Capítulo 2

CIFRAS ENGAÑOSAS.....	27
Algunas complejidades de los números sencillos	29
Confundidos con los gráficos	43
Sobre la confusión	61

Capítulo 3

CIFRAS QUE ASUSTAN	63
Describir problemas sociales en términos atemorizantes.....	65
Riesgos	75
Externalidades y pesimismo pretencioso	88

Capítulo 4

CIFRAS ACREDITADAS	93
Hitos de la investigación científica	96
Ambigüedades en el sistema oficial de registros	105
La fragilidad de la acreditación	117

Capítulo 5

CIFRAS MÁGICAS.....	121
El número especial de trajes de baño.....	124
Cómo jugar con números mágicos dentro de las organizaciones.....	130
Colegios en tela de juicio.....	136
Prácticas policiales basadas en el perfil racial.....	143
¿Por qué es necesaria la magia?.....	149

Capítulo 6

CIFRAS CONTROVERSIALES.....	151
Ciencia y ciencia basura	157
Manipulación y selección cuidadosa.....	163
Cuántos judíos y musulmanes hay en EEUU.....	167
Resultados de la reforma al bienestar social.....	171
Disputas en torno a las cifras o presuposiciones cuestionables	176

Capítulo 7

ALFABETIZACIÓN ESTADÍSTICA.....	179
¿Quién debe alfabetizar estadísticamente?	182
Movimiento por la alfabetización estadística.....	187
ÍNDICE ANALÍTICO	191

Para Loy Bilderback

Agradecimientos

No planeé escribir este libro. Es la continuación de *Damned Lies and Statistics (DLS)*, que se publicó en 2001. Cuando terminé de escribir *DLS*, pensé que las estadísticas eran un tema superado, y pretendía trabajar en un proyecto completamente distinto. Además, soy profesor y los profesores no tienen muchas oportunidades para escribir continuaciones; nos sentimos afortunados si alguien quiere publicar, ni pensamos en leer, lo que alguna vez escribimos.

Sin embargo, casi tan pronto como apareció *DLS*, comencé a recibir e-mails de personas que habían leído el libro. Con frecuencia se referían a estadísticas evidentemente dudosas publicadas en la prensa. Entre mis favoritas: un columnista que advertía que fumar “mata a 1 de 5 estadounidenses cada año”; y una nota de un periódico británico que sugería que “40% de los hombres jóvenes tienen tan poca idea de cómo se abrocha un sostén, que corren el riesgo de herirse seriamente los dedos”. Otros indicaron temas que *DLS* no había tratado: algunos de los mensajes fueron enviados por docentes universitarios que se sentían frustrados por las dificultades para transmitir ciertos contenidos de sus clases.

Además, comencé a ser invitado para hacer presentaciones a grupos o escribir sobre estadísticas; a menudo me solicitaban charlas acerca de asuntos que no me resultaban familiares. Al estudiar

nuevos temas, solían surgir otros que hubiera querido tratar en *DLS*.

Cuando Naomi Schneider, mi editora en University of California Press, me preguntó si quería escribir la continuación de *DLS*, acepté. Ya estaba convencido de que tenía suficientes ideas para otro libro y parecía haber bastantes personas interesadas en el asunto. Temo haber perdido la pista de las fuentes de algunas ideas, pero por lo menos puedo agradecer a aquellos que hicieron sugerencias que, de una forma u otra, fueron incorporadas en este libro, así como dar las gracias a quienes leyeron y comentaron partes del manuscrito. Estas personas, aparte de Naomi, son David Altheide, Ronet Bachman, Joan Best, George Bizer, Barbara Costello, Michael Gallagher, Linda Gottfredson, Larry Griffith, Henry Hipkens, Jim Holstein, Philip Jenkins, Vivian Klaff, Carl Klockars (fallecido), Kathe Lowney, Katherine C. MacKinnon, Michael J. McFadden, Eric Rise, Naomi B. Robbins, Milo Schield y —me temo— otros cuyos nombres debieron trasapelarse. Quiero agradecer especialmente a Vicky Baynes por ayudarme en el misterioso proceso de transformar los gráficos en archivos computacionales. Por supuesto, debo agradecer a estas personas por haberme ayudado, pero no se les debe responsabilizar por mis interpretaciones. Gracias a todos. Espero que este nuevo libro sea de su agrado.

Prefacio

LA GENTE TAMBIÉN CUENTA

El almuerzo se llevó a cabo en un prominente Centro de Estudios muy conservador. Los asistentes eran más o menos conocidos; había leído algunos de sus libros y artículos e incluso había visto algunos en entrevistas por televisión. Escucharon mi exposición acerca de las malas estadísticas y admitieron que el problema era serio. Sólo tenían una crítica: había pasado por alto el asunto de la ideología. Según ellos, las malas estadísticas eran casi siempre promovidas por los progresistas*.

Dos meses antes había sido entrevistado en un programa de conversación de una radio progresista (¡existen!). El anfitrión también pensaba que era un buen momento para poner al descubierto las malas estadísticas, especialmente aquellas dadas a conocer por los conservadores.

Cuando hablo con las personas sobre estadísticas, tengo la impresión de que están bastante dispuestas a criticar las estadísticas dudosas, siempre y cuando las cifras provengan de aquellos con quienes discrepan. Los conservadores creen que las estadísticas presentadas por los progresistas contienen errores graves, tal como los progresistas ansían denunciar las débiles cifras de los conservadores. Cuando los

* En EEUU, como todo el mundo sabe, las tendencias ideológicas referentes a la política se agrupan, a grandes rasgos, en dos vertientes: conservadores, agrupados en el partido Republicano; y los más liberales, con ideas con matices de izquierda, generalmente agrupados en el partido Demócrata y que en esta traducción llamaremos progresistas (N. del E.).

conservadores (o los progresistas) me preguntan cómo detectar las malas estadísticas, imagino que lo que quieren oír es: “Sospeche de todas las cifras promovidas por aquellos con quienes usted no está de acuerdo”. Todos parecen insistir en que las cifras del otro son pésimas (pero, por supuesto, las propias están bien o, por lo menos, lo suficientemente bien). A las personas les agrada ver ejemplos de malas estadísticas de sus oponentes, pero no les interesa criticar sus propias cifras porque la gente se confundiría: criticar las propias estadísticas podría llevar a cuestionar el argumento principal, por lo que es mejor centrarse en los errores del otro y minimizar los míos.

Desgraciadamente, no creo que ningún grupo, facción o ideología en particular, tenga la exclusividad de razonar de manera errada en lo que a estadísticas se refiere. De hecho, al elegir los ejemplos para ilustrar los capítulos de este libro, intenté identificar un grupo de infractores lo más amplio posible. Mi objetivo no es convencerlo a usted de que no se puede confiar en esos tipos (mal que mal, es muy posible que ya esté convencido de ello); lo que pretendo es que concluya la lectura de este libro con la impresión de que *todas las cifras y números —de otros o propios— deben ser manejados con sumo cuidado*.

Esto es engañoso, porque tendemos a asumir que las estadísticas son hechos, pequeñas muestras de verdad que salen a la luz, casi como les sucede a los coleccionistas de piedras¹. Después de todo, una estadística es un número, y las cifras parecen ser pruebas sólidas y objetivas de que alguien debió contabilizar algo. Pero ése es el punto: las personas cuentan. Por cada número que encontramos, alguien tuvo que contar algo. Más que imaginar que las estadísticas son como las piedras, nos convendría imaginarlas como joyas. Las piedras preciosas se

¹ Sostuve algo similar en *Damned Lies and Statistics: Untangling Numbers from the Media, Politicians, and Activists* (Berkeley: University of California Press, 2001), p. 27. Aquí doy ejemplos diferentes, organizados de diversas formas, e intenté minimizar las coincidencias, pero el enfoque subyacente es el mismo. Para entender el presente libro, no es necesario haber leído el primero; sin embargo, ambos se complementan.

pueden encontrar en la naturaleza, pero las joyas tienen que ser creadas por personas. Las joyas deben seleccionarse, cortarse, pulirse y engastarse de manera que luzcan desde distintos ángulos. La gente crea estadísticas de una manera similar: elige qué cosa contabilizar, cómo hacerlo, cuáles resultados compartir y qué palabras utilizar para describir e interpretar esos números. Las cifras no existen independientemente de las personas; comprender las cifras exige saber quién contó qué, qué motivó a realizar ese cálculo y cómo lo hizo.

Todas las estadísticas son producto de la actividad social, del proceso que los sociólogos llaman *construcción social*. Si bien esto puede parecer demasiado obvio, tiende a olvidarse o ignorarse cuando pensamos en la estadística, y especialmente cuando la enseñamos. Una idea errónea, reforzada durante la enseñanza media y la universidad, es que la estadística se considera una rama de la matemática que comienza con la introducción a la teoría de las probabilidades como base del pensamiento estadístico; sobre este fundamento se construye una estructura de mediciones estadísticas cada vez más sofisticadas. A los estudiantes se les enseña la lógica subyacente a cada medición, la fórmula utilizada para calcularla, los comandos de los softwares que pueden extraerla del computador y algunas pautas para interpretar los resultados de esos cálculos. Estas son lecciones bastante complejas: pocos estudiantes comprenden intuitivamente algo más que la estadística básica, y la enseñanza suele centrarse casi exclusivamente en aclarar las complejidades computacionales.

Por lo tanto, la enseñanza de la estadística tiende a minimizar cómo de hecho surgen las cifras sobre aspectos de la vida real. Todas las estadísticas son producto de opciones y compromisos de personas, las que inevitablemente dan forma, limitan y distorsionan el resultado. Los profesores de estadística suelen desestimar esto. Tal como los conservadores del centro de estudios creían que las malas estadísticas eran obra de los malévolos progresistas, los profesores de estadística podrían señalar que es posible que los cálculos o presentaciones de resultados

estadísticos sean tendenciosos o sesgados (es decir, diseñados intencionalmente para engañar). Asimismo, existen innumerables textos que distinguen entre estadísticas y mentiras: *Cómo Mentir con Estadísticas* (además de *Cómo Mentir con Gráficos*, *Cómo Mentir con Mapas*, etc.); *Cómo Distinguir a los Mentirosos de los Estadísticos*; *Cómo Mienten los Números*; e incluso mi libro anterior, *Damned Lies and Statistics*². Se podría concluir que las estadísticas son puras, salvo que desafortunadamente sean contaminadas por las malas intenciones de gente deshonestas.

Quizás sea necesario dejar a un lado el mundo real para que los estudiantes adquieran una manera más de avanzada de pensar sobre las estadísticas. Si se trata de preparar a las personas para pensar críticamente respecto a las cifras que leen en la prensa o escuchan en la televisión, pareciera que las advertencias a tener presente la parcialidad no sirven demasiado. Las estadísticas cumplen un rol importante en los debates del mundo real sobre los problemas y las políticas sociales; las cifras se transforman en evidencia clave para cuestionar los puntos de vista del oponente y promover las opiniones propias. Dado que la gente deliberadamente presenta números distorsionados, o incluso falsos, no podemos descartar la existencia de la parcialidad. Pero tampoco podemos creer que las cifras sencillamente son números verdaderos, presentados por personas sinceras y bienintencionadas (las que naturalmente están de acuerdo con nosotros), o estadísticas falsas dadas a

² Darrell Huff, *How to Lie with Statistics* (New York: Norton, 1954); Gerald E. Jones, *How to Lie with Charts* (San Francisco: Sybex, 1995); Mark Monmonier, *How to Lie with Maps* (Chicago: University of Chicago Press, 1991); Robert Hooke, *How to Tell the Liars from the Statisticians* (New York: Marcel Dekker, 1983); Richard P. Runyon, *How Numbers Lie* (Lexington, Mass.: Lewis, 1981); Best, *Damned Lies and Statistics*; también: Tukufu Zuberi, *Thicker than Blood: How Racial Statistics Lie* (Minneapolis: University of Minnesota Press, 2001). Otros textos tienen capítulos sobre el tema: véase, por ejemplo, "Statistics and Damned Lies", en A.K. Dewdney, *200% of Nothing* (New York: Wiley, 1993), pp. 23-42. A lo anterior se suma un libro de 300 páginas que pretende ayudar a los profesores de estadística a acercarse a sus alumnos y que contiene un capítulo de 17 páginas titulado "Lying with Statistics"; véase Andrew Gelman y Deborah Nolan, *Teaching Statistics: A Bag of Tricks* (New York: Oxford University Press, 2002), pp. 147-163.

conocer a sabiendas por individuos poco honestos (que, desde luego, están en el bando contrario).

El entusiasmo exacerbado es por lo menos tan común como el sesgo deliberado para explicar el porqué la gente suele divulgar las malas estadísticas. Rara vez los números van primero. La gente no suele comenzar creando cuidadosamente fragmentos de información estadística para luego deducir lo que debiera pensarse a partir de ella. Lo más habitual es que el punto de inicio esté conformado por los intereses o preocupaciones de las personas, a partir de lo cual encuentran, o quizás descubren activamente, información estadística relevante para sus fines. Cuando estos números apoyan lo que los individuos creen —esperan o temen— que sea verdadero, es muy fácil que adopten las cifras, pasen por alto o minimicen sus limitaciones, consideren que los números son en primer lugar fascinantes, luego de suma importancia y finalmente acreditados. Muy pronto la gente comienza a hacer circular estas cifras, ahora tan relevantes, y se enfurece si sus estadísticas son puestas en duda. No es necesario mentir de manera intencional a los demás o a uno mismo. *Cuando uno se enfrenta a una cifra atractiva, basta con bajar la guardia crítica, y el impulso hará el resto.*

Al reflexionar sobre las estadísticas, es necesario mantener los estándares críticos. Algunas personas son muy hábiles para hacer esto, siempre y cuando estén examinando los números de sus contrarios. Frente a las cifras propias es mucho más difícil mantener una postura crítica. A final de cuentas, las cifras que manejamos apoyan lo que creemos es verdad. Seguramente, cualquier falla mínima o defecto que pudieran tener debe carecer de importancia. Al menos eso es lo que nos decimos cuando justificamos el doble estándar para juzgar nuestras estadísticas y las de los demás.

Este libro busca promover lo que podríamos llamar un estándar único para la crítica estadística. Sostiene que *debemos reconocer que todas las cifras son productos sociales y que no podemos comprender una estadística a menos que sepamos algo sobre su proceso de creación.* Además, plantea

que *todas las estadísticas son imperfectas y que debemos reconocer sus fallas y limitaciones*. Todo esto es válido estemos o no de acuerdo con quienes dan a conocer las cifras. Debemos pensar críticamente tanto de los números de los demás como de los propios.

Debo confesar que al escribir este libro realicé poca investigación original. Tomé la mayoría de los ejemplos del trabajo de otros analistas, cientistas sociales y periodistas. Mi objetivo al escribir sobre las estadísticas malas es mostrar cómo surgen y se difunden estas cifras. Así como no creo que ésta sea labor de una determinada facción política, tampoco quiero sugerir que se deba culpar sólo a un segmento de la sociedad, tal como los medios de comunicación masiva. Es usual que éstos hagan circular cifras malas, pero también lo hacen los activistas, empresas, funcionarios de gobierno e incluso científicos; de hecho, todas estas personas suelen ser la fuente de las estadísticas que aparecen en los medios. Y debemos recordar que los problemas vinculados con las estadísticas malas suelen salir a la luz gracias al esfuerzo crítico de algunos periodistas sagaces y científicos que reflexionan sobre las cifras, descubren sus fallas y llaman la atención respecto a ellas. Una mirada a mis fuentes revelará que el pensamiento crítico, al igual que las malas estadísticas, puede encontrarse en muchas partes.

Los capítulos de este libro exploran algunos problemas comunes al abordar las estadísticas sociales. Los títulos de los capítulos se refieren a distintos tipos de cifras: las que faltan, las que confunden, etc. Tal como yo los utilizo, estos términos no tienen un significado matemático formal; sólo son encabezados para organizar el debate.

Así, el Capítulo 1 trata sobre lo que yo llamo *cifras que brillan por su ausencia*: estadísticas que podrían ser importantes para la discusión sobre problemas sociales, pero que por algún motivo no surgen en el marco de dichos temas. En este capítulo se identifican varios tipos de cifras que faltan y se intenta explicar esa ausencia.

El Capítulo 2 se refiere a *cifras engañosas*: aquellas que complican problemas básicos que dificultan la comprensión de muchas estadísticas y gráficos simples.

El Capítulo 3 se centra en las *cifras que asustan*: estadísticas sobre riesgos y otras amenazas.

Los Capítulos 4, 5 y 6 examinan la relación entre autoridad y estadísticas.

El Capítulo 4 habla de *cifras acreditadas* y estadísticas que creemos suficientemente buenas como para no cuestionarlas, ya que, por ejemplo, son producto de la investigación científica o se refieren a datos recogidos por el gobierno. Este capítulo plantea que incluso las mejores estadísticas deben considerarse con precaución, es decir, incluso los datos recopilados por expertos pueden llevar a confusión.

El Capítulo 5 aborda lo que yo denomino *cifras mágicas*: esfuerzos para resolver temas mediante estadísticas, como si los números constituyeran una manera de destilar la realidad en hechos puros e indiscutibles.

El Capítulo 6 se centra en las *cifras controversiales*: casos de guerras estadísticas y duelos de datos en que los oponentes se atacan entre sí con números contradictorios.

Por último, en el Capítulo 7 investigo las posibilidades de *alfabetizar estadísticamente* a las personas para mejorar la comprensión masiva de las cifras y enseñar a ser consumidores de estadísticas más reflexivos y críticos.

La lección respecto a que las personas cuentan —que no se trata de “encontrar” datos estadísticos, sino que éstos son creados por alguien— es tanto una advertencia como una promesa. La advertencia es que debemos ser precavidos: *a menos que abordemos las estadísticas de manera crítica, corremos el riesgo de no comprender el mundo que nos rodea*. Pero también hay una promesa: *no es necesario estar a merced de las cifras, es posible pensar en ellas críticamente y apreciar tanto sus fortalezas como debilidades*.

Capítulo 1

CIFRAS QUE BRILLAN POR SU AUSENCIA

El 5 de marzo de 2001, el conductor del noticiario de la CBS, Dan Rather, comenzó la emisión vespertina diciendo: “Los tiroteos en las escuelas de este país se han convertido en una epidemia”. Ese mismo día, un estudiante en Santee, California, asesinó a balazos a dos alumnos e hirió a trece. La cobertura mediática vinculó este episodio a una tendencia alarmante. Entre diciembre de 1997 y mayo de 1998 hubo tres tiroteos en escuelas que recibieron gran atención por parte de los medios: West Paducah, Kentucky (tres muertos y cinco heridos); Jonesboro, Arkansas (cinco muertos y diez heridos); y Springfield, Oregon (después de asesinar a sus padres en su casa, el victimario se dirigió a la escuela, mató a dos personas e hirió a 21). En la primavera siguiente ocurrió la tragedia de Columbine High School en Littleton, Colorado, en que dos estudiantes asesinaron a doce compañeros y a un profesor antes de suicidarse¹. Así las cosas, ¿quién podría poner en duda la afirmación de Rather de que estábamos frente a una epidemia?

¹ Rather, citado en Lynnell Hancock, “The School Shootings: Why Context Counts”, *Columbia Journalism Review* 40 (May 2001): 76. Existen muchas publicaciones acerca de estos episodios, y particularmente sobre el caso de Columbine. Véase, por ejemplo, Daniel M. Filler, “Random Violence and the Transformation of the Juvenile Justice Debate”, *Virginia Law Review* 86 (2000): 1095-1125; y Gary Kleck, “There Are No Lessons to Be Learned from Littleton”, *Criminal Justice Ethics* 18 (Winter 1999): 2, 61-63.

El término “epidemia” sugiere un fenómeno diseminado que crece, algo que se extiende. ¿Estaban en alza los tiroteos en las escuelas? Sorprendentemente, existían evidencias que indicaban todo lo contrario:

- Ya que los tiroteos en las escuelas son delitos violentos, es conveniente examinar las tendencias de criminalidad documentadas por el FBI. Los Informes Estandarizados de Delincuencia (Uniform Crime Reports), registro que lleva el FBI de los delitos denunciados a la policía, indican que la tasa general de éstos, incluidos los más violentos, como homicidios, asaltos y amenazas graves de acometimiento, cayó durante los años 90.
- Asimismo, la Encuesta Nacional de Victimización por Delincuencia (National Crime Victimization Survey), que consulta a los entrevistados si alguien en el hogar ha sido víctima de un delito, revela que las tasas de victimización cayeron durante la misma década; en particular, disminuyeron las denuncias de adolescentes afectados por delitos violentos.
- También se redujeron otros indicadores de violencia escolar. La Encuesta de Conductas de Riesgo en Jóvenes (Youth Risk Behavior Survey), llevada a cabo por los Centros de Control y Prevención de Enfermedades de EEUU (U.S. Centers for Disease Control and Prevention), mostró, durante la década de los 90, una sostenida caída del porcentaje de estudiantes de enseñanza media involucrados en peleas o tenencia de armas al interior de los establecimientos educacionales.
- Finalmente, cuando los investigadores del Centro Nacional de Seguridad Escolar (National School Safety Center) revisaron los informes de los medios de comunicación relativos a los años escolares desde 1992-1993 hasta 2000-2001, identificaron 321 muertes violentas ocurridas en las escuelas. Pero no todos estos incidentes involucraban violencia entre estudiantes; por ejemplo,

incluían 16 muertes accidentales y 56 suicidios, así como episodios que no contaban con la participación de alumnos, como el crimen de una profesora asesinada por su esposo, del que estaba separada y que luego se suicidó, y una persona que no era estudiante que fue asesinada en el patio de un colegio durante el fin de semana. Sin embargo, incluso si consideramos las 321 muertes, el promedio disminuye de 48 fallecimientos violentos anuales durante los años escolares 1992-1993 hasta 1996-1997, a 32 al año desde 1997-1998 hasta 2000-2001. Si eliminamos las muertes accidentales y los suicidios, la tendencia a la baja aumenta, con el promedio anual disminuyendo de 31 muertes en el período inicial a 24 en el último (que incluía todos los incidentes con alta cobertura mediática mencionados anteriormente). Si las muertes violentas son trágicas, también son escasas. Decenas de millones de niños van al colegio; por cada millón de estudiantes, hay menos de una muerte violenta anual en las escuelas.

En otras palabras, existía abundante evidencia estadística que cuestionaba la afirmación de que el país estaba sufriendo una súbita epidemia de tiroteos en las escuelas. Los Informes Estandarizados de Delincuencia del FBI y la Encuesta Nacional de Victimización por Delincuencia son recursos comunes de los periodistas; por ello sorprende que, al cubrir los tiroteos en las escuelas, los medios no incorporen datos de estas fuentes².

² Existen importantes recopilaciones de datos sobre violencia escolar, entre las que se incluyen: Margaret Small y Kellie Dressler Tetrick, "School Violence: An Overview", *Juvenile Justice* (U.S. Office of Juvenile Justice and Delinquency Prevention) 8 (June 2001): 3-12; Phillip Kaufman et al., *Indicators of School Crime and Safety: 2001* (Washington, DC: U.S. Departments of Education and Justice, 2001); Kim Brooks, Vincent Schiraldi y Jason Ziedenberg, *School House Hype: Two Years Later* (Washington, DC: Justice Policy Institute, 1999); National School Safety Center, *School Associated Violent Deaths*, 2001, disponible en www.nssc1.org; y Mark Anderson et al., "School-Associated Violent Deaths in the United States, 1994-1999", *Journal of the American Medical Association* 286 (December 5, 2001): 2695-2702. Para información acerca de los estándares de los periodistas para evaluar las estadísticas sobre delitos,

Aunque dé la impresión que las estadísticas están presente en todos los debates sobre temas sociales, en algunos casos —como la cobertura periodística de los tiroteos en las escuelas— se ignoran datos importantes y disponibles. Podríamos considerar que estas son las *cifras que brillan por su ausencia*. Este capítulo revisa los diversos motivos que hay detrás de estas cifras faltantes, los que incluyen ejemplos abrumadores, conceptos incalculables, fenómenos no contabilizados, números olvidados y cifras legendarias. Además, nos preguntaremos por qué razón las estadísticas potencialmente importantes no están presente en determinados debates públicos, e intentaremos evaluar los efectos de esa ausencia.

EL PODER DE LOS EJEMPLOS

¿Por qué faltan cifras en algunos debates sobre problemas y políticas sociales? Una posible respuesta es que un ejemplo potente puede anular la discusión sobre un tema. Los tiroteos de 1999 en Columbine High School vienen al caso. El alto número de muertos aseguraba que Columbine sería una historia noticiosa importante. Además, la ubicación de la escuela en un suburbio facilitaba el acceso de los periodistas a la escena. Dado que evacuar a los estudiantes y salvar el edificio tomó algunas horas, la prensa tuvo tiempo suficiente para llegar y hacer filmaciones que podrían transmitirse en las semanas siguientes para ilustrar historias relacionadas. Un crimen horrible ocurrido en una próspera comunidad suburbana hacía de la historia algo especialmente atemorizante: si este tiroteo sucedía en una escuela, es seguro que delitos como éstos podrían repetirse en cualquier parte. Por otro lado, la tragedia de Columbine se dio en un contexto de fuerte competencia entre canales de televisión por cable que

véase Kurt Silver, *Understanding Crime Statistics: A Reporter's Guide* (n.p.: Investigative Reporters and Editors Inc., 2000).

transmiten las 24 horas. Debido a ello, la decisión de cubrir en vivo algunos funerales y destinar tiempo a extensas discusiones acerca del suceso y sus consecuencias, ayudó a mantener la historia en el tapete durante semanas.

Para los medios, un suceso dramático puede ser más que un tema para el noticiero, porque los periodistas han desarrollado la habilidad de buscar la significación mayor de un evento, de modo de describirlo como un ejemplo de un problema más extendido. Así, cuando se vincula las historias de tiroteos ocurridos en las escuelas en 1997-1998 con la matanza de Columbine, ésta se vuelve un caso perfecto para ilustrar la violencia escolar. Entonces, y tal como planteaban los comentaristas, si estábamos frente a un problema mayor, éste debía ser reflejo de condiciones sociales subyacentes, es decir, los tiroteos en las escuelas debían considerarse una epidemia con causas identificables. Los periodistas han identificado esas tendencias delictivas al menos desde el siglo XIX; el mismo período en que los criminólogos han concluido que esto no responde a patrones de conducta criminal sino más bien de cobertura periodística. Toda la evidencia estadística disponible sugería que entre el inicio y el término de la década de los 90, la violencia escolar había declinado; por lo tanto, no existía tal ola de tiroteos escolares. Sin embargo, las poderosas imágenes de Columbine dieron por tierra con esa evidencia. Un ejemplo espantoso era “prueba” de que los tiroteos escolares eran en efecto una epidemia.

Pero los ejemplos más convincentes ni siquiera tienen que ser verídicos. Las historias denominadas “mitos urbanos” también moldean nuestro pensamiento sobre los problemas sociales. Los mitos urbanos suelen difundirse a través de canales informales: antiguamente, por transmisión oral; ahora, por medio de faxes y mensajes de correo electrónico. Sin embargo, su característica central se mantiene: debe ser una buena historia, tan buena como para que la gente la recuerde y comunique. Los mitos prosperan porque provocan miedo, desagrado u otras emociones intensas que vuelven memorable y repetible el hecho

al cual aluden³. Muy a menudo los mitos urbanos tienen gran actualidad: cuando en los medios aparece el secuestro de niños, contamos historias sobre raptos en los centros comerciales; si el foco de atención periodística se posa en las pandillas, corremos la voz sobre sus letales ritos de iniciación. De manera muy parecida a como lo hacen los dramáticos reportajes exhibidos por los medios, esas historias influyen en nuestra forma de pensar acerca de los problemas sociales.

El poder de los ejemplos es ampliamente reconocido. Es muy probable que un periodista que prepara una crónica sobre cualquier tema de relevancia social —p.ej., la escasez de viviendas— ilustre el problema con el caso de una persona en particular que no tiene donde vivir. Los periodistas (y sus editores) prefieren ejemplos atractivos y convincentes que generen interés en la audiencia. Quienes defienden determinadas políticas sociales también han aprendido a ayudar a los periodistas, sugiriéndoles ejemplos que pueden ser utilizados para ilustrar ciertos puntos específicos.

De ese modo, quienes promueven el aumento de servicios destinados a los que carecen de vivienda podrían mostrar a una familia sin casa, quizás una mujer con hijos pequeños cuyo esposo está cesante. Como contrapartida, es probable que los políticos que buscan internar a discapacitados mentales sin casa, usen como ejemplo a un individuo trastornado y violento que parece poner en peligro a los transeúntes⁴. La selección de los ejemplos comunica el sentido de la naturaleza de un determinado problema social.

El problema con los ejemplos —ya sea que provengan de sucesos dramáticos, mitos urbanos o la decisión estratégica de periodistas o defensores de ciertas ideas— es que su representatividad es improbable. Los ejemplos convencen cuando tienen poder emocional, cuando

³ Chip Heath, Chris Bell y Emily Sternberg, “Emotional Selection in Memes: The Case of Urban Legends”, *Journal of Personality and Social Psychology* 81 (2001): 1028-1041.

⁴ Cynthia J. Bogard, *Seasons Such as These* (Hawthorne, NY: Aldine de Gruyter, 2003).

asustan o incomodan. Pero los ejemplos más atípicos suelen distorsionar nuestra comprensión de un problema social; cuando nos concentramos en la excepción dramática, tendemos a pasar por alto los casos más comunes o típicos. Los demócratas solían quejarse de la afición del presidente republicano Ronald Reagan a repetir la historia de una “reina del subsidio” que supuestamente había logrado reunir decenas de cheques de la asistencia social por medio de identidades falsas⁵. Al utilizar este tipo de ejemplos para tipificar los fraudes a los servicios sociales, se está sugiriendo que las personas que reciben los subsidios no son merecedoras de éstos o que en realidad no necesitan la ayuda pública. Quienes defendían las prestaciones sociales solían oponer datos estadísticos a las anécdotas de Reagan, demostrando con evidencias que los receptores merecían tales prestaciones (p.ej., exponiendo el bajo porcentaje de personas sanas sin hijos pequeños que recibían beneficios) o que las condenas por fraude eran relativamente escasas⁶. El riesgo está en que el ejemplo poderoso pero atípico —la familia sin casa o la reina del subsidio— puede pervertir nuestra visión sobre un problema, reduciendo una compleja situación social a una fábula simplista y melodramática.

En consecuencia, las estadísticas ofrecen una manera de verificar nuestros ejemplos. Si la investigación acerca de las personas sin casa muestra pocas familias como la mencionada (o pocos individuos que amenazan con violencia), o si los estudios sobre los beneficiarios demuestran que el uso fraudulento de identidades falsas es poco común,

⁵ Compárese Kiron K. Skinner, Annelise Anderson y Martin Anderson, *Reagan, in His Own Hand* (New York: Free Press, 2001), pp. 241, 459; y Tip O'Neill, *Man of the House* (New York: Random House, 1987), pp. 347-348.

⁶ En respuesta a esto, quienes critican la asistencia social proporcionaban sus propias cifras, sugiriendo que los abusos eran muy extendidos. Un análisis planteaba: “Una fuente importante de las variaciones en los cálculos (de uso fraudulento de los subsidios)... ha sido la utilización —por parte de quienes calculan— de múltiples definiciones de lo que se entiende por falta” (John A. Gardiner y Theodore R. Lyman, *The Fraud Control Game* [Bloomington: Indiana University Press, 1984], p. 2).

entonces debemos reconocer el efecto distorsionador de los ejemplos atípicos y darnos cuenta de que la ausencia de cifras puede dañar nuestra habilidad de comprender las reales dimensiones del problema.

LO QUE NO SE PUEDE CALCULAR

En ocasiones, los números faltan porque los fenómenos son muy difíciles de contabilizar. Consideremos otra ola delictiva. Durante el verano de 2002, la atención del público se centró en el secuestro de niños. Primero, se enfocó en el caso de una adolescente que fue raptada una noche desde su dormitorio, un clásico ejemplo dramático de un delito terrible que podría sucederle a cualquiera. Pasaban semanas sin noticias de la niña, y tanto la búsqueda como la cobertura mediática continuaron. De a poco, los informes de otros casos de niños secuestrados o asesinados, que no estaban relacionados con el primer evento, empezaron a ser vinculados con éste, lo que llevó a los medios a hablar de una epidemia de raptos.

Sin embargo, el tema tenía antecedentes. Veinte años antes, ya hubo quienes llamaron la atención sobre el problema de los niños desaparecidos, juntando ejemplos aterrizantes con grandes cálculos estadísticos. Una afirmación bastante extendida sostenía que casi dos millones de niños desaparecían cada año, incluyendo cincuenta mil que eran secuestrados por extraños. Más tarde, periodistas y científicos sociales plantearon que esas cifras eran irracionalmente altas. Como resultado, en 2002, algunos periodistas cuestionaron la afirmación de que se estaba frente a una epidemia de raptos; de hecho, argumentaron que el FBI había investigado más casos de secuestro el año anterior, lo que sugería que esos delitos se estaban tornando menos comunes⁷.

⁷ Para la década de los 80, véase Joel Best, *Threatened Children* (Chicago: University of Chicago Press, 1990). Respecto a 2002, véase Donna Leinwand, "Kidnapping Problem 'Impossible' to Quantify", *USA Today*, August 15, 2002, p. 3A.

Ambas afirmaciones —que había una epidemia de raptos y que éstos estaban declinando— se basaban en evidencias muy débiles. Las estadísticas sobre niños desaparecidos nunca serán precisas, pues su número es muy difícil de contabilizar. Además, existen problemas de definición:

- ¿Qué entendemos por “niño”? Es decir, ¿cuál es la edad límite superior a considerar?
- ¿Qué entendemos por “desaparecido”? ¿Cuánto tiempo debe estar desaparecido un niño para ser contabilizado como tal: algunos minutos, un día, 72 horas?
- ¿Qué tipo de ausencias deben contabilizarse? ¿Vagabundear y perderse? ¿Escaparse? ¿Ser llevado por un pariente durante una riña familiar? ¿Se considera “desaparecido” a un niño que se encuentra en un lugar conocido con el progenitor que no detenta la custodia?

Es necesario que las personas lleguen a un acuerdo sobre qué contabilizar antes de hacerlo, pero no todo el mundo concuerda sobre las respuestas a estas preguntas. Obviamente, las respuestas escogidas afectarán las cifras; utilizar una definición amplia implica que se contabilizarán más niños.

Un segundo grupo de problemas se refiere a las denuncias. Es probable que los progenitores de un niño desaparecido reporten su ausencia a la agencia policial más cercana, usualmente la comisaría o departamento de policía. Pero esas autoridades pueden responder de diversas maneras. En EEUU, algunos estados exigen que esas instancias reenvíen todas las denuncias de niños desaparecidos a un centro de referencia estatal que supuestamente ha de contactar a todos los departamentos de seguridad del estado para facilitar la búsqueda. Los centros de referencia —y algunos departamentos— pueden notificar al Centro Nacional de Información Criminal (National Crime Information

Center), rama del FBI que reúne las denuncias de personas desaparecidas. Algunas denuncias también llegan al Centro Nacional para Menores Desaparecidos y Explotados (National Center for Missing and Exploited Children), grupo que cuenta con financiamiento federal y es conocido por dar a conocer las fotografías de niños desaparecidos, o al FBI (que reclama competencia en algunas, pero no la mayoría, de las denuncias). Además, las autoridades de una misma jurisdicción no manejan de la misma forma todas las denuncias de menores desaparecidos: el caso de un niño de 6 años arrastrado al interior de un vehículo extraño ciertamente será tratado de manera diferente a la denuncia de un adolescente de 16 años que huyó de casa. Podemos sospechar que las políticas de los diversos organismos variarán de manera considerable. Es probable que tanto la jurisdicción en la que desapareció un niño como las características del caso, influyan en si una denuncia particular llega o no a los registros de los diversos organismos involucrados.

Por lo tanto, es muy difícil realizar comparaciones convincentes de las cifras de menores desaparecidos, ya sea de tiempo en tiempo o de un lugar a otro. Los periodistas que advirtieron que en 2002 el FBI tenía menos denuncias de niños secuestrados que en 2001, y que entonces concluyeron que el problema estaba disminuyendo, asumieron de manera errada que los archivos del FBI eran más completos y autorizados de lo que se pensaba. Algunas situaciones —como la de los niños desaparecidos— son muy difíciles de contabilizar, lo que debería aumentar nuestro escepticismo en relación con la exactitud de las estadísticas que pretenden describir tales sucesos.

Esas dificultades pueden generar problemas especiales cuando las personas buscan comparar cosas que son relativamente fáciles de medir con aquellas que son menos cuantificables. Considérese el método de análisis de costo-beneficio como base de la toma de decisiones⁸. En

⁸ Para una introducción a este enfoque, véase Anthony E. Boardman et al., *Cost-Benefit Analysis: Concepts and Practices* (Upper Saddle River, NJ: Prentice Hall, 1996).

principio, parece sencillo: calcular los costos esperados y el valor de los potenciales beneficios para diferentes cursos de acción, y elegir la opción que promete el mejor resultado. Sin embargo, uno de los problemas es que algunos costos y beneficios son más fáciles de calcular que otros. Una adolescente que intenta decidir si va al cine o cuida niños, probablemente pueda asignar un valor monetario bastante exacto a estas opciones —el costo de la entrada al cine y el refrigerio versus sus ingresos al cuidar niños—, pero incluso así la decisión dependerá de factores adicionales asociados con la felicidad: ¿me sentiré más feliz pasando la noche con mis amigos en el cine, o prefiero ganar dinero que me permitirá un mayor beneficio más adelante?

Quando hablamos de gasto público, esos cálculos se vuelven muy complejos. ¿Debiéramos construir más autopistas o apoyar el transporte público? Éste rara vez se autofinancia: si el costo por viaje es demasiado alto, los usuarios dejan de utilizarlo; para conservar el número de usuarios es necesario subsidiar el precio de los pasajes. En ocasiones, los críticos al transporte público argumentan que subsidiar es negativo, que el transporte público es ineficiente, caro y, por consiguiente, no competitivo. Por su parte, los que lo defienden plantean que esta crítica desconoce muchos de los costos y beneficios relevantes. Mientras los usuarios asumen el costo de utilizar el transporte público cada vez que compran un pasaje, las formas en que costeamos el uso de las autopistas son menos obvias (p.ej., por medio del impuesto a los combustibles). Además, las autopistas suman costos a la calidad de vida que suelen permanecer ocultos, como es la mayor contaminación del aire, más accidentes de tránsito y ciudades que destinan grandes áreas a caminos y estacionamientos. Pero esos costos son difíciles de calcular. Incluso si nos pusiéramos de acuerdo en los probables costos de salud que generan la contaminación ambiental y los accidentes de tránsito, ¿cómo podríamos asignar un valor monetario al poder caminar tranquilamente de un lugar a otro? Y, por supuesto, los críticos podrían refutar que también se incurre en costos al construir y mantener

los sistemas de transporte público. ¿Y qué queda de la libertad que ofrece un vehículo particular: la posibilidad de elegir la ruta y el horario según el gusto de cada uno? ¿Acaso no debieran incorporarse estas consideraciones en los cálculos?

Existen dos soluciones a los problemas planteados por los factores intangibles al análisis de costo-beneficio, pero ninguna de ellas es completamente satisfactoria. La primera es dejar a un lado estos factores, ignorar lo que no se puede cuantificar. Pero, ¿debiéramos considerar irrelevantes factores como la calidad de vida, simplemente porque no es posible medirla? La segunda opción es estimar el valor de los costos y beneficios, asignándoles un valor monetario. Este enfoque permite tener presente los factores, pero el proceso es, a todas luces, arbitrario; ¿qué valor monetario debiéramos asignar a la comodidad o a la libertad? Al asignar un valor determinado, ya sea muy bajo o muy alto, es fácil sesgar los resultados de cualquier análisis de costo-beneficio.

Nuestra cultura tiene problemas en asignar valor a cierto tipo de factores. Por ejemplo, cada cierto tiempo, la prensa se escandaliza de que se asigne un valor a la vida de las personas mediante un análisis de costo-beneficio⁹. Esos casos generan indignación: ¿cómo es posible que alguien coloque un precio en dinero a la vida humana?, ¿es que acaso la vida humana no es invaluable? La respuesta depende de dónde y cuándo se hagan las preguntas. En EEUU, la idea de que la vida humana no tiene precio es relativamente nueva. Hace sólo un siglo, los padres de un niño arrollado por un tranvía podían demandar a la empresa de transporte por daños equivalentes al valor económico que ese niño suponía a la familia (los probables ingresos que ese niño habría generado hasta la adultez); por supuesto, hoy en día los progenitores demandarían por el valor (muchísimo más alto) de su dolor y

⁹ Para una crítica a una conocida historia sobre analistas de costo-beneficio amorales, véase Matthew T. Lee y M. David Ermann, "Pinto 'Madness' as a Flawed Landmark Narrative", *Social Problems* 46 (1999): 30-47.

sufrimiento. Incluso el valor monetario de la vida de un niño varía a lo largo del tiempo y el espacio¹⁰.

La mayor dificultad radica en que, inevitablemente, todo tiene sus externalidades o efectos colaterales. Construir un puente o implementar un programa de vacunación infantil tiene riesgos y costos; como también lo tienen las alternativas de no construir el puente o no vacunar a los niños. Ya que debido a las vacunas algunos niños pueden enfermar o morir, la discusión en EEUU acerca de seguir o no vacunando es muy compleja. Los defensores de ambas posturas intentan sortear este debate creando alternativas demasiado simples: los que apoyan la idea de la vacunación pueden declarar que el daño ocasionado casi no existe y que, por el contrario, el riesgo de no vacunar tendría consecuencias terribles; a su vez, los opositores insistirán en que las vacunas son perjudiciales y no aportan demasiado. Obviamente, ambos bandos harían bien en utilizar buenos datos. Más allá de eso, debemos reconocer que cada una de las opciones implica costos y que podemos poner en la balanza y elegir entre opciones imperfectas. Incluso si admitiéramos que una vacuna podría matar a unos pocos niños, pero al mismo tiempo salvar a muchos, ¿cómo podemos incorporar a nuestra decisión la idea de que el valor de cada vida humana es incalculable? ¿Cómo podríamos medir el valor de las vidas que se perderían si el plan de vacunación se implementara en contraposición a la gran cantidad de vidas que se extinguirían si este plan no se pone en práctica? (El Capítulo 3 se extiende sobre esta discusión acerca de las externalidades o efectos colaterales).

En resumen, en algunas discusiones sobre temas sociales faltan cifras porque ciertos fenómenos son difíciles de cuantificar y cualquier intento de asignarles un valor numérico estará sujeto a debate. Pero negarse a incorporar esos factores en nuestros cálculos conlleva riesgos. Una buena solución es reconocer las dificultades a las que nos

¹⁰ Viviana A. Zelizer, *Pricing the Priceless Child* (New York: Basic Books, 1985).

enfrentamos cuando medimos estos fenómenos, discutirlos de manera abierta y sopesar las opciones como mejor podamos.

LO QUE NO SE CONTABILIZA

Una tercera categoría de cifras que brillan por su ausencia se refiere a las que no son contabilizadas deliberadamente, los registros que no se conservan. Consideremos los cálculos sobre afiliación religiosa de la Oficina del Censo de EEUU (Bureau of the Census): no existen. De hecho, en el censo no se aplica ninguna pregunta acerca de la religión. Los argumentos sobre la separación constitucional de Iglesia y Estado, así como el sentido común de que la religión es un tema delicado, han llevado a la Oficina del Censo a omitir las preguntas acerca de la religión cuando se encuesta a la ciudadanía (en contraste con la mayoría de los países europeos, donde esas consultas sí se hacen)¹¹.

Así, cualquiera que intente calcular el nivel de actividad religiosa en EEUU deberá basarse en cifras más o menos exactas, como los registros de las iglesias o los informes de asistencia a los oficios religiosos. Los registros de las diversas denominaciones religiosas varían en torno a qué es lo que se contabiliza: ¿se suman los niños una vez bautizados, o es que uno se convierte en un miembro en la infancia o la adultez? ¿Son eliminadas de los registros las personas que dejan de participar en ciertas actividades religiosas? Esas variaciones dificultan comparar el tamaño de las diversas religiones (como se verá en el Capítulo 6). En ocasiones, otras encuestas aparte del censo consultan a las personas acerca de la frecuencia con que asisten a los oficios religiosos, pero tenemos motivos para sospechar que los encuestados sobrestiman su asistencia (quizás para causar una buena impresión a los encuestadores)¹². El resultado es que, por lo menos en EEUU, es difícil

¹¹ William Petersen, *Ethnicity Counts* (New Brunswick, NJ: Transaction, 1997), pp. 78-81.

¹² C. Kirk Hadaway, Penny Long Marler y Mark Chaves, "What the Polls Don't Show: A

medir con precisión las preferencias o el nivel de participación religiosa de la población. La política de no consultar sobre religión en el censo se traduce en que esa información sencillamente no existe.

La manera en que se redactan las opciones en las encuestas también genera categorías no contabilizadas. Desde 1790, en cada censo se ha consultado sobre raza u origen étnico, pero la formulación de las preguntas —y la selección de respuestas posibles— ha cambiado. Por ejemplo, el censo de 2000 fue el primero en agregar la alternativa “multiétnico” y dar la posibilidad a los encuestados de identificarse como tales. Los impulsores de este cambio arguyeron que en el árbol genealógico de muchos estadounidenses, existen antepasados de diferentes razas y que no era razonable forzar a las personas a ubicarse en una sola categoría racial.

Sin embargo, algunos defensores de esta modificación tenían otro motivo para promoverla. Cuando se les obligaba a optar por una sola categoría, los individuos que sabían que entre sus antecedentes familiares había personas de diferentes etnias, debían simplificar demasiado sus respuestas; es probable que muchos seleccionaran la opción que calzaba con la mayoría de sus antepasados. Por ejemplo, es posible que un individuo entre cuyos abuelos había tres personas blancas y un indio americano, eligiera la opción “blanco”. En una sociedad donde la influencia política de un grupo depende en gran parte de su tamaño, esa opción podría disminuir el número de personas con antepasados indoamericanos (o cualquier otro grupo pequeño pero con un alto índice de mezcla racial) identificadas en el censo. Los activistas indoamericanos apoyaban que las personas pudieran catalogarse en más de una raza, porque creían que esto permitiría identificar una población indoamericana mayor y, por lo tanto, aumentar la influencia política de su grupo. Por otro lado, los activistas afroamericanos no estaban tan entusiasmados con la idea de permitir a las personas reconocer su

origen multiétnico. Basándose en parte en el legado del segregacionismo (que sostenía que bastaba tener un solo antepasado negro para garantizar el no ser blanco), las personas con ascendencia mixta blanca y negra (que incluye a la mayoría de quienes usualmente son clasificados como afroamericanos) tendían a calificarse como “negros”. Si un gran número de estos individuos comenzaba a identificarse como miembro de más de un grupo racial, las personas de raza negra arriesgarían perder su influencia política.

Como suele suceder, las actitudes con relación a modificar las categorías censales dependían de si uno esperaba ganar o perder con un determinado cambio. La reclasificación tuvo el efecto esperado, aun cuando sólo un 2,4% de los encuestados optó por describirse como multiétnico. La nueva clasificación aumentó las cifras de personas catalogadas como indoamericanas: si bien sólo 2,5 millones de encuestados seleccionaron la categoría tradicional de una sola etnia, a ello se pueden agregar los que se identificaron como indios en parte y obtenemos un total de 4,1 millones, un aumento de un 110% desde 1990. Sin embargo, un número relativamente pequeño de personas (menos de 800 mil) optó por la alternativa “blanco y negro”, comparado con casi 34 millones que se definieron como negros¹³.

Existen ocasiones en que sólo ciertos casos no se contabilizan. Los críticos plantean que la tasa oficial de desempleo, que considera solamente a aquellos sin jornada laboral completa que buscaron empleo durante las cuatro semanas previas, es demasiado baja. Insisten que un conteo más preciso debiera incluir a quienes desean trabajar pero que dejaron de buscar empleo, así como a los que quieren un trabajo a tiempo completo pero que debieron arreglárselas con empleos a tiempo parcial; dos grupos que, tomados en conjunto, superan en mucho a

¹³ Para el censo de 2000, véase Margo Anderson y Stephen E. Fienberg, “Census 2000 and the Politics of Census Taking”, *Society* 39 (November 2001): 17-25; y Eric Schmitt, “For 7 Million People in Census, One Race Category Isn’t Enough”, *New York Times*, March 13, 2001, p. A1. Para antecedentes históricos, véase Petersen, *Ethnicity Counts*.

los oficialmente desempleados¹⁴. Es evidente que cada definición traza una distinción entre lo que cuenta y lo que no.

La lección es muy sencilla. Las estadísticas dependen de cómo se recoge la información. Si no se hacen las preguntas, si se realizan de manera que limitan las respuestas, o si la medición considera ciertos casos pero excluye otros, lo que finalmente sucede es que faltan cifras. Sin embargo, es inevitable tomar decisiones relativas a qué datos recoger y cómo hacerlo. Si queremos describir la composición racial de EEUU de una manera que sea comprensible, debemos incluir una gran diversidad en unas pocas categorías. El costo de clasificar cualquier cosa en un grupo particular de categorías es que inevitablemente se perderá información: las distinciones pueden confundir; los que pudieron ser límites arbitrarios son considerados significativos; y, en particular, tendemos a perder de vista las opciones e incertidumbres que surgieron al crear las categorías.

En algunos casos, los críticos plantean que una falla en la recolección de información es intencional, una forma de evitar entregar información que se considera dañina. Por ejemplo, es muy difícil reunir datos acerca de las circunstancias en que civiles han recibido disparos de policías. Podríamos imaginar que los baleos por parte de la policía se dividen en dos categorías: aquellos justificados por las circunstancias y los que no. De hecho, en muchos departamentos de policía se evalúan los baleos para definir si fueron o no justificados. Sin embargo, los esfuerzos para conseguir información nacional sobre estos datos no han fructificado. No todos los departamentos de policía comparten su información (lo que, según algunos críticos, implica que ocultan algo), y la proporción de baleos rotulados como “justificados” varía mucho de departamento en departamento (lo que sugiere que la conducta de la policía es muy diferente entre los departamentos o que el proceso de

¹⁴ Para los argumentos básicos acerca de cómo medir el desempleo, véase David Leonhardt, “Breadline? What Breadline?”, *New York Times*, June 24, 2001, sec. 4, p. 5; y Leonhardt, “Help Wanted: Out of a Job and No Longer Looking”, *New York Times*, September 29, 2002, sec. 4, p. 1.

revisión de los baleos es extraordinariamente distinto entre uno y otro)¹⁵.

Existen muchas maneras de asegurarse de que las cosas no se contabilicen. El modo más sencillo es no recoger la información (p.ej., no hacer ninguna pregunta sobre religión en el censo). Pero incluso cuando la información existe, es posible no compilarla (por el mero hecho de no realizar ciertos cálculos necesarios para generar determinadas estadísticas), rehusarse a publicarla o incluso bloquear el acceso a ella¹⁶. De una forma mucho más sutil, tanto la recolección de información como el análisis pueden ser costosos y tomar mucho tiempo; en una sociedad donde los investigadores dependen de otros para el financiamiento, la decisión de no financiar una determinada investigación puede tener el efecto de relegar ciertos temas al ámbito de lo no contabilizado.

Lo anterior funciona en dos sentidos. Inevitablemente, también escuchamos argumentos de que las personas *debieran* dejar de recoger cierto tipo de información. Por ejemplo, existe una conocida guía que entrega a los futuros estudiantes un ranking de las “universidades fiesteras”. Un programa auspiciado por la Asociación Médica Estadounidense llamado “A Matter of Degree”, y que se opone al consumo abusivo de alcohol en los campus universitarios, sostiene que esta clasificación minimiza la importancia del problema, y quizás contribuye al consumo de alcohol; por lo mismo, ha solicitado que esa guía deje de publicar la mencionada lista¹⁷. Si bien es poco común que los

¹⁵ Fox Butterfield, “When Police Shoot, Who’s Counting?”, *New York Times*, April 29, 2001, sec. 4, p. 5. Para una crítica más amplia acerca de cómo las políticas pueden configurar lo que los funcionarios deciden no contabilizar, véase A.P. Tant, “The Politics of Official Statistics”, *Government and Opposition* 30 (1995): 254-266.

¹⁶ Para un estudio de caso de cómo las decisiones políticas pueden definir la forma en que una agencia oficial (Canadian Centre for Justice Statistics) recoge, analiza y presenta la información, véase Kevin D. Haggerty, *Making Crime Count* (Toronto: University of Toronto Press, 2001).

¹⁷ American Medical Association, Office of Alcohol and Other Drug Abuse, “College Binge-Drinking Prevention Program Calls on *Princeton Review* to Stop Publishing ‘Party Schools List’”, comunicado de prensa, 7 de agosto de 2002.

críticos se preocupen de la mala influencia que las estadísticas puedan ejercer en relación con la moral, algunos creen que es mejor que no se recoja información y, por consiguiente, permanezca sin ser conocida.

NÚMEROS OLVIDADOS

Las cifras que brillan por su ausencia tienen otra forma de ser pasadas por alto; son datos que alguna vez fueron de público conocimiento e incluso familiares, pero que hemos olvidado o no queremos tomar en cuenta. Consideremos la cantidad de muertes causadas por el sarampión. En 1900, la tasa de mortalidad del sarampión era de 13,3 por cada 100.000 personas; el sarampión figuraba entre las 10 enfermedades fatales en EEUU. Sin embargo, a lo largo del siglo, el sarampión perdió su poder letal; las causas de su eliminación como amenaza médica importante fueron, en primer lugar, los tratamientos eficaces y, luego, la vacunación. Éste no es un caso excepcional. A comienzos del siglo XX, muchas de las causas de muerte se debían a enfermedades infecciosas; la influenza/pulmonía/neumonía, la tuberculosis, la difteria y la fiebre tifoidea también se ubicaban en los *top ten*¹⁸. En EEUU, la mayoría de las que fueran enfermedades devastadoras están casi completamente controladas gracias a las vacunas y antibióticos. La variedad de amenazas médicas ha cambiado.

Si pudiéramos recordarlos, los números olvidados tienen el potencial de ayudarnos a poner las cosas en perspectiva. Cuando perdemos de vista el pasado, tenemos más dificultades para evaluar la situación actual. Sin embargo, las personas que intentan llamar la atención sobre los problemas sociales actuales suelen ser reticentes a comparar con el pasado. Después de todo, las comparaciones pueden mostrar que hay progreso. Por ejemplo, durante el siglo XX, la expectativa de vida de los estadounidenses aumentó de manera drástica. En 1900, un recién

¹⁸ Bureau of the Census, *Historical Statistics of the United States* (Washington, DC, 1975), p. 58.

nacido varón podía esperar vivir 46 años; un siglo más tarde, la expectativa de vida masculina se ha extendido hasta los 73 años. El aumento en el caso de las mujeres ha sido aún mayor: de 48 años a 80. Durante el mismo período, la proporción de estadounidenses que terminaban la enseñanza media aumentó de aproximadamente 6% a un 85%. Pareciera que muchas personas sienten temor de hablar sobre el progreso logrado en el tiempo, pues ello invitaría a ser complacientes con la sociedad contemporánea, y por lo tanto, prefieren enfocarse en las tendencias a corto plazo —especialmente si las cifras parecen convincentes dado que indican que las cosas empeoran¹⁹.

Asimismo, comparar nuestra sociedad con otras puede ayudarnos a comprender mejor el tamaño y la forma de nuestros problemas. Una vez más, en la discusión sobre los problemas sociales y para enfatizar la magnitud de nuestros problemas contemporáneos, esas comparaciones tienden a hacerse de manera selectiva. Cuando la información sugiere que EEUU se queda atrás, las estadísticas comparativas son un lugar común, pero podríamos sospechar que aquellos que intentan promover la acción social estarán menos dispuestos a presentar evidencias que demuestren que EEUU está en ventaja. (Por supuesto, aquellos que se resisten al cambio podrían apoyar esas cifras). Las comparaciones a lo largo del tiempo y el espacio son recordadas cuando ayudan a fundamentar sus postulados a quienes defienden alguna postura; de lo contrario, tienden a ser ignoradas u olvidadas.

CIFRAS LEGENDARIAS

Cabe destacar una última categoría. No incluye cifras potencialmente importantes que están faltando, sino que se relaciona con cifras erróneas o irrelevantes que de alguna manera se infiltraron en la discusión

¹⁹ Para ver algunas de las formas en que el progreso nos ayuda a ser más conscientes de los problemas sociales, véase Joel Best, "Social Progress and Social Problems", *Sociological Quarterly* 42 (2001): 1-12.

de los temas sociales. Por ejemplo, se ha vuelto común que los periodistas comparen diversos riesgos contra un estándar muy peculiar: el número de personas fallecidas cada año en todo el mundo por la caída de cocos desde la copa de palmeras tropicales (la cifra anual de muertes por caída de cocos solía ser de 150). ¿Es que 150 personas realmente mueren por esta causa? Parece posible: los cocos son pesados y duros y caen a gran velocidad, por lo que ser golpeado en la cabeza presumiblemente sea fatal. ¿Pero quién contabiliza estas muertes? La respuesta: nadie. Si bien hay literatura médica que reporta heridas —no muertes— provocadas por la caída de cocos, esa cifra de 150 muertes es el equivalente periodístico de un mito urbano²⁰. Empieza a correr de boca en boca como un “hecho verdadero” que se repite como algo que “todo el mundo sabe”.

Otras estadísticas legendarias se atribuyen a fuentes supuestamente acreditadas. La afirmación que en teoría habría hecho la Organización Mundial de la Salud (OMS) acerca de que el color rubio del cabello se debía a un gen recesivo y que por lo tanto las personas rubias se extinguirían en unos doscientos años, fue dada a conocer por importantes medios de prensa, los que reportaban el tema basándose en la cobertura de unos y otros, sin chequear con la OMS (que finalmente negó la historia)²¹.

Las cifras legendarias se instalan con facilidad. Por ejemplo, consideremos la afirmación de que los 56 años es la edad promedio en que una mujer enviuda. A pesar de lo poco probable (después de todo, los hombres viven más de 70 años en promedio, los casados viven más

²⁰ La difusión de la estadística de caída de cocos fue apoyada por una compañía inglesa de seguros de viaje; véase Beverly Beckham, “Travelers Should Watch Out for Coconuts: The Killer Fruit”, *Boston Herald*, April 7, 2002, p. 30. El 19 de julio de 2002, en el sitio web de Cecil Adams, “The Straight Dope” (www.straightdope.com/columns/020719), apareció una discusión de la información disponible.

²¹ Lawrence K. Altman, “Stop Those Presses! Blonds, It Seems, Will Survive After All”, *New York Times*, October 2, 2002, p. A5.

que los solteros y los esposos son en promedio sólo algunos años mayores que sus mujeres), este dato ha circulado por más de veinte años. Apareció en un comercial de servicios financieros transmitido por televisión, en materiales entregados a alumnos de programas de estudios de la mujer y en numerosos artículos de periódicos y revistas; pero su origen se pierde en el tiempo. Quizás ha permanecido porque ningún organismo recoge el dato de la edad en que las personas enviudan, dificultando el cuestionamiento de esa cifra tan repetida. Sin embargo, los demógrafos —utilizando complicadas ecuaciones que incluyen tasas de mortalidad relativas a edades específicas, el porcentaje de personas casadas en diferentes cohortes por edad y la diferencia de edad entre maridos y esposas— han concluido (sin sorprender a nadie) que la edad promedio en que las mujeres enviudan aumentó de 65 años en 1970 a alrededor de 69 en 1988²².

Incluso los números que actualmente se originan en las afirmaciones de los científicos pueden adquirir cualidades legendarias. En parte, esto refleja la dificultad de traducir complejas ideas científicas en afirmaciones supuestamente fáciles. Por ejemplo, la conocida sugerencia de beber ocho vasos de agua al día tuvo su origen en un análisis que recomendaba hacerlo. Pero el análisis también indicaba que gran parte de esa agua podía provenir de los alimentos (p.ej., el pan tiene un 35% de agua y las carnes y verduras contienen proporciones aún mayores). Sin embargo, la idea de que los alimentos contienen la mayor cantidad del agua necesaria para una buena salud, fue prontamente olvidada, a favor de la indicación de consumir la mayor cantidad bebiéndola²³. Asimismo, las repetidas afirmaciones de que el ADN de los seres

²² Robert Schoen y Robin M. Weinick, "The Slowing Metabolism of Marriage: Figures from 1988 U.S. Marital Status Life Tables", *Demography* 30 (1993): 737-746. El primer comentario que encontré al respecto fue el de Judy Klemesrud, "If Your Face Isn't Young: Women Confront Problems of Aging", *New York Times*, October 10, 1980, p. A24.

²³ Edward M. Eveld, "Awash in Water: Eight Glasses a Day? It's Probably More Than We Need, Scientists Say", *Seattle Times*, July 13, 2003, p. L6.

humanos y los chimpancés se asemeja en un 98% —ó 98,4%, 99% ó 99,44%— parecen muy exactas y precisas, pero ignoran los complejos supuestos involucrados en esos cálculos e insinúan que esta cifra tiene un significado mayor del que de hecho tiene²⁴.

Las cifras que circulan con tanta facilidad no son necesariamente válidas ni significativas. En el mundo moderno, con fácil acceso a internet y todo tipo de bases de datos computarizadas, hasta los números que han sido desacreditados de manera minuciosa pueden continuar circulando; son fáciles de recuperar y difundir, pero casi imposibles de erradicar. El problema no son las cifras que faltan; en esas circunstancias, los números están muy presente. Lo que falta es la evidencia necesaria que otorga credibilidad a las estadísticas.

El atractivo de las cifras legendarias radica en que parecieran reforzar o dar autoridad a una afirmación. Es mucho menos convincente decir: “¡Ésa no es una causa importante de muerte! ¡Te apuesto que mueren más personas por la caída de cocos!”, que sencillamente comparar 150 muertes provocadas por cocos que caen con algún otro motivo. Se supone que las cifras dan cuenta de hechos; implican que alguien contabilizó algo. Por supuesto, si eso es cierto, debiera ser posible documentar la afirmación —precisamente lo que no se puede hacer con las cifras legendarias.

Un fenómeno relacionado con lo anterior es que algunas cifras, aunque no parezcan demasiado llamativas, pueden considerarse más significativas de lo que de hecho son. (El Capítulo 5 también se refiere a este tema). Podemos ver esto en los esfuerzos de los burócratas para medir lo que no se puede medir. Por ejemplo, un municipio quiere premiar a los buenos profesores. Pero, ¿qué hace que un profesor sea un buen profesor? La mayoría podemos hacer memoria e identificar algunos profesores que fueron mejores que otros. ¿Qué los hizo mejores?

²⁴ Para una discusión más minuciosa de este ejemplo, véase Jonathan Marks, *What It Means to Be 98% Chimpanzee: Apes, People, and Their Genes* (Berkeley: University of California Press, 2002).

Quizás nos apoyaron cuando estuvimos en problemas, nos estimularon o exigieron en un área específica. Mis razones para destacar a algunos profesores pueden ser muy distintas a las de otra persona. Los profesores pueden ser excelentes de muchas maneras, y es probable que no exista un método confiable para convertir el grado de excelencia en una cifra. ¿Cómo podríamos medir una docencia de calidad o el genio artístico? Incluso los fanáticos del béisbol —esos amantes de las estadísticas que guardan fotos y noticias de manera compulsiva— pueden discutir sobre los méritos relativos de diferentes atletas, y eso que el béisbol tiene registros bastante completos del desempeño de los jugadores.

Pero es altamente improbable que ese tipo de atractivo hacia la imposibilidad de medir el desempeño tranquilice a los políticos o a un público enfurecido que exige mejores escuelas. Por ello, los burócratas de la educación —municipios y departamentos de educación pública— insisten en medir el “desempeño”. En los últimos años, la medida predilecta ha sido el puntaje que obtienen los estudiantes en los exámenes estandarizados. Esto no es completamente absurdo, ya que no es del todo descabellado argüir que un mejor educador hará que los estudiantes aprendan más y, por lo mismo, tengan un mejor rendimiento en las pruebas estandarizadas. Pero los puntajes en los exámenes dependen de muchas cosas además del desempeño de los profesores, incluyendo la vida familiar de los estudiantes. Y es muy probable que nuestros recuerdos acerca de los “mejores profesores” no dependan de cómo ellos influyeron en nuestro rendimiento en los exámenes estandarizados.

Aun cuando son imperfectos, los puntajes en los exámenes pueden ser indicadores de la calidad de la enseñanza, pues ofrecen una medida cuantitativa bastante buena: este estudiante respondió correctamente tantas preguntas, este curso obtuvo buenos puntajes, etc. Por eso, no es de extrañar que los burócratas estén tan pendientes de estas mediciones: son precisas (y es relativamente barato obtener la información), aun cuando su significado no sea muy claro.

Sucede lo mismo en otros contextos. Las universidades quieren que sus profesores realicen investigaciones de alto nivel y sean buenos docentes, pero todos reconocen que esas características son difíciles de medir. Por eso existe la enorme tentación de enfocarse en las cosas que son fáciles de contar: ¿cuántos libros o artículos ha publicado un académico? (Algunos departamentos incluso consideran selectivamente los artículos según la revista en que se publicaron, dependiendo de algún cálculo relativo a la influencia de cada una). ¿Son los puntajes de evaluación de un profesor mejores que el promedio?

El problema con esas mediciones burocráticas es que se pierden de vista sus limitaciones. Comenzamos diciendo que la calidad de la enseñanza debe ser medida y que este método —sin importar sus falencias— es mejor que ninguno. Incluso si al principio hay quienes se resisten, con el tiempo se instala la inercia y las personas aceptan su utilización. No mucho después, esas mediciones se darán por sentado y sus debilidades tenderán a ser olvidadas. Muchas de las cifras que se discutirán en los próximos capítulos pueden ser acusadas de imperfección. Si son presionados, los defensores de un dato estadístico reconocerán que esa crítica es válida, que la medición adolece de fallas. Pero, preguntarán, ¿qué alternativa existe? ¿De qué otra forma podemos medir —de manera rápida, barata y más o menos objetiva— la calidad de la enseñanza (o lo que sea que nos preocupe)? ¿Una estadística imperfecta no es acaso mejor que nada? Tienen razón en esto. Pero los defectos de una estadística nunca deben ofuscarnos; una vez que olvidamos las limitaciones de las cifras, les otorgamos mucho más poder e influencia de lo que merecen. Cabe recordar que *una medida directa y transparente es preferible y que una medida imperfecta es —insistimos— un tipo de cifra faltante.*

¿QUÉ ES LO QUE FALTA?

Cuando las personas utilizan estadísticas, asumen —o, al menos, quieren que su audiencia asuma— que las cifras significan algo. Esto quiere

decir que, como mínimo, alguien de hecho contabilizó algo y que lo hizo de una manera que tiene sentido. La información estadística es una de las mejores formas de hacer comprensibles las complejidades del mundo, de identificar modelos en medio de la confusión. Pero las malas estadísticas nos entregan información errada.

Este capítulo plantea que algunas estadísticas son malas no tanto porque la información que contienen sea errada, sino porque hay algo que falta: lo que no se contabilizó. Las cifras pueden estar ausente en muchos sentidos: un ejemplo poderoso puede hacer que olvidemos buscar las estadísticas; las cosas pueden no ser contabilizadas porque hacerlo se considera difícil o imposible, o porque sencillamente decidimos no contarlas. En otros casos, contabilizamos, pero algo se pierde en el proceso: las cosas que se contabilizaron fueron olvidadas, o podemos esgrimir cifras que no tienen contenido.

En todos estos casos, falta algo. Comprender eso nos ayuda a reconocer una buena estadística. Las buenas estadísticas no son sólo el producto de quienes contabilizan; la calidad de las estadísticas también depende del deseo y habilidad de las personas para contabilizar cuidadosamente y de su determinación acerca de qué contar para que las cifras resultantes sean exactas y significativas.

Este proceso nunca es perfecto. Cada cifra tiene sus propias limitaciones; cada cifra es resultado de opciones que inevitablemente involucran conciliaciones. Se supone que las estadísticas nos ayudan a resumir, a tener una visión en perspectiva de la complejidad del mundo. Pero en el proceso de seleccionar qué se contabilizará y cómo, siempre se sacrifica alguna información. En resumen, siempre falta algo. Hay algo que se hace notar por su ausencia. Al evaluar las estadísticas, no debemos olvidar lo que se ha perdido, aunque esto sólo nos ayude a comprender lo que tenemos.

Capítulo 2

CIFRAS ENGAÑOSAS

Un prestigioso comentarista político escribió recientemente en su columna: “En política, es un lugar común que alrededor del 40% de los republicanos siempre votará por el candidato presidencial republicano y más o menos el mismo porcentaje de demócratas votarán por el candidato de su partido. La batalla es por el 20% del centro”¹. Marcador: porcentajes 1; opinólogo 0. Las cifras que parecen sencillas suelen confundir incluso a quienes se les paga para orientar a los demás —y esto no es infrecuente.

Por ejemplo, los argumentos esgrimidos en la propuesta federal de rebaja de impuestos de 2003 parecían ser contradictorios. Los promotores del proyecto de ley declaraban que, para una familia promedio, la reducción de impuestos sobrepasaría los mil dólares; pero quienes se oponían destacaban que más de la mitad de las familias verían reducidos sus impuestos en menos de cien dólares². En otras palabras, según los que apoyaban el proyecto de ley, el beneficio promedio sería

¹ Cal Thomas, “Democrats Are Losing Center”, *Wilmington (Del.) News Journal*, August 3, 2003, p. A11.

² Para ejemplos de las afirmaciones de promotores y opositores a la rebaja de impuestos, véase David E. Rosenbaum, “Washington Memo: The President’s Tax Cut and Its Unspoken Numbers”, *New York Times*, February 25, 2003, p. A25.

mucho, y en opinión de quienes lo rechazaban, demasiado poco. ¿Le parece confuso? Pues, lo es.

La mayoría de nosotros supone que comprende el significado del término “promedio”. Si bien muchos críticos se lamentan de nuestro analfabetismo numérico o de nuestra incomodidad con las cifras, los estadounidenses consumen cotidianamente una gran cantidad de estadísticas, lo que incluye promedios, porcentajes y cosas por el estilo³. Tasas de delitos, índices bursátiles y promedios de goleadas no sólo constituyen material de los noticieros sino de las conversaciones del día a día. Se supone que entendemos esas cifras —y probablemente lo hacemos, pero a medias.

Sin embargo, la familiaridad puede llevar a confusión. Aunque sean aparentemente simples, las cifras sencillas pueden tender trampas a los desprevenidos. Es posible que se den a conocer estadísticas incorrectas, o se muestren cifras correctas de forma inadecuada. Como resultado, obtenemos confusión. Quizás sepamos que estamos confundidos (nos damos cuenta que no comprendemos los números) o a lo mejor no (creemos que entendemos las cifras cuando en realidad no es así). En algunas oportunidades, las mismas personas que dan a conocer las cifras pueden estar confundidas; en otros casos, saben lo que están haciendo e intentan engañarnos.

A pesar de que la confusión estadística adopta muchas formas, en este capítulo analizaremos sólo algunas. Para comprender algunos de los problemas más frecuentes con que nos topamos, es necesario aprender ciertos sencillos principios matemáticos y lógicos; nuestro objetivo será explorar cómo los procesos sociales —es decir, las personas que contabilizan— contribuyen a aumentar esos errores. El capítulo se inicia con algunos problemas comunes que involucran estadísticas conocidas, como los promedios y porcentajes, para luego

³ Para una introducción al concepto, véase John Allen Paulos, *El Hombre Anumérico: El Analfabetismo Matemático y sus Consecuencias* (Barcelona: Tusquets, 1998).

referirnos a temas específicos que surgen a raíz de los gráficos confusos.

ALGUNAS COMPLEJIDADES DE LOS NÚMEROS SENCILLOS

La estadística más simple es, desde luego, un recuento; alguien suma algo y lo da a conocer: nuestra ciudad tiene tantos habitantes, la policía registró un total de tantas denuncias al año, etc. Los recuentos pueden estar viciados, especialmente cuando los ítems contabilizados están parcialmente ocultos (lo que dificulta efectuar un conteo completo) o cuando son demasiado comunes (lo que puede encarecerlos tanto que sea necesario realizar estimaciones más baratas e imprecisas). También está el tema de qué es lo que debe contarse; es importante comprender cómo se define y mide lo que se contabiliza⁴. Pero, sobre todo, el conteo parece ser algo muy sencillo y directo. El concepto es fácil de entender, todos hemos contado cosas. Un recuento o conteo es una cifra sencilla que corresponde claramente a una pregunta conocida: ¿cuántos o cuántas hay? Es difícil confundirse con un recuento. A pesar de ello, no se puede decir lo mismo de otras estadísticas sencillas; incluso la aritmética fundamental puede llevar a confusión.

Promedios

Un tipo de confusión aritmética muy común tiene que ver con el concepto de promedio. El método habitual para calcular un promedio —que alguna vez aprendimos y generalmente damos por sentado— es sumar una cierta cantidad de datos y dividir esa suma por el número de casos. Por ejemplo, si un grupo de niños rinde una prueba de dictado de 10 palabras, para obtener el puntaje promedio del grupo podemos sumar

⁴ Respecto a temas vinculados con definición y mediciones, véase Best, *Damned Lies and Statistics*, pp. 39-52.

la cantidad de palabras que cada niño escribió correctamente y dividir las por el número de niños; digamos que fueron 8,2 palabras correctas.

El promedio calculado a través de este método tan conocido se denomina técnicamente *promedio* o *media*. El promedio es una medida útil, siempre y cuando los datos no varíen demasiado (p.ej., en nuestro dictado, el puntaje de un niño no puede ser más de 10 ni menos de 0). Pero imaginemos una fábrica con 90 obreros que ganan anualmente 40 mil dólares cada uno, 9 administrativos que ganan 80 mil dólares anualmente y un gerente general que se lleva a casa —me resisto a decir que “gana”— unos 6 millones de dólares. Si calculamos de la siguiente manera el ingreso promedio de las personas que trabajan en la fábrica:

$$\begin{array}{rcl}
 90 \times \text{US\$ } 40.000 & = & \text{US\$ } 3.600.000 \text{ ingreso total de obreros} \\
 9 \times \text{US\$ } 80.000 & = & \text{US\$ } 720.000 \text{ ingreso total de administrativos} \\
 1 \times \text{US\$ } 6.000.000 & = & \text{US\$ } 6.000.000 \text{ ingreso total de gerente general} \\
 \hline
 & & \text{US\$ } 10.320.000 \text{ (ingreso total)/100 (total de personas)} \\
 & & = \text{US\$ } 103.200 \text{ ingreso promedio}
 \end{array}$$

Pero esta media no tiene sentido. Nadie en la fábrica gana el promedio exacto; el salario de los administrativos (US\$ 80.000) se acerca más al promedio, pero la cifra promedio calculada (US\$ 103.000) dista mucho del ingreso de los obreros y de lo que gana el gerente general.

Una solución al problema sería medir el promedio de otra manera: en vez de la media o promedio, utilizar la mediana. La *mediana* es el valor de una variable que ocupa la posición central, o del medio, en un conjunto de datos. Para obtener la mediana, debemos hacer una lista en orden de los valores, desde el menor al mayor, y calcular el valor del medio. En el ejemplo de la fábrica hay 100 personas y los ingresos en los lugares 50 y 51 están en el medio. Ambos ingresos equivalen a US\$ 40.000, por lo que el valor de la mediana de ingreso de la fábrica es de US\$ 40.000. En este caso, la mediana nos da un valor más

exacto de un ingreso típico que el promedio. A fin de cuentas, el 90% de las personas en nuestra fábrica hipotética gana US\$ 40.000, lo que configura el salario típico. Dado que la distribución de los ingresos varía muchísimo, la mediana es la medida preferida para dar un sentido más real de lo que es un promedio. Por ello es que frecuentemente nos encontramos con referencias al “ingreso mediano de un hogar”, etc.

Ya sea que elijamos el promedio o la mediana para expresar un promedio, siempre perdemos algo de información. En el ejemplo de la fábrica, el 99% de las personas ganaba menos del ingreso medio, por lo que US\$ 103.200 es un promedio en un sentido bastante curioso. Asimismo, al utilizar el valor de la media, US\$ 40.000, perdemos de vista a quienes ganan más —ni hablar del gerente general. En este caso, la media es el número preferible, pero hay que admitir que no existe una medida perfecta; ninguna cifra en particular nos da una idea exacta de cómo se distribuye el ingreso en nuestra fábrica imaginaria. Además, la cifra que se utilice depende del énfasis que se le quiera dar: usar el promedio podría servir para destacar los considerables ingresos generados por la fábrica, en tanto que preferir la mediana sirve para poner el énfasis en el modesto salario de los obreros.

No es difícil hallar ejemplos de dichas diferencias: en relación a los comentarios a la disminución de impuestos que mencionamos al inicio de este capítulo, nótese cómo los promotores del proyecto de ley hacían mención a la reducción promedio, en tanto los oponentes se referían a la mediana. Cada vez que nos enfrentamos a un promedio, debíamos poder distinguir si se calculó como tal o como mediana. Idealmente, también habría que considerar si conocer la otra cifra cambiará nuestra impresión de lo que es el promedio.

Porcentajes

Parece ser que el porcentaje es la estadística que mejor manejamos: es sencillo de calcular, muy útil e instintivamente fácil de comprender. En

su forma más simple —cuando dividimos un todo en partes—, presenta pocos problemas. Si nos dicen que alrededor del 10% de las personas escribe con la mano izquierda, deducimos sin problemas que el 90% restante usa la derecha, y por lo tanto, que hay 9 diestros por cada zurdo.

Sin embargo, dado que el porcentaje es una herramienta tan útil y conocida, suele usarse para presentar información algo más complicada. Y no es necesario que las cosas se pongan demasiado difíciles para confundirnos. Imaginemos que hacemos un estudio de mil adolescentes, clasificándolos como delincuentes o respetuosos de la ley y como zurdos o diestros. Supongamos que entre los que usan la mano derecha hay 810 respetuosos de la ley y 90 delincuentes, mientras que entre los zurdos encontramos 80 que respetan la ley y 20 que delinquen. (Antes que alguien se ofenda, permítanme aclarar que estos son datos imaginarios que sólo buscan ilustrar lo que quiero decir). Para revisar esas cifras y entenderlas, ordenemos los datos como en la Tabla 1, que contiene celdas para las cuatro posibles combinaciones de lateralidad y respeto por la ley.

Tabla 1. Números en bruto (un tanto confusos)

	Diestros	Zurdos
Respetuosos de la ley	810	80
Delincuentes	90	20

Esta tabla no ayuda mucho, ¿verdad? Veamos entonces qué sucede con la Tabla 2, en que calculamos los porcentajes para que cada fila sume 100%. (Nótese que de aquí en adelante, junto con los porcentajes, incluyo el número real de personas entre paréntesis en cada celda, lo que permitirá al lector chequear mis cálculos). La Tabla 2 indica claramente que los diestros son mayoría tanto entre los adolescentes

que respetan la ley como entre quienes delinquen. Pero, como sabemos que hay muchos más diestros que zurdos, ese dato no es demasiado interesante.

Tabla 2. Cálculo de porcentajes por fila (todavía confusos)

	Diestros	Zurdos	Total
Respetuosos de la ley	91% (810)	9% (80)	100% (890)
Delincuentes	82% (90)	18% (20)	100% (110)

Observemos ahora qué sucede cuando calculamos los porcentajes hacia abajo para que cada columna sume 100%, como en la Tabla 3. Gracias a este cálculo, vemos claramente la pauta: en nuestro estudio, el porcentaje de zurdos que delinquen es el doble (20%) que los diestros (10%).

Tabla 3. Cálculo de porcentajes por columna (mucho más claros)

	Diestros	Zurdos
Respetuosos de la ley	90% (810)	80% (80)
Delincuentes	10% (90)	20% (20)
Total	100% (900)	100% (100)

La Tabla 2 ilustra lo que queremos decir cuando planteamos que alguien calculó los porcentajes “de mala manera” o “en la dirección equivocada”. En general, los porcentajes deberían calcularse para que cada valor de la variable independiente* sume 100%. En nuestro ejemplo,

* Por variable “independiente” se entiende la causa del comportamiento de la otra variable, llamada variable “dependiente”. Así, por ejemplo, haber nacido en Latinoamérica es la variable independiente de hablar español (N. del E.).

el ser zurdo o diestro es la variable independiente; nadie creará que ser respetuoso de la ley o delinquir pueda hacer que una persona sea zurda o diestra, pero en cambio sí podemos imaginar que la lateralidad afecte de alguna manera el respetar o no las leyes⁵. Por ello, necesitamos calcular nuestros porcentajes de tal forma que las columnas de adolescentes diestros y zurdos sumen, cada una, 100% (como sucede en la Tabla 3).

El decidir cómo calcular porcentajes exige ponerse a pensar. Con sorprendente frecuencia, a la hora de dar a conocer porcentajes mal calculados, uno encuentra personas que debieron haber sido más juiciosas.

En otras ocasiones, algunos porcentajes enrevesados aparecen impresionantes. Supongamos, por ejemplo, que alguien plantea que un alto porcentaje de alcohólicos —60%— sufrió de abuso infantil. (Una vez más, estoy inventando cifras para ilustrar un punto). Tanto quienes afirman esto como las personas que lo lean podrían considerar ese dato como una demostración de que el ser abusado durante la infancia incide en la probabilidad de que las personas se vuelvan alcohólicas. Sin embargo, esta estadística nos entrega porcentajes calculados de una manera errónea.

Reflexionemos un poco sobre esto. El alcoholismo de un adulto no puede ser la causa de que éste haya sido abusado cuando niño; por lo tanto, el abuso —no el alcoholismo— debe considerarse como la causa o la variable independiente. Lo que queremos comparar es el porcentaje de individuos que sufrieron abuso infantil y que se convirtieron en adultos alcohólicos, con aquellos que sin ser abusados se volvieron alcohólicos. O sea, debemos comparar el porcentaje de los que fueron

⁵ Después de imaginar este ejemplo, descubrí que existen estudios que vinculan el ser zurdo con la delincuencia. Hay evidencias de que ambos factores están relacionados (si bien no tanto como sugieren los datos que inventé), pero no existe acuerdo en relación con su veracidad. Por ejemplo, véase William C. Grace, "Strength of Handedness as an Indicant of Delinquents' Behavior", *Journal of Clinical Psychology* 43 (1987): 151-155.

abusados con los que no lo fueron, para que cada uno sume 100%. Si el primer número es más alto que el segundo —como muestra la Tabla 4—, los datos efectivamente sugieren que una historia de abuso podría influir en la probabilidad de llegar a ser alcohólico.

Tabla 4. Datos imaginarios que muestran que el abuso infantil incide fuertemente en volverse alcohólico en la adultez

	Abusados	No abusados
Alcohólicos	20% (120)	10% (80)
No alcohólicos	80% (480)	90% (720)
Total	100% (600)	100% (800)

Cuando escuchamos que la mayoría de los alcohólicos fueron abusados en su niñez, tendemos a creer que las personas que sufrieron de abusos tienen más probabilidades de ser alcohólicas (tal como muestra la Tabla 4). En esta tabla, el porcentaje de alcohólicos entre los que fueron abusados de niños (20%) es el doble de los alcohólicos que no sufrieron abusos (10%). (Nótese también que mis datos imaginarios incluyen a 200 alcohólicos de los cuales el 60% [120 casos] fueron abusados. Por lo tanto, estos datos efectivamente apoyan la afirmación de que el 60% de los alcohólicos fue víctima de abuso infantil).

Consideremos ahora las Tablas 5 y 6, que utilizan cifras diferentes. En la Tabla 5, vemos el mismo porcentaje de personas que se hicieron alcohólicas entre quienes fueron abusados de niños y los que no (20% en ambos casos). Esta información sugiere que el abuso durante la infancia no incide en el alcoholismo adulto. La Tabla 6 muestra que las personas que no sufrieron de abuso cuando niños, tienen más probabilidades de volverse alcohólicas (30%) que las que sí lo fueron (20%).

Tabla 5. Datos imaginarios que muestran que el abuso infantil no tiene efectos en el alcoholismo adulto

	Abusados	No abusados
Alcohólicos	20% (120)	20% (80)
No alcohólicos	80% (480)	80% (320)
Total	100% (600)	100% (400)

Tabla 6. Datos imaginarios que muestran que el abuso infantil tiene pocas probabilidades de incidir en el alcoholismo durante la adultez

	Abusados	No abusados
Alcohólicos	20% (120)	30% (80)
No alcohólicos	80% (480)	70% (189)
Total	100% (600)	100% (269)

A pesar de las notables diferencias, una mirada más cuidadosa de las filas superiores de las Tablas 4-6 muestra que, en cada caso, se mantiene el porcentaje errado (60% de los alcohólicos —120 de 200— fue abusado cuando niño). Aun cuando presentar porcentajes calculados de manera equivocada nos haga pensar que la tendencia general es la que muestra la Tabla 4, es importante entender que las tres tablas, pese a los datos contradictorios, son coherentes con ese porcentaje equivocado. A esto nos referimos cuando calculamos mal los porcentajes; lo que, más que aclarar, confunde.

También podemos engañarnos cuando describimos una secuencia de cambios por medio de porcentajes. Supongamos que un determinado índice bursátil cayó un 50% entre 1980 y 1990, pero luego subió un 95% entre 1990 y 2000. ¿Era mayor el valor en el año 2000 que su valor de 1980? A primera vista, podría pensarse que sí (“bajó un 50% pero luego subió un 95%, y 95 es mucho más que 50...”). No obstante,

la respuesta correcta es que no, el valor en el año 2000 no era mayor que en 1980.

Supongamos ahora que el valor del índice en 1980 era de US\$ 1.000. Una caída de 50% en 1990 habría provocado una disminución de US\$ 500 (el 50% de mil dólares es US\$ 500, los que restamos a los originales US\$ 1.000). Sin embargo, el alza de 95% entre 1990 y 2000 debe medirse contra el valor de 1990 (95% de 500 es 475, lo que sumado al valor de 500 en 1990, da un total de 975, que es menos que la cifra de US\$ 1.000 en 1980). Es decir, *un cambio porcentual siempre debe calcularse en comparación con el valor inicial de la secuencia*. Al describir cualquier serie de transformaciones (como las variaciones del índice mencionado de 1980 a 1990 y luego a 2000), el resultado de una modificación crea una nueva base sobre la cual debe calcularse el cambio siguiente. De ahí que si bien es fácil entender cada cambio en forma individual, necesitamos pensar con cuidado cuando comparamos una serie de variaciones porcentuales.

Intente calcular esto rápidamente: si el total del índice bursátil subió un 50% entre 1980 y 1990, y luego cayó un 95% entre 1990 y 2000, ¿cuál es el valor mayor, el de 1980 o el de 2000? La respuesta, una vez más, es el valor de 1980, y ahora con un margen amplio. Asumamos otra vez que el valor de 1980 era US\$ 1.000. Un aumento de 50% implicaría un alza en 1990 a US\$ 1.500, pero una caída de 95% de US\$ 1.500 daría un total de US\$ 75 para el año 2000. La lección es bien sencilla: una serie de cambios, expresados en forma porcentual, da por resultado cifras que en realidad no son comparables⁶. Es probable que quienes presentan la información de esta forma, estén confundidos o desean engañar a propósito.

⁶ Sin embargo, suele ser correcto comparar dos o más cambios, expresados en porcentajes, ocurridos durante el mismo período. Por ello, si las tasas relativas a dos diferentes delitos son medidas en 1980 y de nuevo en 1990, y la información que manejamos indica que el delito X aumentó en un 25% en tanto que el delito Y aumentó en un 50%, podemos concluir que el delito Y aumentó más rápido que el delito X.

Aunque los conceptos de promedio y porcentaje son conocidos y parecen muy simples, potencialmente pueden confundirnos. Entonces no debe sorprendernos que la confusión aumente a medida que las ideas estadísticas se hacen más complejas.

El significado de la correlación

Una de las formas más importantes de razonamiento ocurre cuando reconocemos que dos cosas están vinculadas (p.ej.: cuando acciono este interruptor hacia arriba, la luz se enciende, y cuando lo acciono hacia abajo, se apaga). Este reconocimiento nos invita a sospechar que una cosa puede ser causa de otra, lo que nos permite entender nuestro mundo y planificar nuestras acciones basándonos en lo que creemos saber (“está demasiado oscuro, creo que accionaré el interruptor”).

Una relación pautaada entre dos cosas puede tener varias formas: cada vez que A sube, B sube; cada vez que A sube, B baja; cuando A sube, B tiende a subir un poco; etc. Los filósofos y científicos clasifican estas relaciones como formas de *correlación*. Cuando decimos que A y B están correlacionados, estamos apuntando a algún tipo de nexo observable entre ambos, ya sea una correspondencia perfecta uno a uno (la luz se enciende cada vez que subo el interruptor, y sólo cuando éste se acciona hacia arriba) o solamente una leve tendencia (quienes sufrieron de abuso cuando niños tienen alguna probabilidad de llegar a ser alcohólicos).

Esas relaciones pueden ser causales. Por ejemplo, comprendemos que accionar el interruptor hacia arriba enciende la luz. Pero ese entendido se basa en más de una observación; también tenemos una teoría para explicar dicha relación (accionar el interruptor cierra un circuito que permite que la corriente eléctrica fluya a través de la ampolla, calentando los filamentos hasta que generan luz). Nuestra teoría podría estar equivocada, pero en este caso confiamos en ella, pues sabemos

que muchas personas han puesto a prueba la teoría de la electricidad y ésta predice los resultados con certeza.

Pero —y este es el punto clave—, si bien no puede haber causalidad sin correlación, la correlación en sí y por sí sola no basta para demostrar la causalidad. El que dos cosas estén relacionadas no supone que una sea la causa de la otra. Retornando a nuestra investigación imaginaria sobre abuso infantil y alcoholismo, supongamos que efectivamente encontramos que aquellos sujetos que fueron abusados cuando niños tienen más probabilidades de ser alcohólicos que los que no sufrieron abuso (en otras palabras, imaginemos que tenemos evidencias del tipo mostrado en la Tabla 4). Esos hallazgos no constituyen una prueba de que el abuso infantil sea causa de alcoholismo en la adultez. Puede ser que la relación abuso infantil-alcoholismo sea *espuria*, es decir, que una tercera variable sea la causa de la variación tanto en el abuso como en el alcoholismo. Por ejemplo, es posible que haya una mayor tasa de abuso infantil en las familias pobres y que las personas que crecieron en la pobreza tengan mayores probabilidades de ser alcohólicas. Si ampliamos la información de la Tabla 4 preguntando a los sujetos de nuestro estudio si acaso fueron criados en una familia pobre, podríamos obtener los resultados (imaginarios) de la Tabla 7.

Tabla 7. Datos imaginarios que muestran que la pobreza durante la infancia, y no el abuso infantil, aumenta las probabilidades de ser alcohólico en la adultez

	Criados en la pobreza		No criados en la pobreza	
	Abusados	No abusados	Abusados	No abusados
Alcohólicos	25% (112)	25% (50)	5% (8)	5% (30)
No alcohólicos	75% (330)	75% (150)	95% (152)	95% (570)
Total	100% (442)	100% (200)	100% (160)	100% (600)

De pronto, lo que parecía una fuerte relación entre abuso y alcoholismo desaparece. En la Tabla 7 observamos que una vez que incluimos la variable pobreza durante la infancia, el abuso infantil no tiene efecto en el alcoholismo adulto. Este método de análisis se denomina *control por una tercera variable*, la que en este caso es la pobreza infantil. Ahora el vínculo entre abuso y alcoholismo parece espurio, porque la correlación entre ambas variables está de hecho explicada por una tercera (pobreza en la infancia). Una vez más, al presentar estas cifras imaginarias, no pretendo apoyar una determinada postura sobre el nexo entre pobreza, abuso infantil y alcoholismo⁷. Más bien, sólo quería demostrar que, incluso cuando calculamos correctamente los porcentajes, una relación aparente entre dos variables puede desaparecer en lo espurio.

Por el contrario, una genuina relación causal no es espuria. Pero esto plantea un problema lógico enorme: nunca podremos demostrar de manera absoluta que una relación no es falsa, porque siempre es posible que alguna variable sin considerar, aunque la hubiéramos tomado en cuenta, revele que el nexo es espurio. Por ello, siempre podremos afirmar que la causalidad no puede ser demostrada de manera absoluta. En esto basó su defensa, durante décadas, una determinada empresa tabacalera, insistiendo en que los estudios sobre el vínculo entre fumar y el cáncer pulmonar no demostraban que fumar provoca cáncer. En rigor, estaban en lo correcto. Desde luego, si persistimos con esa lógica, no podemos saber con certeza absoluta que accionar el interruptor es lo que hace que la ampolleta se encienda.

Cualquier afirmación que indique que se ha identificado una relación causal debe revisarse cuidadosamente. Ya he sugerido algunas maneras de poner a prueba dichas aseveraciones. Una de ellas es *exigir algún tipo de teoría*, es decir, argumentos sobre el proceso causal que

⁷ Para una revisión bibliográfica sobre el tema, véase Willie Langeland y Christina Hartgers, "Child Sexual and Physical Abuse and Alcoholism", *Journal of Studies on Alcohol* 59 (1998): 336-348.

vincula dos variables (la corriente eléctrica fluye a través de un circuito cerrado; el humo del tabaco irrita los tejidos pulmonares).

También es posible *identificar terceras variables y chequear* si tal vez son fuentes de falsedad. Además, podemos *acumular evidencias* realizando más estudios. A pesar de que cualquier investigación tiene sus limitaciones, si reunimos varios estudios —cada uno con sus propias restricciones— cuyos hallazgos se apoyen unos con otros, la confiabilidad de los resultados aumenta. Esta es la razón de por qué las evidencias acerca de la relación entre fumar y el cáncer de pulmón hoy resultan abrumadoras. Existe abundante literatura basada en diferentes métodos: desde conocer el historial de fumador de quienes padecen cáncer al pulmón, comparar la proporción de muertes causadas por este cáncer entre fumadores y no fumadores, hasta provocar cáncer pulmonar en animales de laboratorio exponiéndolos al humo del tabaco, etc. Mientras cada uno por separado tiene sus propias limitaciones, los estudios que utilizan todos estos métodos generan resultados que apoyan las evidencias del nexo entre fumar y cáncer pulmonar, suministrando pruebas contundentes de una relación causal.

Un problema similar se puede observar cuando las personas atribuyen una determinada causalidad después del dato. En esos casos identificamos una relación, y debido a que una variable precede lógicamente a otra, asumimos que debe ser su causa. Esta cadena de pensamientos combina razonamientos tanto correctos como equivocados. El razonamiento correcto se refiere a que una causa debe ocurrir antes que su efecto; si podemos demostrar que B ocurre después de A, entonces sabemos que B no es la causa de A. La falacia es que A no es necesariamente la causa porque precede a B —recordemos que la vinculación entre A y B puede ser espuria.

La mayoría de los heroinómanos dicen haber fumado marihuana en algún momento previo a consumir heroína. No obstante, ¿podemos llegar a suponer que el consumo de marihuana provoca el uso de heroína? Por un lado, el hecho de que fumar marihuana suele preceder

al uso de heroína no es prueba de su causalidad; mal que mal, casi todos los adictos también consumieron helados algún tiempo antes de consumir heroína, y no por ello decimos que los helados son causa de adicción. Pero, indudablemente, esa crítica es imperfecta: la marihuana, al contrario del helado, es ilegal, y es razonable sospechar que utilizar una droga ilegal fomente el uso de drogas más duras. Por otra parte, la relación entre fumar marihuana y consumir heroína es muy débil: sólo un porcentaje muy bajo de quienes fuman marihuana se hacen adictos a la heroína.

Actualmente, algunos críticos al uso de drogas intentan minimizar el tema de la causalidad declarando que la marihuana es la “puerta de entrada” a drogas más duras. Con esto quieren decir que es posible que el uso de marihuana no provoque adicción a la heroína, pero puede constituirse en una puerta por la cual atraviesa la mayoría de los consumidores de heroína en su camino a la adicción⁸. Sin embargo, esta analogía es ambigua, pues no especifica la naturaleza del vínculo entre marihuana y heroína. Después de todo, ¿qué es una puerta de entrada? ¿Debemos imaginar una verja que, de algún modo, podemos mantener cerrada? En otras palabras, si pudiéramos evitar que las personas prueben la marihuana, ¿eso garantizaría que más adelante no probarán la heroína? ¿O es que la puerta de entrada es una vía ya transitada, entre muchas otras rutas alternativas, de modo que mantenerla cerrada no tendrá gran efecto? ¿Y qué debíamos pensar sobre los consumidores de marihuana que no se vuelven adictos? Seguramente, muchos adictos a la heroína han atravesado la puerta de entrada de la marihuana, pero relativamente pocos de los que han transitado por esa puerta se han vuelto adictos. El concepto de “puerta de entrada” es demasiado impreciso para ayudarnos a entender los problemas vinculados a la droga o para sopesar las distintas políticas para combatirla.

⁸ Para una detallada discusión sobre el concepto de puerta de entrada, véase Robert J. MacCoun y Peter Reuter, *Drug War Heresies* (New York: Cambridge University Press, 2001), pp. 345-351.

Al reflexionar acerca de la causalidad, hay que ser muy cuidadoso.

Estos temas apenas comienzan a considerar la complejidad de las correlaciones entre variables. El estudio formal de la estadística —el contenido de la mayoría de los capítulos de los textos de estadísticas— se orienta preferentemente a medir la fuerza de las relaciones entre variables. (En esos textos, el término “correlación” también tiene un significado técnico más restringido, como una manera específica de reflexionar sobre esas relaciones y sus mediciones). Gracias a la amplia disponibilidad de computadores muy poderosos y softwares estadísticos de fácil manejo, esos cálculos tan sofisticados se han vuelto cada vez más comunes. Los procedimientos estadísticos, que hace treinta años requerían de enormes computadores, disponibles sólo en algunas universidades, ahora pueden realizarse en un computador convencional de un típico estudiante. Hoy por hoy, casi cualquiera puede generar sus propias estadísticas de gran sofisticación, aunque no las entienda. Con el objetivo de establecer relaciones cada vez más complejas entre un mayor número de variables, esta capacidad ha complicado de manera creciente los análisis estadísticos, utilizando mediciones que pocas personas son capaces de comprender cabalmente.

No obstante, los principios básicos de las correlaciones entre variables no son tan difíciles de entender. Debemos buscar las pautas y patrones que revelan potenciales relaciones y las evidencias de que las variables realmente están vinculadas. Pero al encontrar esas relaciones, no debemos apresurarnos a sacar conclusiones respecto a la causalidad. Más bien, debemos sopesar la fuerza de la relación y la plausibilidad de nuestra teoría, y siempre hay que intentar disminuir la posibilidad de lo espurio.

CONFUNDIDOS CON LOS GRÁFICOS

Como hemos visto, las palabras —y las cifras— pueden ser muy confusas. Quizás sea tiempo de abordar un enfoque que parece más

elemental y que se entiende con mayor facilidad. A veces utilizamos el término “visualizar” para referirnos a nuestra habilidad de captar información. Por ello, veamos si las estadísticas pueden transformarse de modo de presentarlas más claramente.

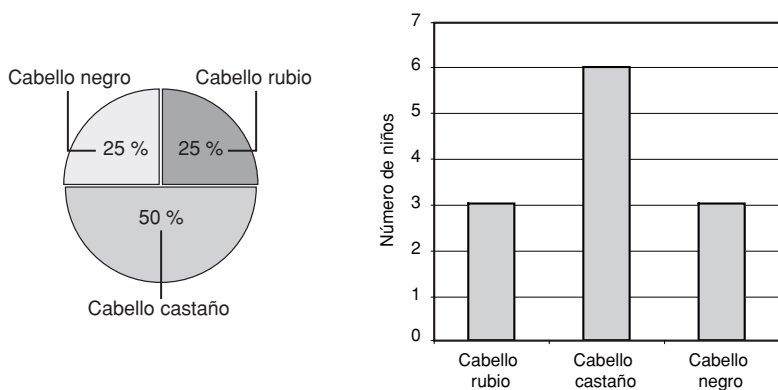


Figura 1. Un simple gráfico de torta (a la izq.) y un gráfico de columnas (a la der.) que representan el color del cabello de un grupo imaginario de niños.

Los gráficos —circulares o en forma de torta; barras o columnas; o de líneas— están entre las formas más conocidas para mostrar información estadística. La idea central es representar los números como imágenes. El estándar básico para calificar los gráficos es extraordinariamente simple: una presentación exacta debe contemplar proporciones visuales equivalentes a las proporciones numéricas que se muestran. Por ejemplo, consideremos un grupo de 12 niños donde 3 tienen cabello rubio, 6 castaño y 3 negro. Podríamos presentar esta información en un gráfico de torta o de columnas (Fig. 1).

En ambos gráficos, las cifras son traducidas a equivalentes espaciales. La mitad de los niños tiene cabello castaño, por lo que su tajada de la torta equivale a la mitad; asimismo, dado que la mitad de los niños está constituida por quienes tienen cabello rubio y los que lo

tienen negro, la columna que representa el castaño es el doble de alta que las demás. En ambas formas de mostrar los datos, las proporciones visuales reflejan las proporciones numéricas.

Efectos perjudiciales de la estética

Los gráficos parecen ser tan obvios y comprensibles que podría pensarse que es difícil echarlos a perder. Si embargo, la triste realidad es otra. Es muy fácil arruinarlos, e incluso en los últimos años se ha hecho aún más sencillo. En gran parte, esto se debe a que los malos gráficos están dominados por la estética. Las personas quieren que sus gráficos impacten, que llamen la atención. Por ejemplo, los gráficos de la Figura 2 parecen poco interesantes, pues no expresan diferencias dramáticas. En contraste, los de la Figura 3, que presentan exactamente la misma situación, son más llamativos.

Ambos pares de gráficos sólo difieren en las escalas del eje vertical. Los dos gráficos del primer par (Fig. 2) están dibujados para mostrar el rango completo de las variables posibles, comenzando con cero en el extremo inferior, hasta un valor algo superior a la cifra mayor graficada. Estos gráficos tienen la virtud de mantener la proporción visual, fiel a la proporción numérica. El problema es que no se leen con facilidad: es muy difícil notar las diferencias entre las barras verticales o la fluctuación entre los puntos del gráfico de líneas.

Casi siempre utilizamos gráficos porque queremos mostrar diferencias o un cambio; generalmente son las diferencias —y no las semejanzas— las que relatan la historia. Si bien, en teoría, el gráfico más fiel a la realidad es el que muestra una escala completa en el eje vertical (Fig. 2), esa escala puede hacer que las diferencias no sean tan evidentes. Una solución muy conocida y utilizada es truncar el gráfico, o sea, cortar la porción inferior de la escala del eje vertical, amplificando y exagerando engañosamente el cambio, generando gráficos como los de la Figura 3. Este enfoque no es ilegítimo, siempre y cuando se rotulen

claramente los valores, de modo que el lector comprenda que el gráfico fue truncado. Al truncar la escala, se logra un gráfico visualmente más interesante; al enfatizar las diferencias y los cambios, se destacan las características de la información.

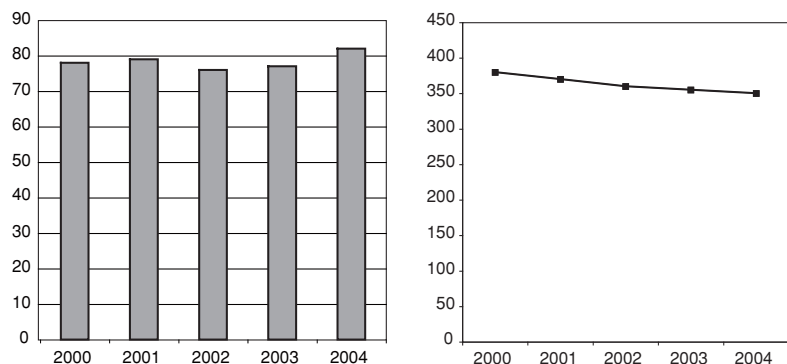


Figura 2. El valor 0 en la base de un gráfico de columnas (a la izq.) y un gráfico de líneas (a la der.) minimiza las diferencias.

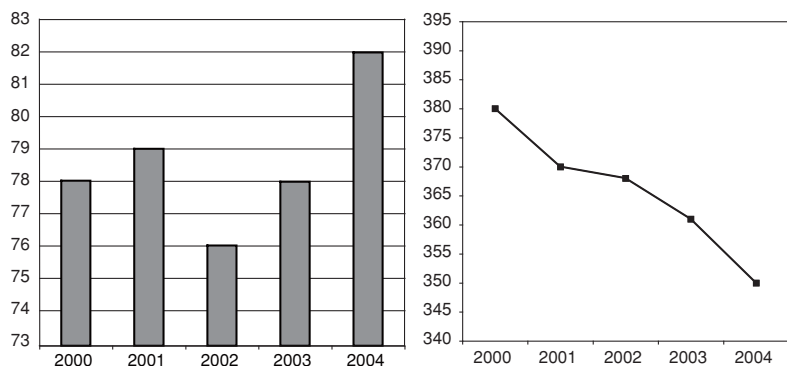


Figura 3. Gráfico de columnas (a la izq.) y gráfico de líneas (a la der.) cuyas bases han sido modificadas para destacar las diferencias.

Algunos expertos sostienen que no es aceptable truncar los gráficos de columnas (porque las alturas relativas de éstas suelen contener la información clave), pero sí lo es en los gráficos de líneas (siempre que

se rotulen con claridad) cuando la atención está puesta en el patrón de cambios. Por ejemplo, diariamente muchos periódicos publican gráficos de líneas que muestran las fluctuaciones del mercado bursátil del día anterior; estos gráficos están truncados y todos los días cambian sus ejes verticales para enfatizar las variaciones de la actividad del día anterior. Si bien sus diferentes escalas significan que el gráfico publicado un día no puede compararse con el de otro, es de suponer que los lectores de periódicos entiendan estas limitaciones.

Varios cambios recientes han acrecentado la influencia negativa de la estética en el diseño de gráficos. Uno de los más importantes es la amplia disponibilidad de softwares para generar gráficos. Hoy en día, los gráficos de torta están por todas partes y el antiguo método de dibujarlos a mano con un compás, una regla y un transportador está convirtiéndose en algo del pasado. Por el contrario, hacer gráficos de torta mediante un computador es muy sencillo y las personas no lo piensan ni un minuto —lo que se nota.

Observemos la Figura 4. Este gráfico de torta acompañaba una nota del periódico referida a niños que habían sido raptados por familiares, y muestra las diferencias en la duración de esos secuestros.

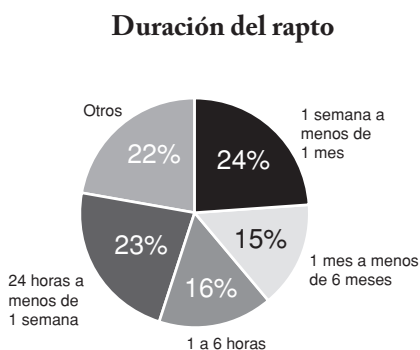


Figura 4. Confuso gráfico de torta que muestra la duración de los raptos de menores por parte de familiares. (Fuente: Ryan Cormier, "A Missing Child, Unbearable Pain", *The News Journal, Wilmington, DE*, 13 julio 2003, p. A9; utilizado con permiso).

El gráfico es confuso por dos motivos. Primero, las tajadas de la torta no siguen ningún orden evidente. Si recorremos el gráfico en el sentido de las manecillas del reloj, leemos: “1 semana a menos de 1 mes”, “1 mes a menos de 6 meses”, “1 a 6 horas”, “24 horas a menos de 1 semana” y “Otros”. La secuencia no tiene ningún sentido; el gráfico confunde más de lo que aclara.

En segundo lugar, el gráfico contiene una extraña categoría: “Otros”.

¿Qué incluye esto? Este gráfico se basó en una tabla de datos publicada por un informe oficial⁹. Al revisar los datos originales, se encontró que el 22% listado como “Otros” en el gráfico incluía la siguiente información:

Secuestros que duraron menos de 1 hora	3%
Secuestros que duraron 7 a 24 horas	4%
Secuestros que duraron más de 6 meses	6%
Niños ubicados pero no devueltos	6%
Casos sin información	3%
Total	22%

Es probable que la persona que hizo el gráfico agrupara estas categorías porque eran cifras relativamente pequeñas, pero el resultado es incoherente: la categoría “Otros” incluye episodios breves, algunos muy largos y otros de los cuales no se sabe nada. No tiene sentido agrupar estas categorías en una sola.

A partir de la misma información pueden derivarse gráficos de torta más útiles y diferentes. Por ejemplo, la Figura 5 nos entrega un gráfico que ofrece categorías en un orden lógico (desde el intervalo

⁹ Heather Hammer, David Finkelhor y Andrea J. Sedlak, “Children Abducted by Family Members: National Estimates and Characteristics”, *National Incidence Studies of Missing, Abducted, Run-away, and Thrownaway Children* (U.S. Office of Juvenile Justice and Delinquency Prevention, October 2002).

más breve al más largo), no mezcla distintos casos en una misma categoría y muestra un patrón claro: los secuestros de menores por parte de familiares tienden a ser breves.

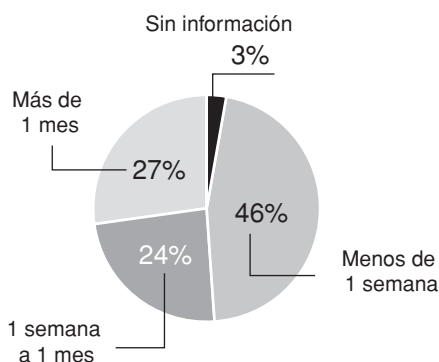


Figura 5. Gráfico de columnas mejorado que muestra duración del secuestro de menores por parte de familiares.

Otro problema es que la mayoría de los softwares para generar gráficos ofrecen una variedad de opciones “atractivas”: por ejemplo, se puede inclinar un gráfico de torta para que se vea desde un ángulo, lo que transforma la torta circular en un óvalo. Esto que puede ser una mejora desde el punto de vista estético, debilita absolutamente la exactitud y utilidad del gráfico, porque los segmentos formados por ángulos iguales desde el centro de un óvalo pueden no tener áreas equivalentes. Por ello, este método de hacer más interesante y atractivo un gráfico de torta viola el principio central de las imágenes estadísticas (proporcionalidad del espacio y las cifras) y entrega una impresión distorsionada, exagerada e inexacta. Otras “mejoras” provocan efectos negativos similares, como mostrar al lector el borde de un gráfico de torta que ha sido inclinado (con lo que se crea la impresión de un disco tridimensional), lo que exagera la importancia visual de las tajadas que pueden verse por el lado.

Asimismo, para remarcar las diferencias en la información, la mayoría de los softwares para generar gráficos trunca automáticamente las escalas del eje vertical en los gráficos de columnas y de líneas. Por ejemplo, no tuve problemas para generar los gráficos de la Figura 3; ingresé algunos datos imaginarios en un conocido programa de hoja de cálculo, le indiqué que realizara un gráfico de columnas y otro de líneas, y ambos aparecieron automáticamente en mi pantalla. En cambio, me tomó mucho más tiempo crear los gráficos completos de la Figura 2, porque, aunque se tratara de las mismas cifras, tuve que sortear la opción por defecto del programa y agregar otros comandos. (Me avergüenza admitir que no pude hacerlo solo; alguien me tuvo que explicar la secuencia de los comandos secretos). En resumen, *es más frecuente que los softwares faciliten el dibujo de gráficos distorsionados que el de gráficos proporcionados*. Y, tal como suele suceder, un importante atractivo de hacer las cosas mal es que resulta mucho más fácil.

Los softwares para generar gráficos son un arma de doble filo: aportan facilidad de uso pero también invitan al abuso. Dado que ofrecen formatos por defecto, no hay que pensar demasiado —o tener mucho cuidado— para crear un gráfico. Como resultado, somos bombardeados continuamente con gráficos innecesarios. Puede ser que una buena imagen valga más que mil palabras, pero la información de un gráfico de torta también se puede expresar en una sola frase. Los múltiples menús de pitos y campanas que ofrecen algunos softwares —como inclinar los gráficos de torta para mostrar su borde, transformar un gráfico de columnas en figuras tridimensionales o convertir el fondo de un gráfico en una colina— suelen distorsionar las proporciones visuales, de manera que se pierde el propósito central del gráfico: ayudar a visualizar las proporciones relativas de los datos presentados.

Un fenómeno parecido se observa en la tendencia de los periodistas a utilizar elementos pictóricos para “avivar” los gráficos. De este modo, en vez de círculos repartidos en segmentos, o meras barras verticales de diferentes alturas, encontramos gráficos de formas muy

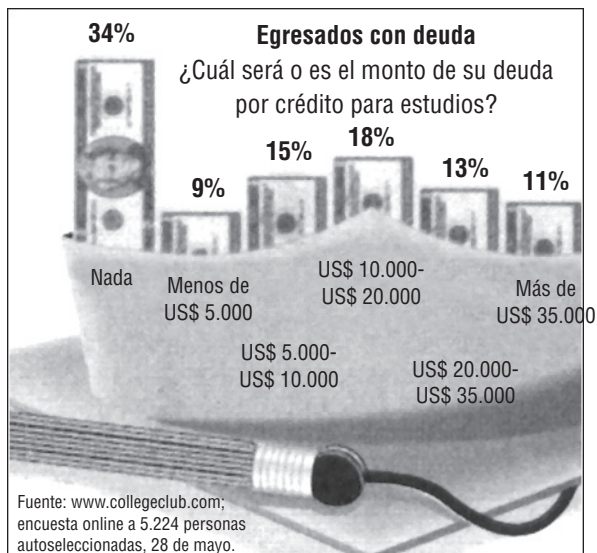
variadas (Fig. 6). El periódico *USA Today* ha popularizado este estilo de gráficos, y aun cuando esas ilustraciones sean muy agradables a la vista, el resultado es pésimo en lo que respecta a transmitir información. A menudo es difícil comprender qué elementos espaciales en la imagen corresponden a las cifras que se supone representan. Uno de los más destacados teóricos sobre gráficos, Edward J. Tufte, ha acuñado el memorable término “basura gráfica” (*chartjunk*) para referirse a todos los elementos que no entregan información alguna y que llenan de desperdicios muchos gráficos¹⁰. En casos extremos, al no permitir descifrar su significado, la basura gráfica hace inservible a un gráfico.

La Figura 6 es un ejemplo típico de las pequeñas ilustraciones que aparecen en las esquinas inferiores de las páginas de los periódicos. Este infograma da cuenta de las respuestas de personas “autoseleccionadas” a quienes se les habría preguntado: “¿Cuál será o es el monto de su deuda por crédito para estudios?”. Desafortunadamente, la mayoría de los resultados de una encuesta online no tienen significado alguno, pues la muestra no es representativa¹¹. No todas las personas tienen acceso a internet; sólo una parte de las que sí lo tienen, se encuentran con una encuesta y sólo algunas (¿cuáles?, ¿quién sabe?) se toman la molestia de responder. Ni siquiera sabemos si la muestra representa de manera adecuada a la población en estudio (que, en este caso, puede referirse al total

¹⁰ Todos se pueden beneficiar con la lectura del magnífico libro de Edward J. Tufte, *The Visual Display of Quantitative Information* (Cheshire, Conn.: Graphics Press, 1983). La larga bibliografía relativa al tema incluye a Jones, *How to Lie with Charts*; y Howard Wainer, *Visual Revelations* (New York: Springer-Verlag, 1997).

¹¹ La encuesta por muestreo es un procedimiento aleatorio de selección de casos orientado a representar, en una escala más pequeña, la estructura de la población que se quiere conocer. La idea es que todos los casos tengan una probabilidad conocida y distinta de cero de ser encuestados. Por esta razón, muchas encuestas por internet no son representativas de la población, ya que al carecer de un diseño muestral adecuado, omiten que hay una parte de la población que no accede a la web, así como tampoco se conoce la probabilidad efectiva que tiene cada caso de ser seleccionado. Para profundizar en este tema, véase Manuel Vivanco, *Muestreo Estadístico, Diseño y Aplicaciones* (Santiago de Chile: Editorial Universitaria, 2005) (N. de la T.).

de alumnos o ex alumnos de una universidad, lo que tampoco podemos comprobar). En resumen, esta información es totalmente inservible.



Por Lori Joseph y Marcy E. Mullins, USA TODAY

Figura 6. El significado de este gráfico se pierde entre tanta basura gráfica.
(Fuente: *USA Today*, 6 agosto 2002, p. 1A; © USA Today, reimpresso con autorización).

Y ésta es la buena noticia. El gráfico en la Figura 6 contiene columnas, representadas por estilizados billetes que parecen salir de un birrete de graduación. (Confieso que me costó mucho descubrir lo que supuestamente representaba esto. Me confundí con el objeto que parece un lápiz con un cordón eléctrico, hasta que me di cuenta que se suponía era la borla del birrete). Pero no hay manera de saber dónde comienzan las columnas en el gráfico; es claro que de algún lugar del interior del birrete, ¿pero de dónde? Podemos distinguir que 18% es más que 13%, lo que a su vez es más que 11%, pero eso ya lo sabemos. Lo que

no obtenemos en este ejemplo es un sentido visual evidente de las proporciones relativas de estas cantidades, porque una parte de las columnas no está a la vista y los contornos desiguales del birrete sugieren que la proporción oculta podría ser distinta entre una barra y otra.

Para hacer las cosas aún más confusas, el ojo del observador es atraído hacia distintos rasgos del birrete que podrían representar la base de las columnas; pero al revisarlo, constatamos que no es así. Aparte de la orilla del birrete, tenemos una sombra curva, el cordón y la borla, el borde del birrete y la sombra de éste. Esto sí que es basura gráfica: obliga al observador a descifrar el significado a partir de las características de la ilustración, esfuerzo que no se compensa, porque esos rasgos ni siquiera se relacionan con la información que el gráfico pretende entregar. En resumen, este gráfico proporciona información irrelevante de una manera ilegible, lo que es un problema cada vez más frecuente en la prensa. Casi a diario, los periódicos nos dan ejemplos incluso peores de esto, con gráficos de torta de formas irregulares, gráficos con columnas de dimensiones indeterminadas, etc.

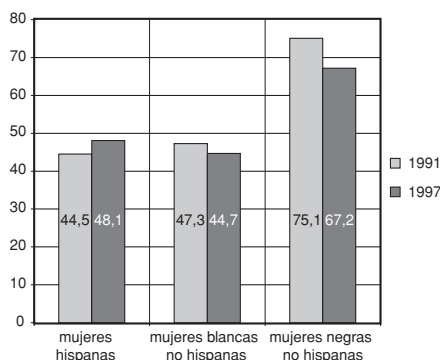
Es posible que no hubiera la intención de hacerlo, pero la combinación de consideraciones estéticas y la gráfica asistida por computador pueden hacer que gráficos impecablemente rotulados se tornen confusos e inservibles. Consideremos la Figura 7, que reproduce gráficos de barras publicados por primera vez por la Asociación Americana de Sociología (American Sociological Association).

Al reimprimir estos gráficos de columnas, he conservado las proporciones relativas expresadas en el artículo original. Los gráficos entregan información desglosada por grupos étnicos (hispanos, blancos y negros) relativa al cambio acaecido entre 1991 y 1997 en la conducta sexual de estudiantes de enseñanza media; el superior muestra el caso de las mujeres y el inferior, el de los hombres. Cada gráfico por separado es claro. Pero la confusión surge cuando se los observa en conjunto. La configuración de la página asigna alrededor de la mitad más de espacio a las columnas del gráfico de las mujeres, mostrando columnas

más cortas en el caso de los varones. Basándonos en la altura de las barras, tenemos la impresión visual de que los varones tuvieron menos actividad sexual, siendo que cuando se comprueban los porcentajes, los hombres tenían una vida sexual más activa. El efecto, en apariencia provocado por la negligente configuración de la página, es que tenemos una percepción visual exactamente opuesta a lo que indican los datos.

Figura 3

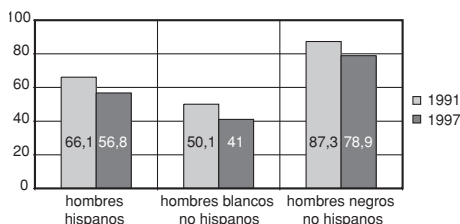
Porcentaje de estudiantes de enseñanza media sexualmente activas según raza



Fuente: tabulaciones de los autores a partir de Encuestas de Conductas de Riesgo en Jóvenes realizadas en 1991 y 1997

Figura 4

Porcentaje de estudiantes de enseñanza media sexualmente activos según raza



Fuente: tabulaciones de los autores a partir de Encuestas de Conductas de Riesgo en Jóvenes realizadas en 1991 y 1997.

Figura 7. Doble estándar evidenciado gráficamente. (Fuente: Barbara Risman y Pepper Schwartz, "After the Sexual Revolution", *Contexts* 1, no. 1 [febrero 2002]: 19, © American Sociological Association, 2002; reimpresso con autorización).

Y este es un ejemplo menor. Es fácil encontrar expresiones gráficas de información numérica en las que la estética y el intento por impresionar han desplazado de tal modo a la información, que son casi imposibles de descifrar. Por ejemplo, las Figuras 8 y 9 reproducen dos presentaciones gráficas —me resisto a llamarlas gráficos— de un “atlas del comportamiento sexual humano”.

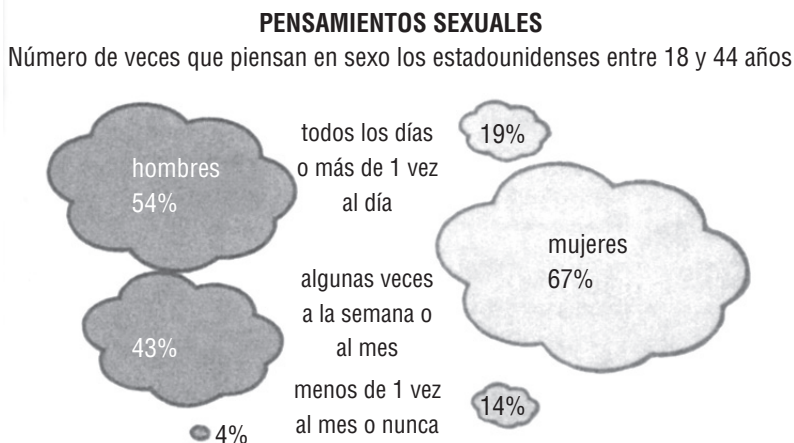


Figura 8. Gráfico en que las proporciones visuales no se relacionan con las cifras que representan. (Fuente: Judith Mackay, *The Penguin Atlas of Human Sexual Behavior* [New York: Penguin, 2000], p. 21; ilustración © Myriad Editions, Ltd., 2000, utilizada con autorización de Viking Penguin).

La ilustración de la Figura 8 utiliza dibujos con forma de nube para mostrar cuán a menudo los hombres y mujeres jóvenes piensan en sexo, con nubes pequeñas representando los porcentajes menores. (¿Por qué nubes? No podemos asegurarlo; la imagen original está impresa sobre un mapa del hemisferio sur, así que a lo mejor son nubes en el cielo, o tal vez pretenden evocar esos globos de palabras en que los caricaturistas escriben los pensamientos de sus personajes). Pero el tamaño de las nubes no es ni remotamente proporcional a las cifras representadas.

Por ejemplo, 67% es casi 3 veces más grande que 19%, pero la nube del 67% es muchas veces más grande que la del 19%. De hecho, este gráfico dificulta la posibilidad de visualizar la escala de diferencias entre las distintas cifras. (A pesar de que no podemos saberlo con certeza, es probable que las formas irregulares de las nubes se deban a un clásico error cometido por los creadores, al intentar dibujar el ancho y la altura de las nubes en proporción a las cifras que representan¹¹. El problema es que nuestro ojo no ve el ancho y la altura, sino el área total: si la nube B es el doble de ancha o alta que la nube A, el área de la nube B se ve cuatro —y no dos— veces más grande que la nube A).

EL ALCOHOL DIFICULTA EL EMBARAZO

Porcentaje de mujeres en Dinamarca, con distintos niveles de consumo de alcohol semanal, que se embarazan en los 6 meses siguientes al abandono de métodos anticonceptivos.

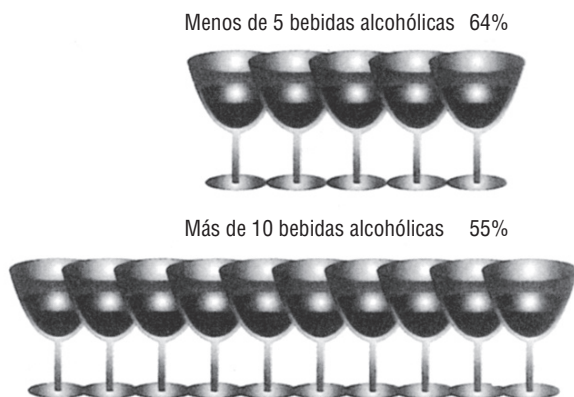


Figura 9. Ilustración que grafica la escala y no los datos. (Fuente: Judith Mackay, *The Penguin Atlas of Human Sexual Behavior* [New York: Penguin, 2000], p. 47; ilustración © Myriad Editions, Ltd., 2000, utilizada con autorización de Viking Penguin).

¹¹ En Huff, *How to Lie with Statistics*, pp. 66-73, hay una interesante introducción sobre el principio de que las áreas de las figuras bidimensionales deben ser proporcionales al valor de las cifras que representan.

Mucho peor es lo que vemos en la Figura 9, “El alcohol dificulta quedar embarazada”. Los datos, muy sencillos, muestran que las mujeres que beben más alcohol tendrían mayores dificultades para embarazarse. Mientras un 64% de las mujeres que bebían menos de 5 bebidas alcohólicas a la semana quedaban embarazadas en los 6 meses siguientes a haber dejado los métodos anticonceptivos, sólo un 55% de las que afirmaron beber más de 10 bebidas alcohólicas semanales dijeron embarazarse en el mismo período.

Para ilustrar estos datos, nos muestran filas de copas de vino. La tasa de 64% de embarazo entre las mujeres que bebían menos, está representada por 5 copas; mientras que la tasa de 55% de embarazo entre quienes bebían más, se ilustra con 10 copas. A primera vista, esto es confuso: ¿por qué utilizar la imagen más pequeña en el caso de la tasa de embarazo más alta? Pero luego todo se aclara: menos de 5 bebidas alcohólicas se ilustra con 5 copas; 10 ó más bebidas, con 10 copas. Esta gráfica debe ser para sujetos (quizás bebedores excesivos) que necesitan visualizar que 5 bebidas son menos que 10. Esto evoca las pinturas rupestres, cuando no se conocían los números escritos y los pastores llevaban un supuesto recuento de sus animales, dibujándolos uno por uno. De esta manera, al optar por representar la variable independiente (el número de bebidas alcohólicas), más que la variable dependiente (la tasa de embarazo), la ilustración de la Figura 9 claramente abandona cualquier esfuerzo por transmitir información.

Selectividad

En general, puede afirmarse que las consideraciones estéticas son las que más daño causan en los gráficos contemporáneos. También hay que reconocer la posibilidad de que, en un esfuerzo por presentar interpretaciones tendenciosas, los creadores de los gráficos manipulen deliberadamente la estética. Pero no nos adelantemos. Cabe recordar que los softwares para diseñar gráficos adoptan las opciones por

defecto —como truncar las escalas del eje vertical—, lo que garantiza la distorsión. Parece ser que muchos gráficos, en realidad mal hechos, carecen de una intención definida. Incluso si aceptamos que el gráfico de torta de la Figura 4, la ilustración del birrete de la Figura 6 y las nubes y copas de vino de las Figuras 8 y 9 son ejemplos de lo que estamos hablando, no es fácil detectar un intento de engaño tras esas muestras.

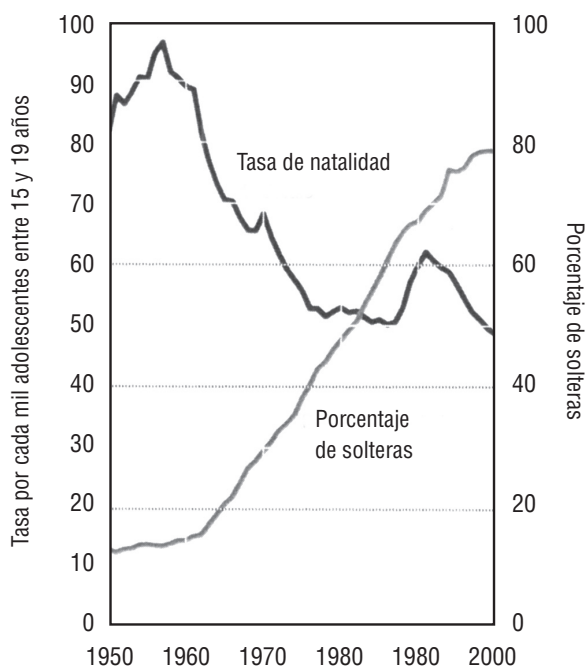
Sin embargo, en otros casos, la opción de cómo presentar la información busca reforzar un punto de vista determinado. Cada gráfico —como cada dato estadístico— refleja una serie de opciones: ¿qué se mostrará?, ¿cómo se mostrará? Es inevitable algún grado de selectividad, pero se puede caer en el abuso.

En la Figura 10 vemos dos líneas en un gráfico —bastante bueno— de una publicación oficial sobre la tasa de natalidad en adolescentes. Una línea muestra la tasa de natalidad (el número de nacidos por cada mil mujeres en un determinado grupo etario) en la segunda mitad del siglo XX. Si bien la línea muestra algunas fluctuaciones, se evidencia que la tasa de natalidad en adolescentes entre 15 y 19 años declinó durante este período: tuvo su nivel máximo en 1957 con 96,3 partos por cada mil mujeres, pero disminuyó a 48,7 en 2000, “el nivel más bajo registrado en el país”¹². Esto podría parecer sorprendente, dada la frecuencia con que los medios muestran historias sobre embarazos y madres adolescentes.

Pero consideremos la segunda línea, que muestra el porcentaje de madres adolescentes que no estaban casadas. En 1957, sólo 13,9% de las embarazadas adolescentes eran madres solteras, pero este número aumentó a 78,7% en 1999. Es decir, a pesar de que ha disminuido la tasa de natalidad en adolescentes, ha crecido el porcentaje de embarazos adolescentes fuera del matrimonio. El problema no es que la tasa de natalidad en adolescentes haya aumentado, porque no ha sido así.

¹² Stephanie J. Ventura, T.J. Mathews y Brady E. Hamilton, “Births to Teenagers in the United States, 1940-2000”, *National Vital Statistics Reports* 49, no. 10 (September 25, 2001): 1.

La preocupación actual es que un porcentaje creciente de nacimientos se da en adolescentes solteras, las que suelen enfrentar mayores dificultades para criar y cuidar a sus hijos. En las décadas anteriores, las parejas se casaban más jóvenes (muchas veces debido a que la novia estaba embarazada); hoy en día se tiende a posponer el matrimonio, incluso cuando la mujer decide llevar a término su embarazo.



NOTA: los datos para 2000 son preliminares.

Figura 10. Las dos líneas en este gráfico cuentan una historia complicada.
(Fuente: Stephanie J. Ventura, T.J. Mathews y Brady E. Hamilton, "Births to Teenagers in the United States, 1940-2000", *National Vital Statistics Reports* 49, no. 10 [25 septiembre 2001]: 2).

¿Entonces qué deberíamos pensar acerca de la tendencia de la natalidad entre las adolescentes? Ninguna de las dos líneas relata la historia

completa. Podríamos interpretar que la tendencia a la baja de la tasa de natalidad adolescente indica que las cosas están mejor, pero habrá quienes planteen que el aumento de la proporción de nacimientos fuera del matrimonio demuestra que las cosas están empeorando. En este caso, las cifras no son incorrectas ni engañosas; ambas líneas en el gráfico se basan en información fidedigna (recopilación de datos federales sobre los nacimientos registrados). Pero ninguna línea, por sí sola, da una idea clara de lo que está sucediendo; para tener una noción de cuán complejos son los cambios que experimenta la sociedad, debemos mirarlas en conjunto.

Pero es fácil imaginar a los defensores de un punto de vista particular sobre el embarazo adolescente —que estamos o no frente a una situación crítica— eligiendo un gráfico con sólo una de esas líneas, con tal de defender su postura. La tendencia que esas personas decidan destacar —la caída en la tasa de natalidad en adolescentes o el aumento del porcentaje de embarazos en adolescentes solteras— inevitablemente influirá en nuestra opinión acerca del tema.

La cantidad de información entregada influye de manera similar. Aun cuando la Figura 10 muestra, en términos generales, que la tasa de natalidad en adolescentes bajó entre 1950 y 2000, también se observa que la tendencia no varió demasiado entre 1975 y 1985. E incluso, entre 1986 y 1991, esta tasa creció antes de caer a los valores más bajos del período. Consideremos cómo cambiaría la forma del gráfico si los promotores presentaran sólo la información de 1975 a 1985, de 1985 a 1991, o incluso de 1975 a 2000. Esta variación sugiere que hay que ser cuidadoso al prestar demasiada atención a gráficos que exhiben cambios en un período muy corto de tiempo, los que pueden terminar siendo sólo insignificantes fluctuaciones dentro de una tendencia que se mantiene a largo plazo¹³.

¹³ Para un ejemplo de tal crítica, véase Bjorn Lomborg, *The Skeptical Environmentalist* (Cambridge: Cambridge University Press, 2001), pp. 22-23.

Siempre vale la pena preguntarse si la información fue cuidadosamente seleccionada para promover un punto de vista determinado, y si existen otros datos que pudieran sustentar interpretaciones diferentes. El ejemplo de información sobre el embarazo adolescente debe recordarnos que los temas sociales, como éste, son complejos y multidimensionales.

SOBRE LA CONFUSIÓN

Los tipos de confusión que hemos analizado en este capítulo son particularmente poco afortunados, porque no había necesidad de ellos. Los porcentajes y los gráficos de torta son herramientas muy sencillas: la mayoría de nosotros suele conocerlos desde la enseñanza básica. Quizás sea esa misma familiaridad la que nos hace ser complacientes: suponemos que comprendemos todo y no reconocemos nuestra confusión. O de pronto, quizás, nos damos cuenta que hay algo malo con las cifras (“¡Eso no puede ser cierto!, ¿o sí?”), y sin embargo, no podemos distinguir dónde está el error. La confusión promueve la frustración, el sentido de que un asunto es demasiado complicado y, por lo tanto, nos rendimos (“Nunca entenderé esto, así que no seguiré intentando”).

Pero antes de darnos por vencidos hay una alternativa. En vez de declararnos incapaces, podemos tomarnos un tiempo para tratar de entender qué es lo que está mal. ¿Es un gráfico confuso? Examinémoslo detenidamente. ¿Qué es lo que se está representando y cómo se han llevado los números a imágenes? ¿Las proporciones visuales son un reflejo exacto de las cifras? ¿Falta información clave? ¿Qué aspecto del tema quisiera saber que no me están mostrando? No hay que olvidar que, al menos en teoría, los gráficos están para aclarar las cosas; si usted está confundido, lo más probable es que sea culpa del gráfico.

Asimismo, las estadísticas básicas —promedios, porcentajes y cosas por el estilo— debieran entenderse fácilmente. Si usted se siente confundido o impactado por las cifras mostradas, reflexione un poco sobre

su significado. ¿De adónde provienen esas cifras? ¿Quién las calculó, por qué se dio el trabajo de hacerlo y cómo lo hizo? ¿Sería distinto si las cifras hubieran sido calculadas o presentadas de otra manera?

Quizás no podamos responder algunas de estas interrogantes, pero incluso eso sería información útil. Si no nos han dicho lo suficiente para responder las preguntas más básicas, es una señal de que algo anda mal. *Si quedamos con la impresión de que las cifras nos están llevando hacia un punto de vista determinado, debemos preguntar por qué se eligieron esas y no otras.* Podemos aprender a tratar la confusión como un desafío, más que como una señal de que es hora de renunciar.

Capítulo 3

CIFRAS QUE ASUSTAN

Hace poco, a un grupo de adultos estadounidenses se les solicitó por medio de una serie de encuestas que calcularan el porcentaje de menores que no tenían seguro médico; también se les preguntó sobre la tendencia actual en la tasa de delincuencia juvenil, la tasa de natalidad en adolescentes y el porcentaje de niños criados en familias monoparentales¹. De las respuestas surgió un patrón bastante claro: frente a cada uno de los temas, la gran mayoría —entre 74% y 93% de los encuestados— opinaba que los problemas eran peores de lo que se veía en la realidad. Por ejemplo, el 76% respondió que en los últimos 5 años había aumentado el porcentaje de chicos que vivían en una familia monoparental. Sin embargo, el porcentaje no había cambiado. Alrededor del 66% creía que el porcentaje de adolescentes que cometían delitos violentos había crecido en los últimos 10 años, mientras que otro 25% dijo que éste se había mantenido más o menos inalterado. No obstante, en ambos casos lo observado había sido una disminución. ¿Cómo se explica la tendencia a imaginar que las cosas son peores de lo que realmente son?

Ya que las estadísticas pueden confundir, es natural que generen inquietud. Además, la mayoría de las cifras buscan, si no asustarnos,

¹ Lina Guzmán, Laura Lippman, Kristin Anderson Moore y William O'Hare, "How Children Are Doing: The Mismatch Between Public Perception and Statistical Reality", *Child Trends Research Brief* (July 2003).

por lo menos inquietarnos respecto al mundo que nos rodea. Casi todo lo que tiene algún interés periodístico es una mala noticia; la emisora local de las “noticias felices” podrá concluir su noticiario con una nota de 45 segundos sobre una abuelita que se lanza al espacio desde una avioneta en paracaídas, pero la historia principal mostrará al reportero en la escena de un asalto en un minimercado con resultados fatales. Las estadísticas siguen esa misma pauta: en general, las cifras que perturban y asustan tienen más cobertura periodística que las que dan a conocer noticias positivas o avances. Entonces no debiera sorprendernos que tendamos a exagerar el alcance de los problemas sociales. Estamos acostumbrados al torrente de estadísticas que indican lo que está mal y nos advierten que las cosas son mucho peores de lo que imaginamos.

Esta tendencia a destacar cifras apabullantes refleja la forma en que se muestran los problemas sociales en nuestra sociedad. Quienes buscan llamar la atención sobre algún problema social deben competir con otras causas para que la prensa, los políticos o el público se fijen en él. En medio de reivindicaciones que compiten de manera disonante, deben convencer a la opinión pública de que su problema particular amerita preocupación. Sus afirmaciones apuntan a elementos que nos resultan familiares: el problema está extendido; tiene consecuencias graves; sus afectados son vulnerables y necesitan protección; todos somos potenciales víctimas; el asunto se está agravando. Las evidencias para apoyar esas afirmaciones suelen originarse en la combinación de ejemplos problemáticos con datos estadísticos (tal como vimos en Cap. 1). Quienes buscan generar inquietud, naturalmente privilegian lo negativo; por consiguiente, prefieren estadísticas que asusten y que representen el problema como muy común o grave.

Pero no sólo quienes defienden una causa prefieren los números que asustan. Los medios rastrean los informes estadísticos —por ejemplo, los resultados del censo— en busca de elementos que tengan interés noticioso, es decir, aquellos que suelen interpretarse como más

preocupantes. Y, como veremos en el Capítulo 4, puede que incluso los investigadores y las autoridades consideren que el enfatizar las cifras que asustan hace más interesante su trabajo.

DESCRIBIR PROBLEMAS SOCIALES EN TÉRMINOS ATEMORIZANTES

Cuando los defensores de un determinado problema social lo describen, el primer dato que solemos escuchar es una estadística relativa a su tamaño, por ejemplo el número de casos o de individuos afectados. Las cifras grandes apoyan la idea de que un problema está extendido y que por lo tanto es serio. Otras estadísticas, como la cantidad o el porcentaje de personas aquejadas por el tema en cuestión, transmiten una sensación de riesgo; entregan un cálculo aproximado, con lo cual crean la impresión de que estamos frente a un problema que me amenaza a mí o a alguno de mis seres queridos. Estas cifras aumentan nuestra sensación de vulnerabilidad. Incluso hay estadísticas, como la tasa de crecimiento, que proyectan el problema hacia el futuro, haciéndonos creer que lo que ahora está mal puede llegar a ser aún peor.

Ese tipo de estadísticas son más convincentes cuando describen el mundo de un modo especialmente alarmante. Mientras más extendidos sean el daño y el sufrimiento percibidos, es más probable que lleguemos a creer que ese problema eventualmente nos afectará a nosotros; y mientras mayor sea la posibilidad de que las cosas empeoren, nuestro miedo aumentará. A su vez, ese temor hace más convincentes las afirmaciones de quienes promueven esos datos y, por ende, influirán más en nosotros.

Si las anteriores generaciones de reformistas enfatizaban que la sociedad tenía la obligación moral de ayudar a sus miembros más vulnerables y desdichados, hoy se estimula a la gente a actuar en base al interés personal e individual. Al advertir que un problema puede afectar a cualquiera, democratizamos el riesgo. “NADIE ESTÁ LIBRE

DEL SIDA” era el mensaje de la portada de la revista *Life* de julio de 1985; otra cita jugosa de la época señalaba que “sólo un par de salarios mantienen a muchas familias viviendo fuera de la calle”. Decir que todos somos vulnerables implica que todos lo somos *de igual forma*; esta afirmación minimiza patrones de riesgo bien documentados, al tiempo que fomenta la sensación de un peligro compartido. Si vemos que el sida o la falta de vivienda está afectando a otro grupo de la sociedad que el nuestro, entonces quienes defienden esa postura deberán apelar a nuestro sentido de responsabilidad moral para interesarnos. Pero si hay una amenaza que pone en peligro a todos, entonces todos estaremos incondicionalmente interesados en el problema. Es bastante revelador que la persuasión moderna suela recurrir tan a menudo al interés individualista por uno mismo, antes que a la preocupación por los demás.

Aun cuando un determinado problema no aparente ser una amenaza directa e inmediata, es posible pintar un futuro en el que las cosas serán mucho peores. Una manera de detectar los problemas que se nos vienen encima es detectar tendencias; aun cuando la situación no esté mal en este momento, podemos vislumbrar signos de que se deteriora. Por supuesto, las tendencias más atemorizantes son las que parecen llevarnos directo a la catástrofe. Los cálculos estadísticos de futuros problemas sociales son difíciles de contradecir: aparte de esperar y ver qué sucede, cuesta desacreditar un escenario terrorífico. Pero una mirada a la historia reciente de los pronósticos revela cuán nebulosa puede ser la visión de futuro de los expertos. Las revistas de mi niñez predecían que en el año 2000 el mundo sería muy distinto: los empleados irían al trabajo en helicópteros personales y vehículos propulsados por energía atómica, aunque no se mencionaban los computadores. El historial de quienes defienden la visualización futura de los problemas sociales no es mucho mejor: sólo recordemos los pronósticos de Y2K* respecto al

* El problema del año 2000, conocido como efecto 2000, error del milenio, problema informático del año 2000 (PIA2000) o Y2K, es un error de software causado por la costumbre que habían adoptado los programadores de omitir el año al almacenar fechas, asumiendo que el

colapso social total que sobrevendría tras la falla simultánea de los sistemas computacionales mundiales cuando el calendario cambiara de 1999 a 2000.

En consecuencia, existe una tensión entre la necesidad de una retórica convincente por parte de los defensores —argumentos que pueden hacer que algunas personas se fijen en ciertos problemas sociales— y las limitaciones de las evidencias disponibles. Usualmente esto se resuelve por la vía de ignorar la evidencia a favor de presentar el mensaje más potente. Para los promotores de una idea, que creen que su causa es justa y consideran que las cifras son perfectamente razonables, las estadísticas que asustan tienen un evidente atractivo. Para los medios de comunicación, esas cifras tienen interés noticioso, por lo que son tema para una buena historia. Es por ello que esos datos característicamente enfrentan muy poca resistencia. Este capítulo revisa tres tipos de cifras habitualmente utilizadas para hacer que ciertos problemas sociales se vuelvan atemorizantes: grandes cálculos, tendencias problemáticas y escenarios apocalípticos.

Medir el tamaño de un problema

La cifra atemorizante más sencilla es el cálculo de la magnitud de un problema social, por ejemplo la cantidad de personas involucradas o su costo monetario total. A simple vista, esto parece sencillo: todos hemos contabilizado cosas, por lo que suponemos que para obtener ese resultado, alguien contó algo. Si el resultado de ese conteo es un número grande, tendemos a creer que existe un gran problema. Así de simple.

Pero los problemas sociales son difíciles de contabilizar. Una de las posibilidades es que los casos sean difíciles de identificar e incluso puede que sea complejo definir y medir lo que se desea contar. Por

software sólo funcionaría durante los años cuyos nombres comenzaran con 19. Lo anterior tendría como consecuencia que después del 31 de diciembre de 1999, sería el 1 de enero de 1900 en vez del 1 de enero de 2000 (N. de la T.).

ejemplo, consideremos las afirmaciones tan publicitadas respecto a que cada año mueren entre 44 mil y 98 mil pacientes en hospitales estadounidenses debido a errores médicos perfectamente evitables. Estas son cifras muy atemorizantes, no sólo por su dimensión sino porque se supone que uno va a los hospitales a mejorarse y no a morir. Pero, ¿cuáles son exactamente esos errores médicos fatales y cómo podríamos identificarlos y contabilizarlos? El hecho mismo de que se nos dé a conocer un rango tan amplio de números totales revela que son aproximaciones, no conteos precisos. La pregunta es: ¿cómo se llegó a esas cifras?

La respuesta es un tanto compleja. Estas estimaciones en particular se obtuvieron a partir de dos estudios de altas hospitalarias en los que se revisaron fichas médicas para identificar “efectos secundarios” (lesiones causadas por errores médicos). A partir de ellos, los investigadores concluyeron que entre 3% y 4% de los pacientes sufrió dichos daños. Sin embargo, ninguno de los dos estudios midió el porcentaje de efectos secundarios evitables o el porcentaje de los eventos que resultaron fatales. Ambas cifras habían sido calculadas después por personas que reinterpretaron a partir de los datos del estudio original². A pesar de que algunos criticaron el origen poco claro de esos datos, fueron los cálculos tardíos de muertes los que llamaron poderosamente la atención, no los resultados de la investigación original sobre los efectos secundarios.

Además, los estudios originales no consideraban el estado general de salud de cada paciente. Una investigación posterior realizó un análisis más detallado que sí lo tomaba en cuenta. Los resultados de esta investigación nos hacen recordar que los pacientes hospitalizados, a fin de cuentas, suelen estar gravemente enfermos y sus expectativas de vida no van más allá de unos días. La causa inmediata de su muerte

² Véase Harold C. Sox Jr. y Steven Woloshin, “How Many Deaths Are Due to Medical Error? Getting the Number Right”, *Effective Clinical Practice* 3 (2000): 277-282.

pudo haber sido un error médico, o “efecto secundario evitable”. Pero hay que tener en cuenta que la precaria salud de esos enfermos los hace especialmente vulnerables a los efectos de los errores médicos. En todo caso, no es común que se elijan esos casos para ejemplificar el riesgo de esos errores.

Quienes defienden esas estadísticas y los medios de comunicación prefieren muestras más dramáticas, destacando casos de pacientes que antes del efecto secundario tenían una larga expectativa de vida, por ejemplo un joven atleta que fue operado de una lesión menor que derivó en un grave daño cerebral³. El estudio que consideraba el estado general de salud de cada paciente sugería que “por cada 10.000 admisiones, un óptimo cuidado [es decir, sin errores]... implicaría alrededor de 1 paciente adicional con una expectativa de vida de 3 meses o más en buen estado de salud cognitiva”⁴. En otras palabras, esos investigadores argumentaban que los errores médicos rara vez tienen resultados fatales en enfermos con mayores expectativas de vida.

Con lo anterior no quiero decir que los hospitales no cometen errores fatales, pues eso sí sucede en cierta medida. Tampoco pretendo descartar algunos estudios o apoyar otros. El punto que quiero enfatizar es que *estimar el tamaño de un problema no es tan sencillo como parece*. Este ejemplo muestra cuán complicado puede ser medir lo que a primera vista parece un fenómeno que no deja lugar a dudas: pacientes que han muerto debido a errores médicos. Incluso, siendo optimistas,

³ Warren Wolfe, “Reporting Hospital Errors Seen as Good Idea”, *Minneapolis Star Tribune*, December 1, 1999, p. 3B. Luego de citar ese cálculo de 44 mil a 98 mil muertes anuales, el artículo destaca que “en los hospitales de Minnesota, entre 1994 y 1997, los accidentes médicos causaron o contribuyeron a la muerte de 26 personas”, pero no se refiere a la enorme diferencia entre las cifras nacionales y el pequeño número de muertes atribuibles a errores ocurridas en ese estado.

⁴ Rodney A. Hayward y Timothy P. Hofer, “Estimating Hospital Deaths Due to Medical Errors”, *Journal of the American Medical Association* 286 (July 25, 2001): 418. Este estudio se basó en pacientes de un hospital de veteranos, que pueden ser mayores que —o al menos, diferenciarse en otros aspectos de— los pacientes de otros hospitales.

si pudiéramos identificar cuáles muertes realmente fueron causadas por negligencias médicas, ¿contabilizaríamos cada error fatal? Algunos responderán que de todas maneras, pues cada paciente muerto debe ser considerado. Sin embargo, otros podrían distinguir entre un error que acorta un día la vida de un paciente terminal en coma, de otro que le roba décadas a una persona joven relativamente sana. ¿Cambiará nuestra opinión sobre el problema de los errores intrahospitalarios si descubrimos que los casos del último tipo son bastante escasos?

No hay respuestas correctas a esas preguntas; individuos bastante razonables estarán en desacuerdo con lo que habría que contabilizar. Dada la enorme ventaja retórica de plantear un problema como algo muy grande —y atemorizante—, es raro que esas sutilezas salgan a la luz en las discusiones sobre los problemas sociales. Y, desde luego, usar ejemplos convincentes para ilustrar un problema puede hacer que las cifras se vuelvan aún más aterradoras.

Problemas con las tendencias

Aun cuando un problema todavía no sea tan grande, puede estar creciendo. Las mediciones a lo largo de un período de tiempo nos permiten identificar tendencias, es decir, patrones o pautas de cambio. Esta es una forma muy importante de razonamiento, y una vez más, no tan sencilla como parece. *El principal problema con la evaluación de tendencias es mantener la comparabilidad de las mediciones*: si no medimos cada vez las mismas cosas de igual forma, nuestras cifras estarán reflejando cambios en el modo de medir, en lugar de diferencias en lo que estamos contabilizando⁵.

Una manera de comprobar las afirmaciones acerca de las tendencias sociales es preguntarse: ¿qué podría estar provocando el cambio?

⁵ Para una discusión más detallada sobre problemas relativos a medir tendencias, véase Best, *Damned Lies and Statistics*, pp. 98-109.

Supongamos que los medios dan a conocer un aumento en las denuncias de, digamos, abuso sexual perpetrado por familiares. Podríamos preguntarnos por qué sucede esto. ¿Existe alguna razón para pensar que el número de parientes abusadores está creciendo? Tal vez. ¿Pero no será que acaso las personas están prestando más atención al abuso de este tipo? (Obviamente, esto último es así, pues existen más reportajes periodísticos sobre el tema). Tal vez la gente se esté acostumbrando al problema y a considerarlo más en serio, por lo que está más dispuesta a denunciar. También puede deberse a que las autoridades están haciendo mejor su trabajo de registrar tales denuncias. Algunos suelen descartar esas explicaciones alternativas, pues plantean que una mayor atención a este tipo de abuso ha creado una suerte de reacción contraria: un incremento de los casos debido al mayor interés público. Unas afirmaciones generan otras que las contradicen, pero el peso de las demostraciones debe caer sobre quienes plantean que existe una tendencia.

Respecto a esto se hacen aparentes un par de normas. En primer lugar, *deberíamos desconfiar de aquellas afirmaciones que sugieren que una tendencia se ha revertido de manera súbita*. En general, las pautas sociales cambian lentamente porque los acuerdos sociales conllevan una considerable inercia. Las redes sociales son redes de conexiones reforzadas por conjuntos de supuestos culturales. *Ni esas redes, ni esos supuestos, cambian súbita y conjuntamente*. Si nos preguntamos por qué algunas personas cometen delitos, convendría tener en cuenta que la mayor parte del tiempo casi toda la gente respeta la ley. Los criminólogos entregan todo tipo de respuestas, señalando la dinámica familiar, la situación económica, las características del sistema de justicia penal, mensajes de la cultura, etc. Los índices de criminalidad y conducta respetuosa de las leyes pueden depender de todas las variables indicadas. La complejidad propia de estos vínculos causales dificulta que las tendencias cambien abruptamente: mientras una causa de la criminalidad puede sufrir una modificación importante, es poco probable que todos

los factores causales vayan a cambiar simultáneamente. A pesar de esta complejidad, es habitual que cuando alguien nos advierte sobre alguna nueva tendencia, suele plantear al mismo tiempo que este efecto se debe a un determinado cambio particular en un factor específico.

Aun cuando surgen nuevas tendencias, es probable que explicaciones simples de una sola variable no puedan dar cuenta de su desarrollo. Por ejemplo, después del aumento de las tasas delictuales en la década de los 80, éstas cambiaron de dirección durante los años 90 y comenzaron a disminuir. Muchos atribuyeron esta nueva tendencia a causas específicas, como el combate contra la droga, las políticas de “mano dura” (reforzamiento de las leyes en contra del desorden público) o la mayor presencia policial en las calles. Pero los criminólogos que pretendían investigar y explicar la caída de las tasas delictuales concluyeron que ésta era producto de una combinación de factores, entre los que se pueden mencionar la prosperidad económica y un cambio en el patrón de consumo de drogas^{6**}.

En segundo lugar, deberíamos desconfiar de explicaciones que atribuyen una tendencia a algún tipo de preocupación típica de nuestra cambiante sociedad. Vivimos en un mundo marcado por cambios medianamente constantes, aunque esto no es nada nuevo. Desde la Revolución Industrial (alrededor de la primera mitad del siglo XIX), el cambio es parte constitutiva de la experiencia estadounidense. Por ejemplo, cuando nos impresionamos de cómo internet ha acelerado las comunicaciones (“¡Todo ha cambiado!”), nos olvidamos de las importantes transformaciones ocurridas por la extensión de las redes

⁶ Alfred Blumstein y Joel Wallman, eds., *The Crime Drop in America* (New York: Cambridge University Press, 2000); Andrew Karmen, *New York Murder Mystery: The True History Behind the Crime Crash of the 1990s* (New York: New York University Press, 2000).

^{**} Recientemente, esta tendencia tan llamativa ha sido atribuida a otro factor: la decisión, en 1973, del Gobierno Federal de EEUU de legalizar el aborto, con lo cual la cantidad de niños no deseados y maltratados ha disminuido significativamente. Véase Steven D. Levitt y Stephen J. Dubner, *Freakonomics* (Buenos Aires: Ediciones B, 2006) (N. del E.).

telefónicas en el siglo XX o el surgimiento del telégrafo en el siglo XIX. Hace más de dos siglos que se habla de que el ritmo del cambio amenaza con destruir el tejido social de Estados Unidos. Aun cuando confiemos en nuestra habilidad para medir tendencias, debemos ser cuidadosos y no apresurarnos a sacar conclusiones sobre sus causas.

¿Se acerca el Apocalipsis?

Las actuales discusiones acerca de los problemas sociales no sólo nos advierten sobre el empeoramiento de las tendencias que nos preocupan, sino que nos dicen que un poco más adelante nos aguarda la catástrofe. Estas inquietudes toman muchas formas: temor ante la espiral de guerras fuera de control (guerra nuclear, invierno nuclear, los peligros de las armas químicas o biológicas de destrucción masiva); desastres medioambientales (sobrepoblación, agotamiento de los recursos naturales, contaminación, calentamiento global); temores médicos (epidemias de nuevas enfermedades como VIH o ébola, problemas de salud causados por la contaminación); preocupación frente al colapso económico; y otras amenazas exóticas, desde los impactos de asteroides y la robótica (seres humanos movidos por la inteligencia artificial) hasta la nanotecnología (materiales creados por la ingeniería que superan las formas de vida biológica), sin olvidar la crisis del Y2K. Aparentemente, el mundo de allá afuera es muy peligroso.

De más está decirlo, pero cuando estas visiones apocalípticas incluyen estadísticas, éstas suelen ser imprecisas. A menudo el método adoptado es el que utilizó Thomas Malthus, párroco del siglo XVIII que explicó que la hambruna era inevitable, pues el crecimiento demográfico sobrepasaría la producción agrícola. El modelo maltusiano era simple y de fácil comprensión; cualquiera que acepte sus suposiciones, concluirá que el resultado —hambruna catastrófica— es inevitable. El único problema es que se ha demostrado que Malthus estaba equivocado: el crecimiento demográfico puede ser y ha sido controlado

(muchos expertos piensan que la población global dejará de crecer en algún momento durante este siglo) y la producción agrícola se ha expandido más rápidamente que la población.

La lección es la siguiente: los escenarios apocalípticos —en especial, los más increíbles (“¡Sí, eso podría suceder!”)— dependen de las presuposiciones que hay detrás. La exactitud de esas presunciones está directamente relacionada con la importancia que le asignemos al escenario. El mundo es muy complicado, mucho más complejo que el más elaborado modelo computacional. Sin embargo, cuando hablamos de problemas sociales —incluso aquellos grandes asuntos que amenazan la vida—, tendemos a simplificar su complejidad.

Ciertamente carezco de conocimientos suficientes como para evaluar las bases científicas de las advertencias frente al calentamiento global, y sospecho que muchos de los que trabajan en los medios de comunicación no están mejor calificados que yo. Dependemos de los científicos para que nos aconsejen sobre estos temas. Con todo, sé lo suficiente —como debiera ser el caso de los periodistas— como para poner en duda que una pequeña cantidad de datos basta para establecer que estamos frente a un calentamiento global de proporciones catastróficas. Por ejemplo, un biólogo dio a conocer un informe relativo al desplazamiento del área ocupada por la *Euphydryas editha*, una mariposa californiana, hacia el norte del país; la prensa inmediatamente lo calificó como evidencia del impacto del calentamiento global. Análisis posteriores han cuestionado esta interpretación. Ahora, el punto es que datos relativos al cambio en el hábitat de una mariposa en particular no son lo suficientemente convincentes como para confirmar o desacreditar el argumento de que la actividad humana está causando un calentamiento global catastrófico⁷. Si las afirmaciones sobre el calentamiento global son ciertas, de seguro existen muchas evidencias como

⁷ Para una crítica de cómo los medios trataron este tema, véase David Murray, Joel Schwarz y S. Robert Lichter, *It Ain't Necessarily So: How Media Make and Unmake the Scientific Picture of Reality* (Lanham, MD: Rowman & Littlefield, 2001), pp. 49-52.

ésta. Pero los medios suelen enfocarse en informes aislados: las historias son fáciles de comprender (las mariposas se desplazaron hacia el norte); son apropiadas para su tratamiento visual (podemos imaginar a un periodista hablando contra un fondo de mariposas filmadas revoloteando alegremente); y pueden ser usadas como una señal de una tendencia mayor y atemorizante.

Las afirmaciones apocalípticas no tienen buenos antecedentes. Y hay que revisar con cuidado cuando se asevera que las estadísticas las apoyan, en especial cuando se plantea que las cifras sencillas y fáciles de entender son prueba de que el futuro nos depara cambios complejos que amenazan nuestra civilización.

RIESGOS

Las estadísticas relativas a los riesgos se han vuelto uno de los tipos más comunes de cifras que asustan. Hablamos sobre “aumento del riesgo”, “factores de riesgo” o estar “en riesgo”. La década decisiva en nuestra comprensión del riesgo debió ser la de los años 60, pues en ese período ocurrió un hito clave. En 1964, el Director General de Servicios de Salud de EEUU dio a conocer el famoso informe acerca del tabaco y la salud. En tanto las tabacaleras insistían en que no había evidencias convincentes sobre una relación causal, los detractores hacía tiempo que advertían que fumar daña la salud. Este informe tuvo gran impacto porque era lo suficientemente acreditado (aun cuando muy pocos estadounidenses podrían explicar en detalle cómo se llegó a esas conclusiones) y porque entregaba una completísima revisión de gran cantidad de datos que llevaban a una conclusión: fumar aumenta el riesgo de contraer una serie de enfermedades.

El citado informe casi coincide con la publicación de otros dos libros centrados en la idea del riesgo: *Primavera Silenciosa* (1962) de Rachel Carson e *Inseguro a Cualquier Velocidad* (1965) de Ralph Nader. El primero advertía que el DDT y otros químicos amenazaban el

medio ambiente, y el segundo atacaba la incapacidad de la industria automotriz para diseñar vehículos más seguros. Estas críticas hablaban de productos de uso cotidiano —cigarrillos, sustancias químicas y vehículos motorizados— que representaban peligros graves que, pese a ello, permanecían ocultos. Esos análisis estimularon la discusión sobre los riesgos. Al final de la década había surgido una organización pro defensa de los derechos de los consumidores que buscaba protección contra las sustancias peligrosas, y el movimiento ambientalista captó nuevos seguidores haciendo énfasis en los riesgos de la contaminación. De manera creciente, los riesgos fueron considerados como ocultos, quizás irreconocibles y peligrosos, aunque eran potencialmente manejables si se lograba su comprensión y conocimiento adecuado y el modo de tratarlos. Los medios de comunicación tuvieron un rol muy importante en el proceso de advertir al público respecto a estos riesgos.

Muchos de los artilugios de la vida moderna —cinturones de seguridad; *airbags* de los automóviles; cascos para ciclistas; alimentos producidos sin grasa, cafeína o pesticidas; restaurantes y lugares de trabajo libres de humo; sexo seguro; dosis diarias de aspirina para niños; chequeos médicos varios— reflejan nuestra interpretación acerca de algunos riesgos y los intentos de minimizarlos. Hay algo cómico en esto de intentar ajustar nuestra vida al último reportaje sobre el último estudio dado a conocer. ¿Es tan perjudicial beber vino? o ¿es beneficioso para la salud tomar una bebida alcohólica al día? ¿Sólo el vino tinto es bueno? (Personalmente, me estoy inclinando ante la idea de que el chocolate amargo prolonga la vida, y si alguien tiene evidencias en contra, no quiero enterarme).

Cuando tratamos de traducir estas palabras a números, entramos al mundo de las probabilidades. Un riesgo es la probabilidad de que algo pueda ocurrir. Por ello, cuando decimos que un fumador tiene un mayor riesgo de contraer cáncer pulmonar, no estamos diciendo que todos los fumadores vayan a desarrollar esa enfermedad, como tampoco

queremos afirmar que ningún no fumador enfermará. Más bien, la idea de un riesgo mayor implica comparar probabilidades: si X de cada 1.000 no fumadores eventualmente desarrollan cáncer pulmonar y si los fumadores enferman en mayor número, entonces el número de casos de cáncer por cada 1.000 fumadores debiera ser mucho mayor. La idea parece simple, pero los números llevan a confusión.

El cálculo de probabilidades es algo no bien comprendido. (Esto explica por qué a los casinos les va tan bien). Tendemos a descubrir y reconocer pautas y a suponer que significan algo. Si lanzamos al aire una moneda 4 veces y obtenemos 4 caras seguidas, algunos supondrán que en la próxima tirada obtendremos sello (porque este resultado de alguna manera está atrasado), en tanto otros dirán que de nuevo saldrá cara (porque hay una “racha”). Un matemático nos diría que todos están equivocados porque cada lanzamiento es independiente de los demás: lo que suceda la próxima vez que lancemos la moneda no tiene nada que ver con lo que sucedió en el lanzamiento anterior. Después de 4 caras, la posibilidad de otra cara en el quinto lanzamiento es 50/50. Si conseguimos una quinta cara, la posibilidad de otra cara al sexto lanzamiento sigue siendo 50/50. Si lanzamos una moneda 6 veces, tenemos 64 posibles secuencias de resultados. Seis caras consecutivas (CCCCCC) es uno de los resultados posibles; CSCSCS, otro. Tendemos a fijarnos en el primero y a considerarlo espectacular, mientras el último parece rutinario, pero la probabilidad de lograr cualquiera de ambas pautas es exactamente la misma: 1 en 64. Con esto no pretendo afirmar que la probabilidad de conseguir 6 caras es la misma que de lograr 3 caras y 3 sellos; 20 de las 64 posibles secuencias involucran 3 caras y 3 sellos (CCCSSS, CCSCSS, etc.). Pero puede darse cualquier secuencia de igual modo, y el hecho de que algunas secuencias parecieran conformar pautas reconocibles no las hace más o menos propensas a ocurrir.

Una vez que comprendemos esto, podemos entender que se produzcan todo tipo de combinaciones aparentemente inusuales, el tipo

de cosas que consideraríamos coincidencias memorables. Si alrededor del 10% de las personas son zurdas, entonces la probabilidad de que nos topemos con 1 zurdo es de 1 en 10 (0,1); con 2 zurdos, 1 en 100 ($0,1 \times 0,1 = 0,01$); y con 3 zurdos, 1 en 1.000 ($0,1 \times 0,1 \times 0,1 = 0,001$). A pesar de estas probabilidades, si nos topamos con mucha gente, ocasionalmente nos encontraremos con 2 ó 3 zurdos consecutivamente. Podemos esperar que sucedan cosas extrañas, si bien rara vez ocurren.

Convertir estos principios en estadísticas (los llamados cálculos del riesgo) por lo general lleva a la confusión. Consideremos el siguiente problema respecto a mujeres que se han hecho mamografías para detectar cáncer (datos aproximados de mujeres de 40 años, o más, que no tienen otros síntomas):

La probabilidad de que una de estas mujeres tenga cáncer de mamas es de 0,8%. Si una mujer tiene cáncer mamario, la probabilidad de que su mamografía salga positiva es de 90%. Si una mujer no tiene este cáncer, hay un 7% de probabilidades de que su mamografía resulte positiva. Imaginemos una mujer con una mamografía positiva, ¿cuál es la probabilidad de que tenga cáncer mamario?⁸.

¿Confundido? No hay de qué avergonzarse. Cuando se planteó este problema a 24 médicos titulados, sólo 2 acertaron. La mayoría se equivocó rotundamente: un tercio respondió que había un 90% de probabilidades que una mamografía positiva indicara cáncer de mamas; otro tercio mencionó cifras entre 50 y 80%. La respuesta correcta estaba en torno al 9%.

Revisemos con atención el problema. Primero, hay que señalar que el cáncer mamario es relativamente poco común (0,8%); vale decir, por

⁸ Ejemplo tomado de Gerd Gigerenzer, *Calculated Risks* (New York: Simon & Schuster, 2002), p. 41.

cada 1.000 mujeres, 8 tendrán esta enfermedad. La probabilidad de que esas mujeres reciban una mamografía positiva es de 90%; digamos que 7 de cada 8. Eso hace que 992 mujeres no padecen de cáncer. De este grupo, alrededor de un 7% también recibirá un resultado positivo en su mamografía, lo que equivale a cerca de 69 casos llamados falsos positivos. Es decir, 76 mujeres en total ($7 + 69 = 76$) obtendrán un resultado positivo en el examen, aunque sólo 7 de éstas —alrededor de un 9%— efectivamente tendrán cáncer mamario. Lo que quiero destacar es que al medir riesgos, es usual que haya que realizar una serie de cálculos. Incluso los profesionales capacitados (como los médicos) no están acostumbrados a calcular el riesgo y cometen errores con facilidad.

Desafortunadamente, *los mismos médicos suelen entregar ese tipo de información sobre el riesgo a sus pacientes, quienes tienen aún menos capacitación para entender su significado, con lo cual pueden quedar bastante alterados*. Es muy probable que una mujer que acaba de recibir su mamografía positiva esté muy alterada y su capacidad de comprender las cifras y calcular el riesgo en que se encuentra, sea incluso menor a la de los médicos (quienes, como vimos recién, en la mayoría de los casos tampoco tienen la respuesta correcta).

Medir riesgos

Entonces, ¿cómo calculamos los riesgos? ¿De adónde surgen esas cifras? Esta última es una pregunta espinuda. Idealmente, la ciencia opera a través de experimentos. Supongamos que queremos saber si alguna actividad implica un riesgo para la salud, por ejemplo ingerir bebidas gaseosas dietéticas (lo que yo suelo hacer). Imaginemos un diseño experimental en el cual dos grupos de niños han sido seleccionados al azar y criados en idénticas circunstancias, salvo que el grupo experimental toma bebidas gaseosas dietéticas y el grupo control, no. Les hacemos un seguimiento en la adultez y hasta la vejez, y determinamos si los grupos tienen distintos tipos de problemas de salud. Obviamente,

ese experimento es irrealizable; sería demasiado costoso en tiempo y dinero, y ni hablar de lo inmoral que sería entrometerse así en la vida de las personas. Por estos motivos, *los cálculos del riesgo casi nunca provienen de experimentos con seres humanos*.

En lugar de eso, los investigadores deben diseñar métodos alternativos para estudiar el riesgo. Por ejemplo, pueden identificar a individuos enfermos y averiguar si ingieren más bebidas dietéticas que quienes están sanos, o comparar las tasas de enfermedad en comunidades con alta y baja ingesta conocida de este tipo de bebidas, o experimentar con ratas de laboratorio, dándoles de beber ese producto a algunas, y a otras no. Todos estos diseños implican un arreglo metodológico limitante; son maneras imperfectas de determinar si los consumidores de bebidas gaseosas dietéticas corren mayores riesgos de enfermar que los que no consumen gaseosas dietéticas. Por una parte, esto es inevitable: cada estudio tiene limitaciones de diseño; y por otra, las imperfecciones al medir el riesgo saltan a la vista (porque nunca se puede estudiar a las personas bajo condiciones experimentales estrictamente controladas). Por ende, los resultados de las investigaciones son imprecisos y deben ser tratados con mucho cuidado. Con respecto a esto, quiero plantear dos advertencias fundamentales.

Primero que todo, los resultados de la investigación no deben considerarse convincentes a no ser que revelen la existencia de un riesgo importante. Imaginemos un estudio en que los sujetos que ingieren gaseosas dietéticas parecen ser más propensos a sufrir una determinada enfermedad que los que jamás han probado esa sustancia. Dado que nuestro estudio no puede haber controlado todos los aspectos de la vida de esas personas, no podemos estar seguros de que la diferencia radique en ingerir o no la gaseosa dietética. Para utilizar el término que presentamos en el Capítulo 2, diremos que el vínculo entre bebida dietética y enfermedad puede ser espurio. Por ejemplo, podemos sospechar que los consumidores de bebidas dietéticas están más preocupados de su peso. Quizás hacen menos ejercicio, comen más o tienen

una predisposición genética a subir de peso. ¿Cómo podemos estar seguros de que sus problemas de salud sean el resultado de su elección al momento de tomar una bebida, antes que de uno u otro de los factores recién mencionados? No es posible. Por lo tanto, antes de concluir que la gaseosa dietética es la causa de una mayor incidencia de enfermedades entre los consumidores de bebidas gaseosas, debemos tener evidencias relativamente convincentes.

¿Pero qué hace que las evidencias sean convincentes? Una convención comúnmente usada en este tipo de investigación epidemiológica requiere que los riesgos identificados sean 3 veces más que en el grupo de comparación (o sea, 200% mayor).

Veamos. ¿Se confundió? Un ejemplo: si X es 5, entonces 3 veces X es 15, que a su vez es 10 más que 5. Ya que 10 es un 200% de 5, 15 es 3 veces 5 ó, dicho de otro modo, 200% más grande.

Éste no es un estándar arbitrario. Dado que la investigación mencionada no es realmente experimental, es fácil sospechar que las relaciones que parecen causales sean espurias. Y mientras más débil sea el nexo, es más probable que se trate de un hallazgo accidental, en particular si el riesgo estudiado es poco común. Según la teoría estadística —demasiado técnica para explicarla aquí—, las probabilidades de que una relación aparente (que involucra un riesgo poco común) no sea válida, disminuyen sólo cuando los riesgos identificados son, al menos, un 200% mayor⁹.

Con sólo entender esto ya disponemos de un arma poderosa para evaluar los informes periodísticos acerca de las investigaciones más recientes. Supongamos que mañana usted recibe el periódico y lee que una revista médica publicó un estudio que indica que las personas que ingieren bebidas gaseosas dietéticas tienen un 20% más de

⁹ Para un debate sobre este punto, véase John Brignell, *Sorry, Wrong Number! The Abuse of Measurement* (Great Britain: Brignell Associates, 2000), pp. 46-51. Acerca de las políticas sobre el tema, véase Gary Taubes, "Epidemiology Faces Its Limits", *Science* 269 (July 14, 1995): 164-169.

probabilidades de sufrir un determinado problema de salud. Esa frase confundirá a algunos individuos que, por ejemplo, creerán que un 20% de quienes beben esos productos se enfermarán. De hecho, ese dato no significa nada de esto.

Supongamos que en la población general, 5 personas de cada 10.000 tienen esa enfermedad. Si quienes toman bebidas cola dietéticas corren un 20% más de riesgo, habrá 6 casos de esa enfermedad entre 10.000 sujetos que ingieren esos productos (20% de 5 es 1, entonces un aumento de 20% es igual a $5 + 1$, o sea, 6). En otras palabras, lo que puede ser una estadística impresionante —“¡un riesgo 20% mayor!”— en el mundo real de la práctica, se refiere a un número muy pequeño de gente: 1 caso adicional por cada 10.000 individuos. (De hecho, podemos sospechar que los investigadores y los periodistas prefieren decir “20% mayor” a un “factor de riesgo de 1,2”, que significa lo mismo, precisamente porque el resultado parece mayor y se ve más dramático).

Pero recordemos: para que una investigación sea considerada seria, debe reportar un riesgo 200% mayor. Por ejemplo, si la tasa es de 5 casos de enfermedad por cada 10.000 personas, la investigación debe revelar una tasa de enfermedad de al menos 15 casos por cada 10.000 bebedores de gaseosas dietéticas (15 es 3 veces —200%— mayor que 5).

¿Es éste un estándar razonable? Veamos, los fumadores tienen 1.900% veces más probabilidades de desarrollar cáncer que los no fumadores. Cada vez que lea en la prensa un informe que indique un riesgo menor a 3 veces, ó 200% mayor, usted tiene razón en estar escéptico de los resultados.

Una segunda advertencia se refiere a la importancia de insistir en la existencia de estudios múltiples. Cualquier estudio único puede estar equivocado. Los científicos saben que para comprobar la validez de un hallazgo, la investigación debe poderse replicar: repetirla y obtener resultados similares. (Recuerden el entusiasmo que cundió en 1989

cuando se informó del descubrimiento de la fusión fría y que colapsó, precisamente, cuando los investigadores de otros laboratorios no pudieron replicar esos resultados).

También ayuda la investigación por medio de triangulación. Esto se refiere a utilizar diversos métodos; ya que cada método en sí mismo tiene defectos, las diversas fallas de los distintos métodos pueden eliminarse entre sí. Por ejemplo, el vínculo entre fumar y el cáncer pulmonar está suficientemente probado porque ha sido confirmado en repetidas oportunidades por medio de estudios que utilizan diversos métodos.

A veces los investigadores comparan los resultados de distintos estudios a través de lo que se denomina *metanálisis*. La suposición lógica es que si varias investigaciones muestran un efecto de manera constante, incluso si éste no es contundente (es decir, el riesgo no cumple el requisito de ser 200% mayor), esos resultados debieran darnos mayor confianza de que se trata de una relación real. El problema con esta lógica es que no siempre los investigadores quieren publicar los resultados decepcionantes, o tienen mayores dificultades para hacerlo. Este *sesgo editorial* dificulta encontrar resultados publicados de dichos estudios. En consecuencia, el metanálisis tiende a incluir sólo los estudios más exitosos, aquellos con los resultados más apropiados para publicar. La técnica del metanálisis no es ilegítima, pero tampoco implica un gran respaldo a un nexo determinado. De ahí que un metanálisis de varios estudios que señale, por ejemplo, un 20% de aumento del riesgo, no debiera hacernos confiar por completo en los resultados.

Contextualizar los riesgos también sirve. Cada vez que nos subimos al automóvil para ir al trabajo, estamos corriendo un riesgo. Todos sabemos que los accidentes de tránsito causan muchas muertes. Hasta cierto punto, podemos minimizar los riesgos si obedecemos las leyes del tránsito y utilizamos el cinturón de seguridad; pero si bien la posibilidad de morir en algún viaje es baja, el riesgo nunca llega a valor 0. Aun así, esos riesgos cotidianos —el tipo de cosas que damos por

sentado— pueden ser mucho mayores que aquellos riesgos tan publicitados que en ciertas oportunidades se vuelven el foco de la atención pública. Cuando sentimos miedo, tendemos a fijarnos en aquello que nos asusta más que en el riesgo real de vernos afectados, una reacción que se ha denominado “descuido de la probabilidad” (*probability neglect*)¹⁰.

Un buen ejemplo de esto, en el otoño de 2002, fue la horrorizada reacción del público frente a las noticias que indicaban que había un francotirador en una zona cercana a Washington, DC. Dado que suponemos que el riesgo de morir asesinado por un francotirador es igual a 0, la información relativa a que tal riesgo existía, atemorizó a las personas. Incluso en una región con millones de habitantes, el riesgo de ser baleado era muy bajo. Durante las semanas en que efectivamente los ataques del francotirador cobraron víctimas fatales, el riesgo de morir en un accidente de tránsito en la ciudad de Washington era mucho mayor, aun cuando eso no fuera tema de los noticiarios. Es probable que prolonguemos bastante nuestras expectativas de vida si seguimos las indicaciones que nos han dado durante toda la vida (no fume, use cinturón de seguridad, aliméntese de manera razonable y haga ejercicio), que si nos ponemos a resguardo de un francotirador o evitamos ese aditivo para las comidas del que tanto se habla en los titulares de los medios.

El riesgo de divorcio

Otro motivo por el que la noción de riesgo lleva a la confusión es que no siempre estamos seguros de cuál sea la mejor forma de calcular los riesgos. Consideremos un tema bastante sencillo que parece ser algo más complicado: ¿qué porcentaje de matrimonios acaba en divorcio? No hay ningún organismo oficial que registre en particular

¹⁰ Cass R. Sunstein, “Probability Neglect: Emotions, Worst Cases, and Law”, *Yale Law Journal* 112 (2002): 61-107.

los matrimonios y entonces pueda identificar con precisión cuántos de ellos terminan en divorcio, justamente el tipo de información que quisiéramos tener para responder nuestra pregunta. Al carecer de información completa y precisa, los analistas se ven obligados a utilizar las cifras disponibles. Ya que solicitar una licencia matrimonial u obtener un divorcio son actos legales, los organismos oficiales llevan un registro de ellos y varias jurisdicciones realizan una cuenta detallada de esos actos. Por consiguiente, desde hace tiempo los analistas han optado por dividir el número de divorcios en un año determinado por la cantidad de matrimonios realizados en ese mismo período, de modo de obtener una cifra preliminar respecto a la probabilidad de que un matrimonio termine en divorcio. Aproximadamente desde 1960, la cantidad de divorcios ha sido alrededor de la mitad de los matrimonios, y a esto se le ha denominado “tasa de divorcio”.

El problema es que cuando hablamos de una tasa, lo que hacemos habitualmente es dividir alguna cifra de hechos (muertes o delitos) por la población en riesgo. Así, tanto las tasas de fallecimiento como las tasas delictuales suelen presentarse como una cifra por 100.000 personas en la población; por ejemplo, el FBI informó que en EEUU la tasa de asesinatos en 2000 era de 5,5 asesinatos por cada 100.000 individuos.

¿Pero quiénes constituyen la población en riesgo, cuando queremos calcular una tasa de divorcio? Por cierto, no incluye sólo a quienes se casaron ese mismo año; de hecho, sabemos que relativamente pocas parejas se divorcian en el transcurso del año calendario en que contrajeron matrimonio. Entonces, la población en riesgo son todas las parejas casadas, una cifra bastante grande. Si calculamos la tasa de divorcio dividiendo el número de divorcios durante un año determinado por el total de parejas casadas, sin importar la duración del matrimonio, entonces la tasa de divorcio será mucho menor que 50%. Todos los que opinan sobre este tema coinciden en que el matrimonio es una institución mucho más estable, y que el divorcio es mucho menos común, de lo que creemos.

Revisemos detenidamente esta aseveración. Imaginemos una comunidad en la que se celebran dos matrimonios cada año y que uno de ellos termina en divorcio durante ese mismo período de tiempo. En este caso, es cierto que la mitad de los nuevos matrimonios se divorcia; pero también es cierto que, con cada año que pasa, el número total de parejas casadas crecerá en 1 punto. Así, después del primer año, si dividimos ese único divorcio por el número total de parejas casadas, obtendremos una tasa cuyo valor será menor que 50% a pesar del hecho de que la mitad de los matrimonios termina en divorcio. Este argumento sugiere que la crítica convencional utilizada para desestimar la alta tasa de divorcios es deficiente.

Claramente, medir el riesgo de divorcio es un problema difícil que requiere pensar con atención y disponer de mucha información. En 1996, se realizó una encuesta a una muestra muy grande, casi 70 mil personas mayores de 15 años residentes en 37 mil hogares. Los encuestados fueron consultados respecto a todos los matrimonios y divorcios en sus historias personales. Por ejemplo, un sujeto informó que se había casado una sola vez, 40 años antes, y que permanecía con su cónyuge; mientras otro entrevistado, en la actualidad soltero, mencionó matrimonios en 1970 y 1985, los que terminaron en 1980 y 1992 respectivamente¹¹. A partir de estos datos, los investigadores pudieron identificar cohortes de matrimonios que se realizaron durante diversos períodos (p.ej., primeros matrimonios ocurridos en 1945-1949) y calcular la proporción de matrimonios en cada cohorte que en 1996 estaban divorciados. (Desde luego que es posible que una pareja que permanecía casada al momento de la entrevista pudiera divorciarse

¹¹ Rose M. Kreider y Jason M. Fields, "Number, Timing, and Duration of Marriages and Divorces, 1996", Bureau of the Census, *Current Population Reports*, P70-80 (Washington, DC, February 2002). Para un estudio basado en una muestra diferente que obtuvo resultados similares, véase Matthew D. Bramlett y William D. Mosher, "First Marriage Dissolution, Divorce, and Remarriage: United States", *Advance Data from Vital and Health Statistics* 323 (May 31, 2001).

más adelante). Si bien estos datos son incompletos, dado que se originaron en una muestra y no en la población completa, se trata de una buena muestra y la información entrega una idea de lo que sucede a lo largo del tiempo con parejas casadas (lo que es, recordemos, el tipo de información que pedíamos al inicio de esta discusión).

Estos datos sugieren que podemos esperar que casi la mitad de las parejas actualmente casadas terminen divorciándose. Los investigadores encontraron importantes diferencias por cohortes que muestran cómo ha cambiado la sociedad: básicamente, las personas de una cohorte tendían a mantenerse casadas por más tiempo que las de la cohorte siguiente. Así, sólo el 34% de los hombres de 60 años habían terminado su primer matrimonio, pero el número comparable para los hombres de 50 años era de 40%. En el caso de las mujeres casadas por primera vez entre 1945-1949, el 70% permanecía casada 30 años después; pero entre las que se casaron por primera vez entre 1960-1964, sólo el 55% (un poco más de la mitad) seguía casada. Es demasiado pronto para decir qué proporción de parejas casadas en primeras nupcias durante 1980-1985 celebrará su trigésimo aniversario de matrimonio, pero podemos hacer proyecciones basándonos en lo que ya sabemos: sólo el 73% de las mujeres casadas en dicho período permanecía casada 10 años después, comparado con el 90% de aquellas casadas entre 1945-1949 cuyos matrimonios duraron al menos 10 años. Basándose en estos datos, los investigadores proyectaron que, mientras una gran proporción de los matrimonios más antiguos permanecía intacta, alrededor de la mitad de los más recientes acabarían en divorcio.

De ahí que responder una pregunta aparentemente sencilla —¿cuál es la probabilidad de que un matrimonio acabe en divorcio?— resulta ser un asunto bastante complicado. Pero este tipo de complejidad es pasada por alto en los reportajes periodísticos, que con ligereza informan del riesgo de esto o de aquello; una observación que debería hacernos reflexionar. *Es muy fácil asustarse con las estadísticas sobre los riesgos*. Mientras digerimos la advertencia de turno respecto a una nueva

amenaza, debemos tener presente las dificultades para calcular los riesgos recién descubiertos.

EXTERNALIDADES Y PESIMISMO PRETENCIOSO

A menudo las cifras que asustan nos advierten que el mundo está cambiando. Puede ser inquietante, incluso terrorífico, pensar acerca del cambio, en particular desde que los medios de comunicación se han centrado en que lo que viene será para empeorar las cosas. Una de las ideas más útiles cuando se trata de considerar el significado del cambio, es la noción de efectos colaterales o externalidades, lo cual quiere decir que cada cambio implica costos y beneficios. Es imposible hacer una comparación justa entre lo que hubo antes y lo que sigue si no consideramos los costos y beneficios equiparables.

Uno de los métodos clásicos de promover un cambio específico es contrastar los costos de lo que tenemos ahora con los beneficios que puede traer la modificación propuesta; asimismo, el cambio puede ser resistido enfatizando los beneficios de la situación actual y los eventuales costos que vendrán con el cambio. Para protegernos —para hacer una comparación justa—, necesitamos comparar manzanas con manzanas. En otras palabras, si consideramos que los beneficios actuales son importantes, debemos compararlos con los beneficios futuros, y los costos de ahora deben compararse con los que vendrán.

Es muy fácil olvidarse de esto. Muchos detractores se han vuelto suspicaces respecto al cambio tecnológico y señalan sus costos. La comparación, que suele estar implícita, nos lleva a un pasado idílico en el cual las personas vivían en armonía con la naturaleza y donde la vida era más sencilla y mejor. Esta visión, a través de la bruma del tiempo, es un tanto borrosa; los críticos pueden ver los beneficios del pasado, pero tienen dificultad en calcular los costos implicados. Así calculan, por ejemplo, la cantidad de muertes debidas a la contaminación ambiental provocada por las centrales de energía, pero olvidan contabilizar la

tasa de mortalidad ocasionada por la contaminación intradomiciliaria producto del uso de cocinas a leña y carbón. La comparación que hacen los críticos suele poner en la balanza las muertes actuales o futuras causadas por el cambio, contra un pasado en que la muerte se daba por hecho. El error contrario ocurre cuando los impulsores de un cambio destacan los beneficios e ignoran sus costos.

Entre las formas más comunes de convertir las estadísticas en aterradoras, se encuentran: (a) comparaciones que ignoran las externalidades o efectos colaterales, (b) los grandes cálculos, (c) las tendencias terroríficas, (d) los escenarios apocalípticos y (e) los riesgos mal definidos. Dado que las cifras que asustan son bastante convincentes, y que usualmente tenemos dificultades para distinguir las externalidades y riesgos relativos, muchas de las conversaciones actuales sobre los problemas sociales son atravesadas por la pesimista suposición de que las cosas están cada vez peor.

La experiencia de muchos estadounidenses contrasta con estas advertencias negativas. Con esto no pretendo sugerir que el mundo mejora cada día en todos los ámbitos posibles y para toda la gente: en nuestra sociedad hay mucha adversidad y sufrimiento. Sin embargo, en promedio, los estadounidenses están viviendo más años que sus antecesores, gozan de mejor salud y educación, y tienen estándares de vida más altos¹².

En resumen, hay una brecha entre la convicción de que nuestra vida está bastante bien y la sociedad allá afuera que parece abrumada por los problemas. Esta brecha aparece regularmente siempre que las encuestas de opinión pública hacen consultas sobre las experiencias personales de los individuos y su percepción respecto a la situación del país. Las personas tienden a sentirse razonablemente satisfechas con los profesores de las escuelas a las que asisten sus hijos, pero se

¹² Para una revisión más detallada de estos temas, véase Best, "Social Progress and Social Problems".

muestran muy preocupadas por la calidad de la educación; suelen tener una buena opinión de los parlamentarios que los representan, pero consideran que el Congreso es un desastre; creen que sus vidas van por buen camino, pero no así la sociedad. En general, el modo en que pensamos sobre la sociedad está teñido por esta suposición pesimista.

Como profesor, a lo largo de los años me ha tocado leer miles de trabajos de investigación y ensayos, y me doy cuenta de que los alumnos sencillamente suponen que la delincuencia empeora (así como la pobreza y los suicidios adolescentes), sin importar si las tasas delictuales caen o aumentan. Es como si todos creyéramos que vivimos cómodamente en el País de Jauja, aun cuando estamos convencidos de que la sociedad en su totalidad está destinada a irse al infierno.

En las últimas décadas hemos sido expuestos a una serie de escenarios apocalípticos y advertencias de que el mundo, tal cual lo conocemos, se va a acabar. Algunas amenazas se han diluido (¿recuerdan los temores de 1970 respecto a una nueva Edad de Hielo?), pero seguimos escuchando acerca de múltiples vías a la extinción: invierno nuclear, calentamiento global, sobrepoblación, epidemias, colapso económico, terrorismo. Las estadísticas que asustan ocupan un sitio preferente en estas afirmaciones. Un hallazgo aislado, como el informe de que una especie de mariposa había cambiado su hábitat, puede ser presentado como un presagio de un desastre inminente. Incluso las buenas noticias pueden ser interpretadas como vaticinadoras de catástrofes: si las tasas delictuales están cayendo, esa situación no va a durar mucho; y si dejaran de caer, de seguro es una señal de que la delincuencia volverá a aumentar. Nuestra disposición para especular sobre grandes consecuencias a partir de pequeños eventos denota que estamos siendo constantemente advertidos de que se avecinan cosas muy serias. Y, una vez más, nos sentimos frustrados por lo que parecen ser afirmaciones contradictorias: el alcohol es perjudicial para la salud, pero una copa de vino tinto es buena; un poco de alcohol en cualquier forma es saludable para el corazón (pero no para su hígado).

Las cifras que atemorizan florecen por doquier porque son parte integral del modo en que conversamos acerca de la vida social. Los defensores de diversas causas buscan asustarnos porque insisten en que enfrentamos amenazas reales y debemos ser remecidos para salir de nuestra complacencia. Están menos dispuestos a admitir otras consideraciones, pues desean competir por nuestra atención. Estamos rodeados de defensores de múltiples cosas, donde cada grupo intenta atraer nuestra atención hacia un problema determinado. Cada causa busca destacarse por sobre las demás. Asustar a las personas no es la única manera de ganar esta competencia, pero suele funcionar, en especial si quienes promueven alguna idea utilizan estadísticas que justifican el miedo. Por cierto, estamos en lo correcto cuando suponemos que las cifras que asustan seguirán siendo un componente importante de la forma en que conversamos acerca de los problemas sociales. Estas cifras no van a desaparecer; lo único que podemos hacer es acercarnos a ellas con escepticismo y sopesar si el temor que infunden está realmente justificado.

Capítulo 4

CIFRAS ACREDITADAS

Un par de veces al mes, el editor de alguna revista especializada me consulta si podría hacer las veces de árbitro o evaluador de un determinado artículo. La mayoría de la gente sabe o debería saber que los profesores universitarios deben “publicar o morir”. La carrera académica se basa en gran medida en las publicaciones en revistas especializadas. La evaluación por parte de pares externos es una arista, generalmente oculta, del proceso de publicación. Lo habitual es que después de terminar una investigación y escribir el informe sobre los resultados, los autores envíen sus manuscritos a revistas académicas especializadas. A su vez, los editores de esas revistas reciben más artículos de los que pueden publicar y por eso se apoyan en la evaluación de pares para hacer la selección. El editor envía una copia del manuscrito a diversos revisores (en sociología, las revistas más importantes acostumbran enviarlas a 3 ó 4 evaluadores externos). Y tal como lo sugiere el término, se supone que estos evaluadores son pares de los investigadores, es decir, profesionales entendidos en el tema y, por lo tanto, calificados para juzgar la calidad de un informe de investigación. Es posible que los evaluadores externos no estén de acuerdo entre sí, pero lo habitual es que un editor rechace un manuscrito únicamente cuando hay una mayoría de comentarios negativos.

Los autores, en especial aquellos cuyos manuscritos han sido rechazados, suelen poner en duda la rectitud del proceso de revisión. Algunos, por ejemplo, cuestionan el hecho de que esas evaluaciones sean anónimas (usualmente lo son, si bien los autores y los revisores a menudo adivinan la identidad del otro); otros plantean que una evaluación negativa puede ser reflejo del desacuerdo personal o político entre un revisor y un determinado autor. Sin embargo, el sistema de revisión por pares subsiste porque parece surtir mejor efecto que cualquier otro en lo que a selección de artículos para publicación académica se refiere. Es probable que funcione mejor en la mayoría de las revistas con más prestigio. Existen miles de revistas académicas, pero los intelectuales de varias disciplinas, especialidades y subespecialidades reconocen que algunas tienen muchos más lectores que otras; y la competencia por publicar en las más leídas y, por lo tanto, con mayor renombre, es intensa. Se supone que esas publicaciones son particularmente selectivas.

Los pares que evalúan son una suerte de guardabarrera cuya labor es señalar al editor las falencias de los manuscritos. ¿Es que el autor desconoce alguna investigación reciente sobre el tema? ¿Elegió métodos discutibles para realizar su estudio? ¿Utilizó quizás técnicas inapropiadas para analizar los resultados? Las dudas de los revisores en torno a esos temas advierten a los editores respecto a la publicación de artículos poco consistentes.

Cuando las personas hablan de una revista “acreditada o refrendada”, se están refiriendo a la integridad del proceso de revisión de esa publicación. Los editores y los pares que evalúan no pueden supervisar el proceso completo de la investigación, ni dar fe de la exactitud de cada una de las palabras del manuscrito; pero sí pueden sopesar lo que leen, estar alertas ante algunos signos alarmantes y aceptar para su publicación sólo los artículos más convincentes. Pero aun así, ocurren errores. A lo largo del tiempo, algunos resultados publicados son cuestionados; y rara vez se destapa públicamente un fraude científico con cargos de falsificación deliberada de resultados con el único objetivo

de ser publicados. A pesar de todo, el sistema parece funcionar bastante bien.

La autoridad de las instituciones sociales depende de ciertos acuerdos para asegurar su integridad. Con el objetivo de mantener a raya a las instituciones y a las personas que forman parte de ellas, se han diseñado y creado, entre otras técnicas, la separación de los poderes, la exigencia de juramentos a los titulares de cargos importantes y el estímulo a la ética profesional. Debido a la confianza que tenemos en ellas, podemos, entre otras cosas, creer en las estadísticas que esas instituciones generan. Ya de niños aprendemos a ser escépticos frente a las afirmaciones (incluidas las estadísticas) que aparecen en los avisos publicitarios: creemos que son tendenciosas y están tergiversadas. Por el contrario, confiamos más en las estadísticas que dan a conocer los científicos o los organismos públicos. Consideramos que esa información es más autorizada porque suponemos que dichas instituciones son imparciales y profesionales y se preocupan más de la exactitud de sus cifras.

Por ello podemos denominar *cifras acreditadas* a las estadísticas generadas por quienes consideramos una autoridad en una materia determinada, y que nos llegan por medio de canales institucionales que parecen velar por la exactitud de los números. En general, esas estadísticas evitan el tipo de torpes errores que revisamos en los capítulos anteriores. Las cifras calculadas por las autoridades suelen no tener errores de cálculo y, en general, los métodos de recolección y entrega de datos son apropiados. En su mayoría, cuando esas estadísticas llegan al público, ya han sido revisadas y chequeadas por colegas, evaluadores pares, editores y otros. Estas cifras son lo mejor que podemos encontrar en cuanto a estadísticas.

No obstante lo anterior, incluso las cifras acreditadas deben ser tratadas con cuidado. En un esfuerzo por identificar algunas de las preguntas que debieran formularse en relación con esas cifras, este capítulo revisa algunos ejemplos de estadísticas hallados en revistas

profesionales e informes gubernamentales. Comienza con una amplia discusión acerca de un artículo de una importante revista médica, que fue resultado del proceso de evaluación por pares.

HITOS DE LA INVESTIGACIÓN CIENTÍFICA

Cada mañana leo el *Wilmington News Journal*. Es el periódico más importante de Delaware, pero Delaware no es un estado grande y hay medios impresos mucho más importantes que el *News Journal*. Aun así, es un periódico moderno bastante típico. El 25 de abril de 2001, un artículo de portada resumía un informe publicado en la edición de esa semana del *Journal of the American Medical Association (JAMA)*. En él se planteaba que: “En EEUU, casi uno de cada tres alumnos entre 6º de enseñanza básica y 2º de enseñanza media había sido objeto de maltrato escolar entre pares (*bullying*), o lo ejercía sobre otros estudiantes”. Pero éste no era el único tema. Dos meses después, el *News Journal* publicó un reportaje acerca de otro artículo de *JAMA* que se titulaba: “20% de los niños que navegan en internet denuncian haber recibido propuestas sexuales por esa vía”. Un par de meses más tarde, un titular del *News Journal* se refería a otro artículo de *JAMA*: “Estudio sugiere que 1 de cada 5 muchachas sufre de abuso durante una cita”¹.

El *News Journal* no tiene los medios para contratar a un periodista que lea la edición semanal de *JAMA* y detecte historias de interés

¹ Los reportajes del *Wilmington News Journal* (todos tomados de agencias noticiosas) fueron: Lindsay Tanner, “Study Takes a Rare Look at Bullying’s Broad Effects”, April 25, 2001, p. A1; Tammy Webber, “Sexual Solicitation Reported by 20% of Kids Who Use Web”, June 20, 2001, p. A6; y Lindsay Tanner, “1 in 5 Girls Abused by a Date, Study Suggests”, August 1, 2001, p. A5. Los artículos de origen fueron: Tonja R. Nansel et al., “Bullying Behaviors Among U.S. Youth”, *Journal of the American Medical Association* 285 (April 25, 2001): 2094-2100; Kimberly J. Mitchell, David Finkelhor y Janis Wolak, “Risk Factors for and Impact of Online Sexual Solicitation of Youth”, *Journal of the American Medical Association* 285 (June 20, 2001): 3011-3014; y Jay G. Silverman et al., “Dating Violence Against Adolescent Girls and Associated Substance Abuse, Unhealthy Weight Control, Sexual Risk Behavior, Pregnancy, and Suicidality”, *Journal of the American Medical Association* 286 (August 1, 2001): 572-579.

noticioso: entonces, ¿cómo consigue esos artículos mi diario local? La respuesta es que se basa en las agencias de noticias. Pero eso genera otra interrogante: ¿cómo cubren las agencias noticiosas los avances científicos y médicos? Por ejemplo, *JAMA* envía reseñas o comunicados de prensa de la última edición con la esperanza de que, ante los ojos editores del diario, resulten aptos para su publicación. (No todas las revistas científicas hacen esto; la principal competencia de *JAMA* es el *New England Journal of Medicine*, que para publicitar sus resultados entrega a los medios una copia impresa previo a su lanzamiento)².

Es de suponer que *JAMA* quiere mostrarse ante el público como una revista que publica investigaciones relevantes. Entre otras cosas, esto hace de *JAMA* un medio deseable donde los investigadores más destacados ambicionan publicar y en el cual se puede dar a conocer el trabajo a los pares y a una mayor cantidad de público. Además, los investigadores que logran publicar sus investigaciones en revistas de gran prestigio, complacen a sus patrocinadores (organismos gubernamentales o fundaciones privadas que financian las grandes investigaciones). El saber que su apoyo económico culminó en publicaciones de alta visibilidad, confirma a los patrocinadores que su dinero fue bien utilizado.

Es importante tener presente que este es un proceso extremadamente competitivo. Para decidir cuáles proyectos apoyarán, las agencias de financiamiento realizan una selección muy acuciosa de los postulantes. Los potenciales autores envían 10 veces más manuscritos a *JAMA* de lo que esta revista efectivamente puede publicar; y sus editores no sólo deben decidir cuáles elegir, sino también optar por los que ameritan un comunicado de prensa. Los periódicos se ven saturados de comunicados de prensa y deben determinar cuáles podrán ocupar el limitado espacio disponible. Si una investigación llega a transformarse aun-

² En relación a los esfuerzos de las principales revistas médicas para obtener cobertura mediática de sus artículos, véase Ellen Ruppel Shell, "The Hippocratic Wars", *New York Times Magazine*, June 28, 1998, pp. 34-38.

que sea en un pequeño artículo en el *News Journal*, significa que ha superado varias etapas donde el rechazo es más común que la aceptación.

Recordemos que el *News Journal* publicó reportajes sobre artículos de *JAMA* que señalaban que el maltrato escolar afecta al 30% de los alumnos, que el 20% de los jóvenes que utilizan internet han recibido propuestas sexuales y que 20% de las estudiantes de enseñanza media han sido violentadas en una cita. A modo de experimento, imaginemos que cada uno de esos artículos consideró que el problema que trataba era un décimo —o incluso un tercio— de la frecuencia señalada. Es decir, imaginemos que el maltrato escolar entre pares afectaba a un 3-10% de los estudiantes y que las propuestas sexuales y la violencia ocurrían en un 2-7% de las citas. Ahora que los resultados son menos impresionantes, ¿el *News Journal* habría publicado artículos sobre esos estudios? Es probable que no. ¿Los editores de *JAMA* habrían hecho circular comunicados de prensa con esos hallazgos? De nuevo podríamos decir que no. De hecho, con esos resultados tan poco interesantes, podemos sospechar que los editores de *JAMA* habrían estado menos dispuestos a publicar los artículos, los autores no hubieran enviado los manuscritos a una revista tan selectiva y, ante los resultados publicados, las fuentes de financiamiento se habrían impresionado menos con la recepción. En otras palabras, podemos imaginar justificadamente que todos los involucrados en el proceso de publicación —los editores del *News Journal* y de *JAMA*, investigadores y patrocinadores— prefieren estudios que entregan cifras más impresionantes.

Permítanme aclarar que no estoy sugiriendo que haya algo fraudulento en el proceso. Es cierto que existen escándalos en que los investigadores han falseado sus hallazgos, pero no me refiero a eso. Más bien, quiero plantear que hay maneras más impactantes que otras de presentar los resultados de una investigación. Un informe que dé cuenta de un gran problema tiene más posibilidades de llamar la atención.

Veamos entonces cómo se producen las grandes cifras. Para eso consideremos el estudio acerca del maltrato entre pares. El artículo en

JAMA, “Bullying Behaviors Among U.S. Youth” (Conductas de maltrato entre jóvenes estadounidenses), entregó resultados a partir de una muestra representativa de estudiantes entre 6º de enseñanza básica y 2º de enseñanza media (casi 16 mil jóvenes). Los autores estaban asociados al Instituto Nacional de Salud Infantil y Desarrollo Humano (National Institute of Child Health and Human Development), el cual financió el estudio. La investigación recibió bastante cobertura de medios escritos, radio y televisión, los que destacaron el hallazgo de que un 30% de los jóvenes “denunciaban participación de moderada a frecuente en este problema”. Existen otras investigaciones respecto al tema, pero pocas cuentan con muestras tan grandes y bien diseñadas. Dada la composición y tamaño de la muestra, así como la publicación del artículo en una revista tan prestigiosa, podríamos considerar que es uno de los trabajos existentes más representativos sobre el maltrato entre pares.

Pero, ¿qué es exactamente el maltrato escolar entre pares? De acuerdo a una publicación federal: “Este fenómeno puede asumir tres formas: física (golpear, patear, escupir, empujar, quitar pertenencias personales), verbal (provocar, burlarse, poner sobrenombres, amenazar) y psicológica (difundir rumores, manipular las relaciones sociales, excluir socialmente, extorsionar o intimidar)”³. Cualquier persona que recuerde con claridad sus experiencias en la enseñanza media y lea esa definición, se sorprenderá de que sólo un 30% de los encuestados en el estudio de *JAMA* dijera sentirse afectado por este fenómeno. El *maltrato escolar entre pares* es un término tan vago como amplio, y es probable que mucho de lo que se clasifica como tal, se refiera a conductas relativamente comunes.

Aun así, la proporción de estudiantes que señalen haberse visto involucrados en estas conductas, dependerá de las preguntas que se les

³ Nels Ericson, “Addressing the Problem of Juvenile Bullying”, Fact Sheet 27, June 2001, U.S. Office of Juvenile Justice and Delinquency Prevention, Washington, DC.

formulen. En el cuestionario del artículo de *JAMA*, la sección relacionada con el maltrato comenzaba con la siguiente explicación:

Aquí encontrará algunas preguntas respecto al maltrato escolar entre pares. Decimos que un estudiante está *siendo maltratado* cuando otro estudiante, o un grupo de ellos, dice y/o hace cosas malintencionadas y desagradables a ese estudiante. También estamos frente al fenómeno cuando un estudiante es objeto de burlas repetidas de una manera que no es de su agrado. Pero *no es maltrato entre pares* cuando dos estudiantes de aproximadamente la misma estatura y sexo discuten o pelean (énfasis en el original)⁴.

Luego se les consultó a los estudiantes con qué frecuencia —en el período escolar considerado en el estudio— eran objeto de maltrato o maltrataban a otros. Se hicieron preguntas diferentes con relación al problema ocurrido dentro y fuera de la escuela, a pesar de que en el artículo de *JAMA* las respuestas fueron combinadas, no especificando el porcentaje al interior de los establecimientos. Para cada pregunta, las respuestas posibles describían las distintas frecuencias de involucramiento: “nunca”, “una o dos veces”, “en algunas ocasiones”, “una vez a la semana...” y “varias veces a la semana”. Al presentar los resultados, los autores definieron como participación frecuente las experiencias de al menos una vez a la semana, y como participación moderada, las de “en algunas ocasiones”.

Estas categorías fundamentan la base del hallazgo central de la investigación: aproximadamente el 30% de los jóvenes tiene una participación moderada o frecuente en esta conducta. Los autores concluyen que “este es un problema serio en la juventud estadounidense” y que “la prevalencia del maltrato escolar entre pares observado en el estudio sugiere la importancia de la intervención preventiva en relación con

⁴ Nansel et al., “Bullying Behaviors”, p. 2095.

estas conductas”⁵. En otras palabras, el problema está muy extendido y debemos hacer algo al respecto.

Sin embargo, ¿se puede decir que este artículo demuestra que el maltrato entre pares es un problema serio y extendido? El dato estadístico clave —que el 30% de los jóvenes está involucrado en ello— depende de tres manipulaciones, es decir, tres elecciones u opciones metodológicas:

1. Los estudiantes podían ser maltratadores (13% lo reconoció), víctimas (10,6%) o ambos (6,3%). El decidir contabilizar como “partícipes” del maltrato a quienes lo ejercían y a las víctimas, marcó una gran diferencia; si los autores hubieran optado por considerar sólo a las víctimas, el resultado habría sido en torno al 17% ($10,6 + 6,3 = 16,9$), y no el 30%.
2. Los autores incluyeron el maltrato “moderado” (que sucedía “en algunas ocasiones”, vale decir, más de 1 ó 2 veces en el actual período escolar, pero menos de 1 vez por semana) y el “frecuente” (que ocurría, al menos, semanalmente). Haber adoptado una definición menos amplia habría restado dramatismo a los resultados: sólo 8,4% de los consultados dijeron ser víctimas frecuentes, y no el 17%.
3. Recordemos que los autores combinaron las respuestas sobre maltrato entre pares ocurrido dentro y fuera del colegio. Si bien los investigadores no chequearon esta información, al menos algunos de los encuestados que señalaron ser hostigados frecuentemente habrían explicitado que eso sucedía sólo fuera de la escuela. De haber sido así, incluso menos del 8,4% habría reportado sufrir este problema de manera frecuente en el colegio. Es decir, los autores tomaron una serie de decisiones que les permitieron estimar que el problema afecta de manera importante al 30% de los estudiantes.

⁵ Ibid., pp. 2098, 2100.

Otras opciones —p.ej., si sólo nos concentráramos en las víctimas de maltrato frecuente al interior de los establecimientos educacionales— habrían resultado en una cifra cuyo valor, en el mejor de los casos, llegaba a la cuarta parte de eso.

Con esto no estoy sugiriendo que estemos frente a una investigación de mala calidad ni que el maltrato entre escolares carezca de consecuencias serias en determinadas situaciones. (Algunos informes sostienen que los responsables de los tan publicitados tiroteos en las escuelas actuaban de ese modo en respuesta al maltrato del que eran objeto). Pero las cifras que resultan de las investigaciones sociales deben interpretarse con sumo cuidado. El hallazgo de que un 30% de los estudiantes está involucrado en el maltrato, no debe entenderse como una especie de dato absoluto con existencia propia, sino más bien como el resultado de una serie de decisiones metodológicas. Esos resultados dependen de cómo se redactaron las preguntas, la secuencia que se siguió para formularlas y las decisiones que se tomaron al interpretar y resumir los resultados para su publicación. Los resultados tan comentados de los artículos de *JAMA* sobre propuestas sexuales en internet y violencia en las citas también se vieron afectados por opciones metodológicas.

También es importante conocer los intereses subyacentes a una investigación. Existe un movimiento contra el maltrato entre escolares que plantea que éste es un problema serio que no ha recibido la atención necesaria. Es muy posible que sin esa preocupación, el gobierno no hubiera financiado una investigación de tal envergadura. Desde luego, no existe algo así como un *lobby* a favor del maltrato escolar entre pares; nadie plantea que esta conducta sea algo deseable. Por consiguiente, podemos esperar que la mayoría de los investigadores que se dedican al tema busque demostrar que es un asunto serio y que los editores prefieran publicar artículos que apoyan esa idea.

No había nada poco profesional o malintencionado en el estudio de *JAMA*. Cualquiera que lea el artículo encontrará la información que

acabo de presentar. Sin embargo, para hacer un resumen o un comunicado de prensa, es necesario sintetizar el artículo, por lo que se destacan solamente algunos de los muchos hallazgos de una investigación. Recalcar la cifra de 30% hacía más llamativo el artículo, mientras que otros hallazgos, menos impresionantes, eran ignorados o desestimados en la cobertura de prensa. Por ejemplo, en el estudio de *JAMA* se da a conocer que el porcentaje de estudiantes que informaba haber sufrido estos actos de maltrato caía de manera drástica en tanto aumentaba la edad de los alumnos: 13,3% de los alumnos de 6° de enseñanza básica y sólo 4,8% de los de 2° de enseñanza media decían sufrir frecuentemente este problema. Por lo tanto, el maltrato entre pares disminuye a medida que los jóvenes maduran, algo nada sorprendente, pero que sí podría ser importante de considerar ante los urgentes llamados a tomar medidas antimaltrato de pares en las escuelas.

Al indagar cómo llegaron los resultados de esta investigación a la portada de mi periódico local, quiero destacar el papel de las opciones en la forma en que se reporta una investigación. Todos los estudios responden a una larga serie de decisiones y opciones. Los investigadores deben determinar qué es lo que quieren estudiar, elección que puede revelar sus propios intereses intelectuales, lo que sus colegas piensan que vale la pena investigar y su posible financiamiento. Lo anterior incluye decisiones metodológicas: cómo seleccionar una muestra y recolectar la información, cómo definir y medir los conceptos, cómo analizar e interpretar los resultados. Más aún, las opciones de investigación se ven restringidas por lo que ya se sabe sobre el tema y por los recursos de los que dispone el estudio, tales como tiempo, dinero, personal y otros. Pero esas elecciones siempre tienen consecuencias e inevitablemente influyen en los resultados. Cada investigación tiene sus limitaciones; siempre se puede argumentar que, de haberse tomado otras decisiones, los resultados habrían sido distintos. Por eso los científicos insisten en replicar las investigaciones (repetir un estudio para confirmar los resultados) y recopilar una gran cantidad de datos producto de estudios basados en distintas opciones. En la medida

que aumentan los estudios con resultados consistentes, crece la confianza en dichas conclusiones.

La mayoría de nosotros no sigue muy de cerca la expansión gradual del conocimiento científico. Más bien, nos informamos sobre los avances por medio de los resúmenes de estudios individuales que aparecen, por ejemplo, en las portadas de los periódicos. Y hay que pensar que también los periodistas han tomado decisiones: considerando todas las historias que compiten por ser dadas a conocer y el limitado espacio disponible. ¿Finalmente cuáles serán las historias que ameritan cobertura? En vista de estas restricciones, el avance continuo del conocimiento científico no parece ser particularmente atractivo para periodistas y editores, en tanto que un hallazgo que efectivamente parece revolucionario sí resulta ser una novedad. Los medios buscan lo dramático o de interés humano. Un artículo que destaque el maltrato escolar entre pares como un fenómeno extendido, parece ser una buena historia debido a que gran parte del público podrá considerarlo como un tema importante: esas personas tienen hijos o al menos conocen a algunos niños, lo que hace la investigación más interesante. (Por cierto, las revistas que entregan comunicados de prensa sobre sus artículos deben estar al tanto de los temas que preocupan a los medios; un buen comunicado de prensa debe llamar la atención de los periodistas hacia los aspectos relevantes de un estudio específico).

Y como si esto fuera poco, compliquemos más las cosas. Los científicos también tienen su propia agenda: generalmente desean reconocimiento y que sus investigaciones sean publicadas, lo que a la postre puede significar una cátedra permanente, promoción y financiamiento para más proyectos de investigación. Algunos comentaristas tienden a identificar los intereses de los investigadores con ideologías políticas, preocupados de que los estudios sean diseñados para apoyar posturas conservadoras o progresistas. Pero esa es sólo una pequeña parte de la historia. Los investigadores también pueden adherir a determinadas escuelas teóricas o metodológicas en sus disciplinas; por ejemplo, puede haber discusiones muy duras entre grupos que privilegian modelos estadísticos irreconciliables

utilizados para interpretar los resultados de un determinado estudio. Suele suceder que tales preferencias de los investigadores sean moldeadas por esas lealtades y preocupaciones que, por lo general, pasan inadvertidas —e incomprendidas— ante los ojos de quienes no pertenecen a ese circuito, es decir, la gran mayoría de las personas*.

Se supone que la literatura científica se corrige a sí misma. Los científicos saben que ningún estudio es perfecto, sin embargo, están convencidos de que, con el tiempo, el proceso de investigación producirá un conjunto de hallazgos en los que podremos confiar plenamente. Cuando los investigadores tienen reparos en relación con los resultados de un determinado estudio, pueden solicitar que la información sea revisada y que se realice un análisis alternativo, o pueden decidir que es mejor efectuar otro estudio. De la comunidad científica surge lentamente un conocimiento de mutuo acuerdo, vale decir, acreditado. Pero este proceso es más lento y más complejo que la forma usual en que la mayoría nos enteramos de los resultados de una investigación científica definida; es decir, gracias a los resúmenes que transmiten los contenidos de los comunicados de prensa.

AMBIGÜEDADES EN EL SISTEMA OFICIAL DE REGISTROS

Incluso la información recogida y compilada de la forma más profesional, puede ser malinterpretada. Tomemos en consideración los registros de muertes. En EEUU, la ley exige que para cada fallecimiento conocido haya un certificado de defunción; en cada caso se espera que alguien con autoridad, un médico o forense, asigne una causa de muerte. Estos registros llegan al Centro Nacional de Estadísticas de la Salud

* Un caso poco conocido en nuestro medio han sido las cifras de resultados de metanálisis de estudios terapéuticos con antidepresivos de última generación publicitados por algunos grandes laboratorios farmacéuticos. Lo esencial del engaño ha sido que sencillamente no se contabilizaron los hallazgos de investigaciones con resultados terapéuticos negativos. Véase, por ejemplo, “Drug Companies & Doctors: A Story of Corruption”, de Marcia Angelli, en *The New York Review of Books* (January 15, 2009) (N. del E.).

(NCHS, por su sigla en inglés) por los conductos burocráticos normales. Por su parte, esta entidad publica un informe anual, *Vital Statistics of the United States*, que resume los registros de nacimientos y muertes. Este informe, que antes consistía en tres gruesos volúmenes tipo guía telefónica, ahora está disponible en formato digital y cualquier persona puede acceder a él a través del sitio web del NCHS. Hace tiempo que el gobierno recopila estos datos y ha corregido casi completamente los errores de softwares del sistema. Contabilizar nacimientos y muertes es una tarea bastante sencilla, y esta información es posiblemente lo más completa, exacta y acreditada que podríamos esperar⁶.

Y a pesar de todo, las desventuras pueden ocurrir, y de hecho ocurren. Puede haber problemas. Consideremos, por ejemplo, los reportajes que indican que la tasa de suicidios de afroamericanos jóvenes a fines del siglo XX había aumentado. En 1998, los Centros para el Control y la Prevención de Enfermedades (CDC, por su sigla en inglés) informaron que entre 1980 y 1995, se había duplicado la tasa de suicidios entre los afroamericanos de 10 a 19 años, y que esta alza era mucho mayor que la que se daba entre los jóvenes de raza blanca⁷. Fue un hallazgo preocupante. En la mayoría de los adultos, el suicidio adolescente causa un impacto enorme, por ser un acto que expresa aislamiento, angustia y desesperanza. Además, la mayoría de los estadounidenses querría pensar que las relaciones raciales no han estado empeorando. Entonces, ¿cuáles serían, hacia el fin del siglo, los motivos del alza de la tasa de suicidio entre los adolescentes negros?

Los reporteros que tomaron los datos del CDC intentaron explicar la aparente tendencia consultando a los expertos que de alguna manera

⁶ A pesar de la excelente calidad de la información de este informe, aún persisten ambigüedades. Para una discusión sobre estos temas, véase William Petersen, *From Birth to Death: A Consumer's Guide to Population Studies* (New Brunswick, NJ: Transaction, 2000).

⁷ U.S. Centers for Disease Control and Prevention (CDC), "Suicide Among Black Youths: United States, 1980-1995", *Morbidity and Mortality Weekly Report* 47 (March 20, 1998): 193-196.

son los encargados —en nuestra cultura— del tema del suicidio: psiquiatras y psicólogos clínicos. Estos especialistas hablaron de un “síndrome de esclavitud postraumático que puede manifestarse en una amplia gama de conductas autodestructivas”. También mencionaron: “Rupturas familiares, escasas oportunidades económicas, depresiones no diagnosticadas, pena encubierta provocada por la violencia urbana e, incluso, estrés adicional por ingresar a la clase media”; por ende, “las familias afroamericanas emergentes quizás carezcan del apoyo familiar y comunitario tradicional”⁸. El informe del CDC sugería: “Es posible que los jóvenes negros de familias que ascienden socialmente, adopten del resto de la sociedad algunas conductas comunes en ésta para enfrentar la depresión y la desesperanza, como es el caso del suicidio”⁹. En otras palabras, hubo muchas explicaciones tardías.

Sin embargo, también es posible que este aumento no esté reflejando un cambio en la conducta de los jóvenes afroamericanos, sino más bien una variación en cómo las autoridades procesan sus muertes¹⁰. Veamos la Figura 11, que presenta la cifra de jóvenes afroamericanos de 10 a 19 años que se suicidaron entre 1970 y 1999. (Para ordenar lo más claramente posible los pasos de mi argumentación, decidí graficar las cifras actuales de muertes registradas, en lugar de las tasas de muerte. Las tasas no cambiarían el patrón de la información en lo que respecta al punto que estoy discutiendo). Tal como indica el

⁸ Angela G. King, “Suicides Up Among Blacks”, *New York Daily News*, October 22, 2000, p. 32; M.A.J. McKenna, “Black Teen Suicide Rate Skyrockets, CDC Says”, *Atlanta Constitution*, March 20, 1998, p. 1A; Mary Jo Kochakian, “Experts Can’t Explain Rise in Suicide Rate Among Black Teens”, *Hartford Courant*, March 31, 1998, p. F4.

⁹ U.S. Centers for Disease Control and Prevention, “Suicide”, p. 195.

¹⁰ Para discutir esto, he tomado muchas ideas del análisis realizado por el heterodoxo sociólogo Mike A. Males, cuyos libros desafían los estereotipos mediáticos sobre la juventud. Por ejemplo, véase Males, *Framing Youth: Ten Myths about the Next Generation* (Monroe, Maine: Common Courage, 1999). Para recalcular estos datos utilicé clasificaciones algo distintas a las que usa Males. Compárese Males, “The Myth of the (Black) Teen Suicide Epidemic”, en www.alternet.org/story.html?StoryID=11149, July 10, 2001; e información adicional en <http://home.earthlink.net/~mmales/>

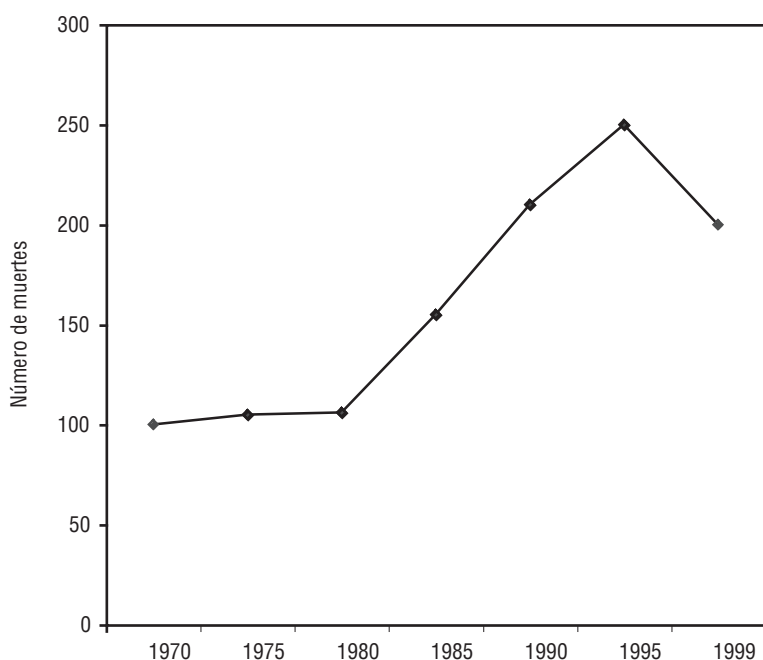


Figura 11. Suicidios de jóvenes afroamericanos entre 10 y 19 años, 1970-1999.
(Fuente: National Center for Health Statistics).

gráfico en la Figura 11, no estamos hablando de muchos casos: según el NCHS (que, recordemos, registra *todas* las muertes), en 1970 hubo algo más de 100 suicidios en este grupo y en 1995 fueron alrededor de 250; en tanto el censo de 2000 identificó a 6,2 millones de afroamericanos entre 10 y 19 años.

La Figura 12 reproduce el primer gráfico, pero agregando una segunda línea que muestra la cantidad de muertes en este grupo etario consignadas como “sin causa determinada”. En la lista del NCHS, “indeterminada” es una categoría residual para aquellas muertes que no tienen una causa específica. La comparación entre ambas líneas es impactante: mientras la línea de los suicidios subió hacia fines del siglo XX, el número de muertes por causas indeterminadas cayó.

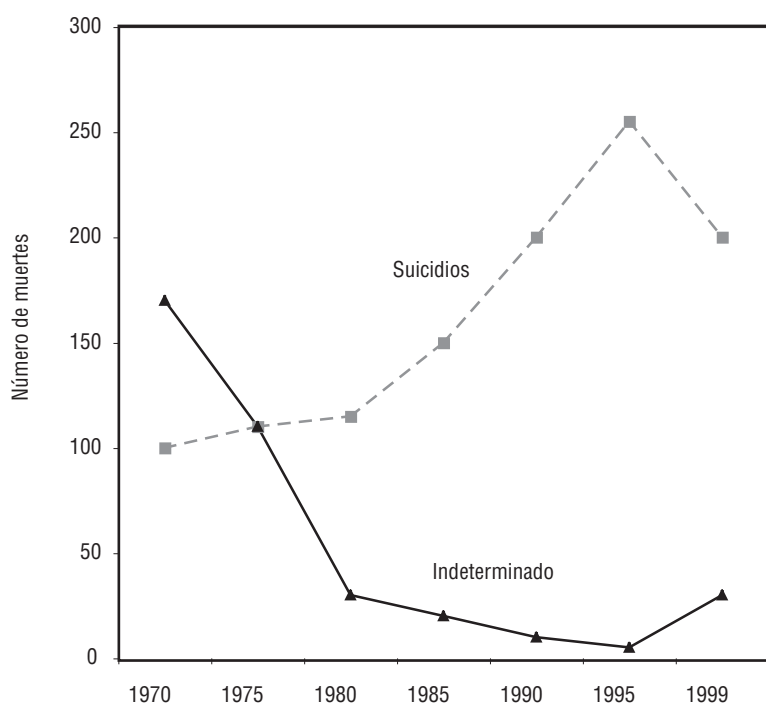


Figura 12. Suicidios y muertes por causas indeterminadas de jóvenes afroamericanos entre 10 y 19 años, 1970-1999. (*Fuente:* National Center for Health Statistics).

La Figura 13 agrega otra dimensión al análisis. Muestra las muertes registradas en el mismo período atribuidas a las siguientes causas accidentales: asfixia por inmersión, disparos, envenenamiento y caídas. Elegí estas cuatro categorías de accidentes fatales, pues también son formas habituales de suicidarse. Por ejemplo, cuando un funcionario certifica la muerte de un joven que se ahogó, esa muerte puede quedar registrada como un accidente o como un suicidio. La Figura 13 revela que en el período mencionado, las cuatro categorías de muerte accidental entre los jóvenes afroamericanos habían declinado hacia fines del siglo.

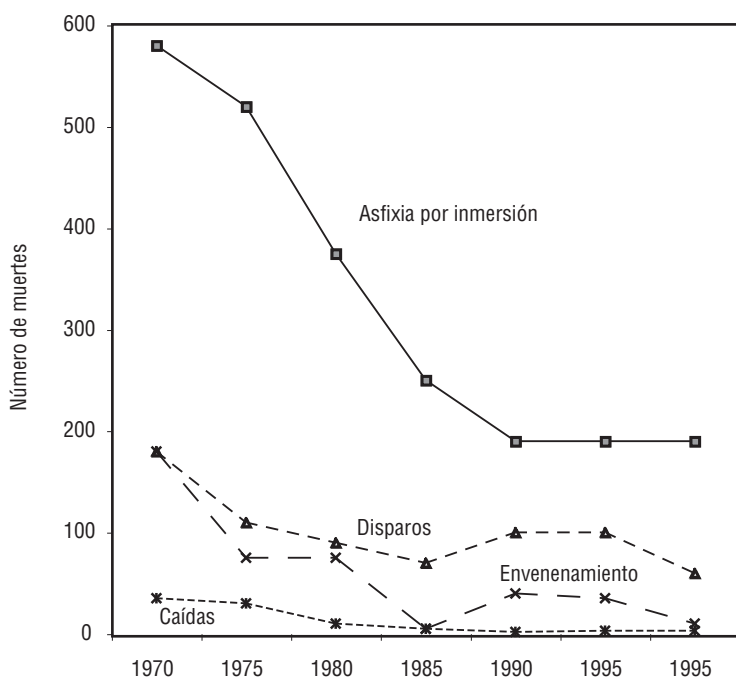


Figura 13. Muertes accidentales como resultado de asfixia por inmersión, disparos, envenenamiento y caídas de jóvenes afroamericanos entre 10 y 19 años, 1970-1999. (Fuente: National Center for Health Statistics).

Finalmente, la Figura 14 combina toda la información de los tres gráficos anteriores en uno de barras. Cada barra está dividida en tres segmentos: el inferior corresponde a los suicidios; el del medio, a las muertes por causas indeterminadas; y el superior, a las defunciones por las cuatro causas accidentales ya mencionadas. La altura total de cada barra representa el número de muertes atribuidas a la sumatoria de todas estas causas. Una vez más, vemos que el número de suicidios entre jóvenes negros aumentó en el período, aun cuando las muertes debidas a las categorías discutidas disminuyeron; vemos que, en términos generales, la cifra total de muertes debidas a estas causas se redujo

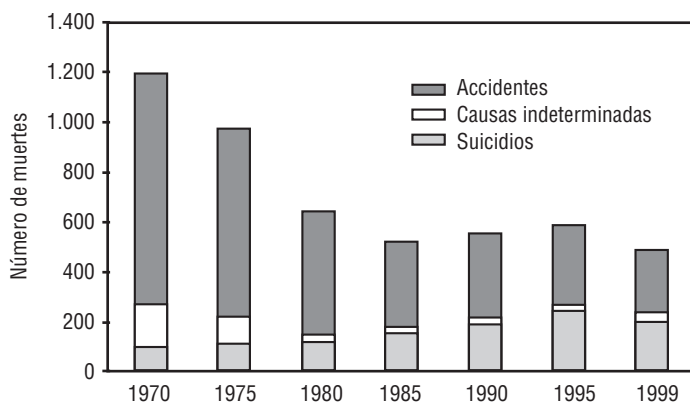


Figura 14. Muertes debidas a cuatro causas accidentales, causas indeterminadas y suicidios de afroamericanos entre 10 y 19 años, 1970-1999. (*Fuente:* National Center for Health Statistics).

a la mitad. Esto es muy importante, porque todas estas defunciones representan incidentes que podrían clasificarse como suicidios (dependiendo, desde luego, de que hubiera otra información disponible). En todo caso, probablemente sea más difícil clasificar una muerte como suicidio que asignarle alguna otra causa, ya que para los familiares es complejo admitir un suicidio. Es más sencillo clasificar una muerte como accidental, o incluso sin causas determinadas.

En otras palabras, el misterio del incremento de los suicidios adolescentes entre los afroamericanos puede resultar ser no tan misterioso. Sus orígenes podrían radicar no tanto en las presiones psicológicas sufridas por jóvenes de familias afroamericanas emergentes, como en un cambio en el modo en que las autoridades clasifican sus causas de muerte. Mientras que antes esas muertes pudieron haber sido consideradas relativamente poco importantes —pasando como accidentes o por una causa indeterminada—, ahora las autoridades pueden llevar a cabo diligencias más cuidadosas para llegar a asignar al suicidio como

causa concreta¹¹. Obviamente, no podemos *probar* que esta sea la explicación correcta del alza en el número de suicidios; ello implicaría revisar las evidencias utilizadas para asignar las causas de muerte en miles de casos, y lo más probable es que la mayor parte de esa información esté extraviada hace tiempo. Sin embargo, este ejemplo nos hace recordar que incluso la información más completa y autorizada —como son los registros de muertes del NCHS— no puede hablar por sí misma.

Más bien, lo que ocurre es que las cifras deben ser interpretadas. En este caso, alguien en el CDC notó un aumento en la cantidad de muertes clasificadas como suicidios entre adolescentes afroamericanos y decidió que era reflejo de un incremento genuino en la conducta suicida. A su vez, ese comportamiento debía ser explicado identificando los cambios en la vida de los jóvenes que habrían influido en su mayor tendencia al suicidio. Un aumento en el número de muertes clasificadas como suicidios no necesariamente refleja una mayor cantidad de actos autodestructivos; también puede revelar cambios en la manera en que las autoridades clasifican las muertes como suicidios. Dado que las cifras provenían de registros oficiales aparentemente acreditados, la mayoría de los comentaristas fallaron al no cuestionar el incremento, a pesar de que necesitaron de explicaciones muy enrevesadas para dar cuenta de ello. Más aún, las estadísticas que asustan relacionadas a la temática racial son tan comunes que muchas personas simplemente asumen su veracidad.

Incluso el registro aparentemente más sencillo de datos puede terminar siendo extremadamente complejo. Consideremos un segundo

¹¹ Un hecho revelador: el CDC se percató de que los suicidios de adolescentes negros habían aumentado notoriamente en el sur de EEUU. Males plantea que un análisis por separado de los registros del estado de California, muestra pocos cambios a lo largo del tiempo. Este estado tiene, desde hace años, un programa profesional para registrar las muertes anuales. Por el contrario, los estados del sur han sido más lentos en asignar esta tarea a profesionales del área médica; en muchas jurisdicciones sureñas, la labor del forense ha sido un cargo electivo, usualmente llevado a cabo por un empresario de pompas fúnebres. Véase Males, “The Myth of the (Black) Teen Suicide Epidemic”.

ejemplo: el esfuerzo para crear la nómina de muertes atribuibles a los ataques terroristas del 11 de septiembre de 2001 en el World Trade Center de Nueva York. Los registros de las aerolíneas permitieron dar los nombres y contabilizar casi inmediatamente a quienes iban en ambos vuelos. ¿Cuántas personas murieron al caer las torres? Nadie posee una lista de control de los individuos que se encuentran al interior de un rascacielos. Incluso durante un día laboral cualquiera, los que están allí —empleados en sus oficinas, visitas, clientes— conforman una población flotante bastante grande. El choque de los aviones contra los edificios implicó la evacuación de miles de personas, al tiempo que cientos de bomberos, policías y demás personal de emergencia ingresaba a las torres. Además, cuando los edificios cayeron, la destrucción fue tal que muchos cadáveres desaparecieron sin dejar rastro alguno.

En este caso, contabilizar los muertos se convirtió en un asunto muy espinudo. A los pocos días, las autoridades tenían varios listados de personas desaparecidas. Los nombres provenían de empresas que entregaron listas de empleados que presumiblemente estaban en los edificios, así como de amigos y familiares que carecían de noticias de gente que sospechaban podía encontrarse en el World Trade Center. Los informes siguieron llegando hasta el 24 de septiembre, fecha en que las listas alcanzaron a contabilizar 6.453 nombres.

Luego la lista comenzó a decrecer. Las autoridades detectaron nombres que se repetían (p.ej., había una decena de informes relativos a una misma mujer, cada uno de los cuales tenía una dirección y número telefónico de contacto diferentes). Hubo personas reportadas como desaparecidas que estaban vivas (se localizó a más de 1.500 extranjeros dados por desaparecidos por diversas embajadas). También se ubicaron alrededor de 70 casos fraudulentos de individuos que pretendían recibir los beneficios de los sobrevivientes. Además, se agregó un puñado de nombres de sujetos que habían sido enviados a hospitales fuera del estado y que luego murieron. Para el 11 de septiembre de 2002, es decir, un año después, el total había disminuido a 2.801 personas,

incluidos 35 a 40 individuos de quienes no se sabía si estaban vivos o no. Aun después de un año de ardua investigación, todavía no se tenía un total exacto; y, de hecho, el conteo seguía cambiando¹².

En el otoño de 2001, la nómina de muertos se transformó en un tema polémico. Durante un par de semanas, algunas autoridades seguían repitiendo los cálculos iniciales que hablaban de 5 ó 6 mil muertos; con ello buscaban justificar las represalias en contra de los terroristas. Incluso hubo quienes criticaron las notas de prensa que al comienzo predijeron (de manera acertada) que el número final de muertos sería mucho menor. Es un hecho que lo horroroso del ataque no depende de la cifra de muertos; la nómina final terminó siendo la mitad de lo que se calculó inicialmente, pero eso no aminora en absoluto la dimensión de la tragedia.

En tanto algunas autoridades fueron muy cuidadosas con el conteo de las víctimas fatales, otras difundieron datos bastante discutibles acerca de la magnitud de la catástrofe; decían que el ataque al World Trade Center había dejado huérfanos a 10 mil —o incluso 15 mil— niños, la mayoría de los cuales necesitaría ser adoptado. Este cálculo no resistía ningún análisis. Aunque consideráramos la cifra más alta de fallecidos (6.500), un total de 15 mil huérfanos implicaba que cada una de las víctimas debía tener un promedio de más de 2 hijos. Aún más, si tomamos la acepción típica de huérfano —un menor de edad que ha perdido a padre y madre—, es obvio que esa afirmación era improbable; muchos de esos hijos debían ser adultos, no todas las víctimas

¹² Eric Lipton, periodista del *New York Times*, escribió una serie de artículos documentando esta variación en la nómina de muertos. Por ejemplo, véanse los siguientes artículos de Lipton para el *New York Times*: “Toll from Attack at Trade Center Is Down Sharply”, November 21, 2001, p. A1; “In Cold Numbers: A Census of the Sept. 11 Victims”, April 19, 2002, p. A14; “Death Toll Is Near 3,000, but Some Uncertainty Over the Count Remains”, September 11, 2002, p. G47; y “Sept. 11 Death Toll Declines as Two People Are Found Alive”, November 3, 2002, p. A17. Aparte de las personas que murieron en el World Trade Center, 184 fallecieron en el ataque al Pentágono y 40 más en el accidente del cuarto avión en Pensilvania, pero era más sencillo contabilizar e identificar a esas otras víctimas. Las muertes de los 19 secuestradores también fueron confirmadas fácilmente.

tenían hijos y la mayoría de los que estaban casados habrían dejado un cónyuge que continuaría cuidando a los hijos. Si bien miles de familias sufrieron la pérdida de un ser querido, las autoridades de Nueva York vinculadas a temas familiares no pudieron identificar a un solo niño que necesitara adopción o una familia sustituta¹³. El ataque al World Trade Center fue pavoroso; así y todo, circularon estadísticas que exageraban el alcance del daño.

He optado por centrarme en las estadísticas de muertes porque parecen más sencillas: es más fácil contar muertos que medir la pobreza, el desempleo, el delito y la mayoría de las otras cosas que las autoridades contabilizan. Las estadísticas oficiales suelen ser las cifras más completas, las mejores —las más acreditadas—, pero eso no significa que sean exactas.

Las autoridades oficiales tienen grandes ventajas al recolectar estadísticas. En comparación con los proyectos de investigación implementados por los científicos, las agencias de gobierno cuentan con presupuestos abultados que les permiten contratar personal para recoger, compilar, analizar e interpretar la información. La obligatoriedad de colaborar con la recolección de datos puede ser impuesta por ley; los ciudadanos deben aportar información para el censo, así como registrar los nacimientos y muertes. En lo que a datos se refiere, las estadísticas oficiales tienden a ser relativamente completas.

Pero los registros oficiales son producto del sistema político y por eso están irremediablemente moldeados por consideraciones políticas. Cada decisión relativa a recoger información oficial puede ser objeto de discusión política. ¿Qué información necesitamos? ¿Cómo debemos recogerla? ¿De qué manera la recopilamos? ¿Cuáles resultados damos a conocer? ¿Cómo hacemos eso y a quiénes informamos los hallazgos? ¿Qué tipo de recursos debemos destinar a este proceso? Recolectar

¹³ Nina Bernstein, "Thousands of Orphans? An Urban Myth", *New York Times*, October 26, 2001, p. B1.

información toma tiempo y cuesta dinero, por lo que suponemos que alguien considera que el esfuerzo de la recolección de datos bien vale la pena. En algunos casos puede haber un amplio acuerdo respecto a que esa información debe recogerse, pues sirve a un interés general; por ejemplo, la mayoría de la gente aprueba que se registren los nacimientos y las defunciones. La gente incluso puede estar de acuerdo en lo que debe contabilizarse y cómo hacerlo.

Sin embargo, muy a menudo las cosas son más complicadas que eso, con intereses que compiten entre sí para dar forma a las estadísticas. Por ejemplo, en el Capítulo 1 se mencionaba que, con motivo del censo y para aumentar las cifras de sus grupos, las minorías étnicas tienden a promover la recolección de datos acerca de la pertenencia a un grupo determinado. También está el caso del Índice de Precios al Consumidor (IPC). El IPC usualmente se utiliza como índice para reajustar los beneficios gubernamentales y aquellos beneficios pactados en negociaciones colectivas según el alza en el costo de la vida. Esto transforma el método de cálculo del IPC en una materia que trasciende el interés académico. Los programas de gobierno y los empleadores que deben pagar salarios a sus empleados basándose en la variación del IPC, tienden a estar más de acuerdo con cálculos que minimizan el crecimiento del índice; en tanto aquellos cuyas ganancias o beneficios están ligados al aumento del IPC, favorecen los que maximizan el crecimiento del IPC. Los economistas que sugieren modos de alterar la fórmula del IPC para reflejar los cambios en el estilo de vida de las personas (tales como el ajuste al impacto de los computadores personales o la telefonía celular), verán que su trabajo es criticado no sólo en términos académicos, sino también por razones políticas¹⁴.

Y, por supuesto, la opinión de las autoridades respecto a su papel puede variar. En un extremo están quienes se ven a sí mismos como

¹⁴ Dean Baker, ed., *Getting Prices Right: The Debate Over the Consumer Price Index* (Armonk, NY: M.E. Sharpe, 1998); Jolie Solomon, "An Economic Speedometer Gets an Overhaul", *New York Times*, December 23, 2001, sec. 3, p. 4.

profesionales imparciales, que recogen estadísticas de manera objetiva. En el otro lado, hay quienes se consideran agentes activos de algún grupo, como la actual administración política, y pueden tratar de generar estadísticas que apoyen sus políticas. (Nótese que esto no requiere llegar al fraude o al engaño descarado: puede tratarse de algo tan inocuo como optar por contabilizar algunas cosas y no otras, o difundir determinadas cifras y enfatizar su relevancia). Es probable que la mayoría de los funcionarios se ubiquen en algún lugar intermedio; intentan ser competentes y rigurosos en su trabajo, y, sin embargo, suelen considerar que su trabajo está permeado por sus propios intereses e inclinaciones o por presiones políticas provenientes de otros.

LA FRAGILIDAD DE LA ACREDITACIÓN

Las estadísticas acreditadas dependen de la confianza que tengamos en las instituciones que las generan. Por ejemplo, los contadores certifican que los registros financieros de una empresa estén en regla, lo que da seguridad a los inversionistas de disponer de la información necesaria para tomar decisiones financieras apropiadas en relación con esa empresa. En 2001-2002 se supo que Enron y otras grandes empresas habían llegado —con la aprobación de sus contadores— a ciertos acuerdos financieros, lo que provocó tal escándalo que no sólo arruinó a las empresas directamente involucradas, sino que amenazó la confianza de los inversionistas en la economía del país. A partir de esa experiencia, el gobierno aprobó una ley de reforma a las empresas que exigía a la plana ejecutiva mayor de las grandes sociedades anónimas certificar personalmente, bajo pena de sanción criminal, que los registros de sus empresas fueran auténticos. En otras palabras, dado que uno de los niveles de protección institucional había resultado insuficiente para garantizar la exactitud y confiabilidad de los registros financieros, fue necesario crear otro nivel de certeza. La confianza en la autoridad depende de tales señales.

Esas garantías pueden parecer acuerdos sociales bastante frágiles. El común de las personas no puede chequear o replicar las cifras que producen los científicos, funcionarios de gobierno, contadores y otras autoridades; el costo en tiempo y dinero sería excesivamente alto. En vez de ello, confiamos en el profesionalismo de esas autoridades y en su promesa de velar por los intereses de sus clientes, la ley, sus pares y ellos mismos, para generar cifras de la mejor calidad posible. Las estadísticas provenientes de países más pobres que carecen de recursos para financiar la recolección y análisis de datos, son poco más que conjeturas¹⁵; pero una sociedad rica en recursos espera de sus autoridades —y generalmente obtiene— estadísticas de buena calidad. En EEUU, la cantidad de malas estadísticas constituyen un escándalo. Recordemos lo que sucedió cuando, luego de las elecciones del 2000, la gente se dio cuenta de que incluso los sistemas mecanizados de conteo de votos podían inducir a error (p.ej., el no contabilizar votos con perforaciones incompletas). Asimismo, los informes sobre fraudes científicos, grandes errores contables o funcionarios que llevan registros imprecisos, se transforman en temas de gran interés para los medios.

A pesar de nuestras expectativas y la confianza que depositamos en las estadísticas acreditadas, los ejemplos presentados en este capítulo demuestran que éstas tienen sus límites. La recolección de datos nunca es perfecta; la “cifra oscura” de los casos ocultos y no contabilizados siempre está presente¹⁶. *Todos los análisis implican decisiones acerca de qué contabilizar y cómo hacerlo, y esas opciones siempre afectan las cifras resultantes*. Esas preferencias suelen reflejar presión por parte de las autoridades. En algunos casos, la presión proviene de una sola dirección, lo que nos induce a apoyar el mismo conjunto de opciones; pero en otros casos, las presiones en competencia provienen de quienes esperan, por

¹⁵ Por ejemplo, véase Norimitsu Onishi, “African Numbers, Problems, and Number Problems”, *New York Times*, August 18, 2002, sec. 4, p. 5.

¹⁶ Acerca de la cifra oscura, véase Best, *Damned Lies and Statistics*, pp. 33-34.

ejemplo, una cifra abultada o, por el contrario, de quienes prefieren un cálculo más bien escueto. No olvidemos que las autoridades también tienen expectativas mutuas: para mejorar la calidad de la investigación publicada, los científicos utilizan el proceso de revisión por pares; para evaluar las cuentas, los contadores emplean parámetros sobre los cuales hay un acuerdo general; etc. Es inevitable que incluso las estadísticas más acreditadas y prestigiosas reflejen de alguna manera todos estos procesos sociales. Las personas podrán discrepar, y de hecho lo hacen, acerca de la mejor forma de implementar un censo, medir la cesantía o evaluar el peligro de maltrato entre escolares. El contabilizar, aun cuando genere estadísticas acreditadas, es un proceso social.

Resumiendo, la pregunta que hay que hacer en relación con cualquier número —incluso aquellos que parecen más acreditados— no es: “¿Será verdad?”. Más bien, debemos preguntar: “¿Cómo se llegó a este resultado?”. Si algunas cifras son más acreditadas que otras, es porque tenemos más confianza en el proceso que las produjo. Pero esto no equivale a decir que deberíamos imaginar que cualquier cifra ofrece soluciones mágicas a nuestros problemas.

Capítulo 5

CIFRAS MÁGICAS

Cualquier persona informada sabrá que las recesiones económicas existen. En los buenos tiempos, los comentaristas especulan sobre el riesgo de que se inicie una; en los tiempos malos, se preguntan si la recesión en curso ya estará por concluir. Quien decide si éstas comienzan o terminan es el Business Cycle Dating Committee de la Oficina Nacional de Investigación Económica del gobierno federal. Este comité, usualmente anónimo, hizo noticia en el verano de 2003 cuando propuso modificar los criterios utilizados para definir cuándo una recesión estaba por concluir¹.

Para identificar cuándo empiezan y terminan las recesiones, el comité usaba varios indicadores mensuales de la actividad económica, incluyendo la cantidad de personas empleadas. Esta fórmula no incluía la medición del producto interno bruto (PIB, o valor total de los bienes y servicios producidos en EEUU) porque era medido trimestral y no mensualmente. Dado que las recesiones previas habían estado marcadas por bajas, tanto en el empleo como en el PIB, el no incluir este último indicador no influía en la determinación de su inicio o término.

¹ John M. Berry, "Number Crunchers vs. Recession: Seeking Official End, Panel Wrestles with One Stubborn Stat", *Washington Post*, July 11, 2003, p. E1; y Berry, "Recession Ended in November of 2001, But Panel Still Sees Economy Slumping", *Washington Post*, July 18, 2003, p. E1.

Sin embargo, la recesión que comenzó a fines de 2001 rompió la tendencia: gracias a la mejoría en la productividad, a finales de 2002 el PIB empezó a aumentar, pese a que el empleo seguía bajando. Dado que la fórmula utilizada por el comité se basaba en la medición del empleo que seguía cayendo, la evaluación oficial indicaba que la recesión todavía no había terminado, no obstante muchos observadores creían que la economía había tocado fondo hacía unos meses, justo antes de que el PIB comenzara a subir. Desde entonces, el comité decidió incorporar a sus cálculos las estimaciones mensuales del PIB, un paso que llevó a declarar que la recesión había terminado, pero el comité reconocía que aún había disminución en el empleo.

Una vez más, vemos cuánto impactan las decisiones respecto a qué contabilizar. Bajo la antigua fórmula utilizada por el comité, la economía de 2003 continuaba recesiva; pero según la nueva propuesta, la recesión estaría superada. El comité traduce cifras —en este caso, medidas económicas— en rótulos oficiales acerca del estado de la economía. Al hacerlo, el comité le otorga gran importancia a esos números.

Nuestra cultura depende de las cifras y, por lo tanto, las trata con seriedad. Si bien sospechamos que las estadísticas son imperfectas, sabemos que no podemos vivir sin ellas. La economía, y el resto de nuestro mundo, son demasiado complejos como para comprenderlos sin recurrir a los números. Necesitamos de las estadísticas para tener una base que nos ayude a comprender lo que sucede y tomar decisiones. Contar y medir puede ayudarnos a decidir qué hacer. Cuando nos llama la atención algún nuevo problema social, el primer impulso es cuantificarlo, es decir, medir su alcance. Decimos que las estadísticas nos permiten “agarrar” el problema “por el mango”, como si el llevarlo a cifras, a números, de alguna manera nos diera los medios para controlarlo o siquiera indicarnos cómo llegar a controlarlo. En la práctica, actuamos como si las cifras tuvieran poderes inauditos para esclarecer y poner en evidencia las decisiones correctas: como si realmente tuvieran propiedades *mágicas*.

Entonces, números mágicos son aquellos que imaginamos exactos y acreditados, que prometen hacer comprensibles y manejables nuestros problemas. Al darnos una base para tomar decisiones complejas, los números mágicos supuestamente pueden cambiar la ambigüedad en certeza. Nos dan un estándar a partir del cual podemos formarnos una idea del mundo. O, al menos, eso es lo que nos gusta creer.

Esto sugiere que debiéramos estar alertas ante cifras mágicas que seguramente aparecerán ahí donde están las fallas geológicas de nuestra cultura, aquellas áreas complejas donde el conflicto, la incertidumbre y la preocupación son especialmente intensos, donde tenemos la necesidad de un fundamento más sólido para fundamentar nuestros actos. Por ejemplo, cuando alguien señala un problema social, nos está obligando a enfrentar aseveraciones que dicen que nuestra sociedad no está tan bien como debiera, que hay que hacer algo para que las cosas cambien. Nuestra cultura aspira a la perfectibilidad: de ahí que insistamos en tareas complejas y difíciles como esa obligación que nos hemos impuesto: “Que ningún niño se quede atrás”; y, como si fuera poco, también le declaramos la guerra a la pobreza, a las drogas e incluso al cáncer. Dadas estas pretensiones tan nobles, al destacarse un determinado problema social, salta a la vista una crítica que exige acción. Por cierto, no todas las personas estarán de acuerdo en si ese problema específico amerita atención, o pueden no coincidir respecto a las soluciones más apropiadas. Entonces, no es de extrañar que las discusiones en torno a temas sociales casi siempre incluyan estadísticas.

Quienes defienden una postura en particular suelen recurrir a las cifras para reafirmar sus argumentos, para hacerla aparecer como más verídica. Recordemos que, en nuestra cultura, se supone que las estadísticas son factuales; las cifras sugieren que alguien cuantificó el

* “*Leave No Child Behind*”, propuesta de ley, aprobada en EEUU en 2001, que estipulaba que los programas de educación pública básica y media debían, entre otras cosas, mejorar sus estándares de rendición de cuentas y responsabilidad en relación con la calidad y el acceso (N. de la T.).

problema y conoce su dimensión. Ellas (las cifras) pueden hacernos sentir menos confundidos sobre qué es lo que se debe hacer.

Es posible que en situaciones sin ambigüedades y que no dejan lugar a dudas, las cifras no sean necesarias. Cuando las acciones de las personas son regidas por rituales, por las órdenes emanadas de quien detenta el poder o por las convenciones morales compartidas, se reduce el terreno donde elegir, para las dudas y la angustia. No obstante, nuestro mundo moderno se caracteriza por su complejidad y diversidad, por afirmaciones irreconciliables y estándares cambiantes. La incertidumbre es algo común en nuestra vida y solemos buscar en las estadísticas su mágica capacidad de esclarecer, de transformar la inquietud en confianza, de convertir lo borroso en hechos. Es más, estas estadísticas ni siquiera tienen que ser números particularmente buenos. Tendemos a creer que cualquier cifra es mejor que ninguna y, en ocasiones, para disminuir nuestra confusión, nos aferramos al primer dato disponible. El problema es que una certeza inspirada por cifras mágicas puede ser en realidad una guía bastante inadecuada para tomar decisiones respecto al mundo real.

Este capítulo revisa tipos de cifras que ocasionalmente adquieren cualidades mágicas; son ejemplos que nos pueden ayudar a comprender la naturaleza de lo que yo denomino los números mágicos. Comenzaremos con una decisión y una elección que, tarde o temprano, la mayoría de las familias deben enfrentar.

EL NÚMERO ESPECIAL DE TRAJES DE BAÑO

Elegir una universidad es un proceso que provoca angustia. El costo, en dinero, de una carrera de pregrado de cuatro años es alto y sube cada año. Si bien se puede decir lo mismo del costo de un auto nuevo, los clientes que entran a una automotora con dinero suficiente rara vez son rechazados; sin embargo, la mayoría de quienes postulan a los planteles de élite son rechazados. Esta incertidumbre —“¿Quedaré seleccionado?”—

lleva a los estudiantes, en su intento por asegurar su ingreso, a postular a varias universidades. Aquellos que han sido aceptados por más de una, tienen la posibilidad de elegir entre las diversas alternativas.

Existen cientos de universidades e instituciones de educación superior y el proceso mismo de postulación es costoso. ¿Qué hacer? ¿A cuáles deben postular los estudiantes? En el último tiempo, ha surgido una pequeña industria de guías para elegir universidades. Entre ellas, se destaca *U.S. News & World Report*, que cada otoño publica un número especial para estudiantes donde los planteles han sido clasificados en base a información estadística. Esta edición, que vende muchísimos más ejemplares que su versión semanal, es conocida por las oficinas de admisión de las universidades como el “número especial de trajes de baño”^{**}. La clasificación de la guía se basa en categorías fundamentadas en puntajes numéricos, criterios que son “aparentemente” objetivos (p.ej., “Mejores Universidades del País: Categoría Doctorado”).

Ahora, preguntémonos sobre los criterios que alguien podría usar para seleccionar una universidad. ¿Qué hay con la calidad de la educación? Tomando en cuenta todos los aspectos, una buena educación siempre debiera ser más deseable que una educación de calidad inferior. Pero es muy difícil definir qué entendemos por una buena educación, más aún pensar en medirla y clasificar los planteles siguiendo este criterio. De hecho, la calidad de la educación probablemente dependerá de muchas cosas bastante complejas de predecir. Podríamos sospechar que el mismo año, en cada una de las universidades del país, habrá alumnos beneficiándose de recibir lo que para ellos es una buena educación; asimismo, habrá otros expuestos a una pésima experiencia educativa. Pero, habiendo dicho eso, ¿cómo podemos traducir esa experiencia a cifras? Recordemos la discusión en el Capítulo 1 acerca de contabilizar lo incalculable.

^{**} Referencia a la muy vendida edición especial de la revista estadounidense *Sports Illustrated*, que una vez al año publica un número con fotografías de deportistas famosos en traje de baño (N. de la T.).

U.S. News resuelve el dilema de manera muy elegante, clasificando las universidades de acuerdo a criterios fácilmente cuantificables, solución bastante común para problemas como éste. Por ejemplo, los centros de estudios desean promover a los buenos profesores e investigadores, pero es muy difícil medir la calidad de su enseñanza o de sus conocimientos. Para tomar esas decisiones, la mayoría de los planteles se basa enormemente en criterios que generan cifras —puntajes de las evaluaciones docentes que hacen los propios alumnos, o número de publicaciones de un determinado profesor—, aun cuando todos los involucrados reconocen que esos indicadores están relacionados muy indirectamente con la calidad del cuerpo docente. En todo caso, al ser numéricas, estas medidas tan imperfectas permiten clasificar a los académicos, por ejemplo: el profesor A tiene mayor puntaje en su evaluación docente que el profesor B, y el profesor Y tiene más publicaciones que el profesor Z. El peligro está en que las cifras parecen objetivas; suele suceder que aquello que podemos expresar en números, se convierta en la medida decisiva. Esto es así porque la ausencia de números hace que los otros criterios parezcan demasiado arbitrarios.

¿Qué tipos de números podrá encontrar *U.S. News* para clasificar a las universidades? La revista utiliza una fórmula bastante complicada, pero, para nuestros objetivos, nos centraremos en tres clases de cifras incluidas en ella.

La primera se refiere a las calificaciones de los estudiantes aceptados para ingresar a la universidad. Ya que la revista busca indicadores que puedan traducirse en cifras disponibles para todas las sedes universitarias, aparecen en escena dos mediciones: el puntaje de las pruebas de postulación rendidas por los estudiantes (Prueba de Aptitud Académica [SAT, por su sigla en inglés]) y su desempeño durante la educación media. Con ello se supone que los planteles que aceptan a los alumnos más sobresalientes (más altos puntajes de postulación y mejores resultados en la enseñanza media) merecen estar más arriba en la clasificación. Nuevamente se está midiendo un concepto cualitativo

a partir de las únicas mediciones cuantitativas disponibles, lo cual, esencialmente, no las hace más objetivas.

El segundo conjunto de medidas se refiere al proceso mismo de admisión de las universidades: (1) los aspirantes postulan a un centro de estudios, (2) la universidad admite a un número restringido (les comunica que son bienvenidos a formar parte de su institución) y (3) algunos de esos postulantes escogidos aceptan ir a esa universidad. Desde luego que el número de estudiantes que decide matricularse es de vital importancia para el establecimiento, ya que éste planifica su presupuesto suponiendo que, para el próximo período, contará con un número determinado de alumnos con su matrícula pagada. Si se matriculan pocos, los ingresos serán menores de lo planificado y el plantel deberá reducir sus gastos; si son demasiados, es probable que falten profesores, dormitorios, equipamiento u otros, para acomodarlos a todos. Dado que muchos estudiantes son aceptados en más de una universidad, todas las instituciones saben que un porcentaje de los postulantes elegirá otro establecimiento. Eso significa que para asegurarse una cantidad de alumnos suficiente en el próximo período, las universidades deben aceptar más estudiantes de los que en realidad pueden acoger.

A su vez, estas etapas generan tres cifras distintas: la cantidad de postulantes, el número de alumnos admitidos y los que finalmente optan por un determinado establecimiento. Al dividir la segunda cifra por la primera, obtenemos la proporción de estudiantes aceptados (*tasa de admisión*). Al dividir la tercera cifra por la segunda, obtenemos la proporción de estudiantes admitidos que aceptan la invitación a ingresar (*tasa de rendimiento*).

Para clasificar las instituciones docentes, *U.S. News* utiliza las tasas de admisión y rendimiento. La revista supone que una universidad que acepta a un pequeño porcentaje de todos los que postulan (baja tasa de admisión) es muy selectiva, pues sólo elige a los mejores. Y si un gran número de los aceptados termina por elegir esa entidad (alta tasa

de rendimiento), pensamos que dicha institución es apetecible para los estudiantes.

Hubo una época en que las tasas de admisión y rendimiento eran cifras de uso interno y privado, utilizadas exclusivamente por los administradores de las universidades para planificar. Si se supone que la tasa de rendimiento de este año será similar a la del año anterior, tendremos una idea bastante razonable de cuántos postulantes es necesario aceptar para obtener la cantidad requerida de estudiantes de primer año. Pero ahora, gracias a *U.S. News* y el resto de la industria de guías de admisión a las universidades, estas cifras no sólo son públicas sino que se consideran como una base razonable para comparar la calidad de las instituciones de educación superior y forman parte de las fórmulas para calcular el ranking de calidad.

Ese énfasis lleva a las universidades a intentar ascender en la clasificación mejorando las cifras que *U.S. News* emplea en sus cálculos. Una manera de hacer esto es captando más postulaciones: independientemente de si en la actualidad recibe muchos candidatos de excelencia, aumentar la cifra total (aunque en definitiva se acepte a la misma cantidad de postulantes) permite declarar una tasa de admisión más baja, con lo cual ese plantel termina apareciendo como más selectivo. Asimismo, uno de los motivos de las universidades para preferir los programas de decisión temprana^{***} es que atraen postulaciones de estudiantes que tienen más probabilidades de recibir una oferta o propuesta de admisión. Al aumentar la cantidad de estos postulantes, se logra que crezca la tasa de rendimiento y que, por lo tanto, mejore la importancia del establecimiento en el ranking².

El tercer elemento en la fórmula utilizada por *U.S. News* para calcular el escalafón de las universidades es el más importante: evaluación

^{***} A través de estos programas, los estudiantes postulan 3 ó 4 meses antes de lo acostumbrado e indican su primera opción de centro de estudios, recibiendo los resultados con anterioridad a quienes lo hacen en los plazos usuales (N. de la T.).

² James Fallows, "The Early-Decision Racket", *Atlantic Monthly* 288 (September 2001): 37-52.

por parte de pares, que tiene una ponderación de 40% del puntaje de un establecimiento³. Pero ¿de adónde sacan a estos pares evaluadores? Supuestamente la revista envía formularios para votar a dos personas con cargo directivo en cada institución (decano, rector u otro encargado de admisión) a quienes se les solicita asignar puntajes a otros planteles del país. (En mi campus de la Universidad de Delaware, estas papeletas son enviadas al director de relaciones públicas y al encargado de admisión). Las interrogantes surgen de inmediato: ¿qué es lo que califica a esas personas para evaluar la calidad de otras universidades?, ¿por qué no enviar los votos a quienes están involucrados más directamente en el proceso educativo con los alumnos?, ¿a partir de qué base podrían esos individuos evaluar instituciones que posiblemente ni siquiera conocen? Hoy en día, las instituciones más sagaces dirigen sus estrategias de marketing directamente hacia estos votantes. Durante las semanas previas a la llegada de las papeletas de votación, las personas que deben emitir su opinión comienzan a recibir publicidad (varias universidades envían libros muy bien impresos en papel satinado en los que ensalzan sus virtudes). Se supone que una campaña eficaz redundará en puntajes más altos por parte de los calificadores, lo que a su vez incidirá en un mejor lugar en la clasificación.

Dado que estos tres elementos son fácilmente modificables, las universidades tienden a enfocarse exclusivamente en ellos. Sin embargo, la fórmula de *U.S. News* también incorpora otros factores, como la cantidad de miembros del cuerpo docente, el gasto incurrido por estudiante y la tasa de graduaciones. Estos números son difíciles de cambiar y sería demasiado caro y complejo hacerlo año a año para evidenciar un cambio sustancial. Por el contrario, estimular una mayor cantidad de admisiones y publicitar las virtudes de las sedes universitarias ante quienes deben votar, son tácticas relativamente económicas

³ La fórmula utilizada en 2002 aparece explicada en Robert J. Morse y Samuel M. Flanigan, "The Rankings", *U.S. News & World Report, America's Best Colleges, 2003 Edition*, September 2002, pp. 79-81.

que podrían generar cambios rápidos y favorables en los puntajes.

Desde luego, ninguna de estas manipulaciones tiene algo que ver con la calidad de la educación ofrecida por una universidad. Sin embargo, son importantes, pues la fluctuación anual de la ubicación de una universidad en el número especial dedicado a la caza de postulantes puede influir en la elección de plantel por parte de los estudiantes. Y esto efectivamente es así, aun cuando es mucho más probable que las fluctuaciones en el ranking reflejen cambios en cómo la oficina de admisión maneja sus asuntos o cuán bien la institución se promociona ante quienes deben votar, y no reflejen lo que realmente sucede dentro de las salas de clases.

Este ejemplo muestra cómo funcionan los números mágicos. Ellos ayudan a disminuir la incertidumbre. En este caso, los estudiantes potenciales y sus familias, que quieren tomar la decisión correcta, se enfrentan a un desconcertante abanico de posibilidades. La clasificación de *U.S. News* parece ofrecer una base objetiva para resolver: la edición especial traduce la calidad de la educación en una fórmula compuesta por elementos cuantificables que produce puntajes que permiten clasificar a las universidades. Tras bambalinas, algunas instituciones pueden estar intentando mejorar su puesto en la clasificación, no necesariamente por la vía de mejorar la calidad de la educación en sus sedes, sino más bien ofreciendo más vacantes o solicitando los votos a los encargados de relaciones públicas de otras universidades. No obstante, esto permanece oculto. Y quienes quieran confiar en la edición especial, podrán seguir creyendo que sus decisiones se basan en estadísticas bien hechas y aparentemente consistentes.

CÓMO JUGAR CON NÚMEROS MÁGICOS DENTRO DE LAS ORGANIZACIONES

Es muy probable que este escalafón de universidades surta su efecto mágico sobre algunos estudiantes y sus familias. Hay muchos futuros

alumnos que, sin lugar a dudas, desconocen las guías, e incluso entre quienes las consultan, son muy pocos los que tomarán decisiones basándose únicamente en ellas. La importancia asignada a las estadísticas de estas guías —el grado de poder mágico que ejercen— varía entre las personas. En contraposición a ello, otras cifras tienen mayor influencia; pueden afectar a mucha gente, en una forma más o menos simultánea, dentro de algunas organizaciones e instituciones.

En tanto los individuos buscan disminuir su incertidumbre por medio de las cifras, muchas grandes organizaciones e instituciones dependen de las estadísticas para manejar su funcionamiento cotidiano⁴. Necesitan de las cifras para evaluar cómo van las cosas. Las empresas deben calcular sus costos y ventas, sus ganancias y pérdidas; en cambio, el gobierno y otros organismos sin fines de lucro tienen sus propios presupuestos e itinerarios. Una organización debe generar informes de avance, reportes de eficiencia, evaluaciones y todo tipo de documentos que incluyen cálculos complejos. Cuanto más grande, más cuesta llevar la cuenta de todo lo que sucede; y para resumir y ordenar la complejidad y evaluar su propio desempeño y el de los otros, más dependen de las cifras sus administradores y miembros. Estos números condensan la realidad en cantidades que a simple vista parecen muy sencillas, con lo cual se convierten en un fundamento para la toma de decisiones. Aun así, el proceso subyacente no es tan diferente al de los confundidos estudiantes de enseñanza media que recurren a las guías ya mencionadas: la ambigüedad y la incertidumbre estimulan a las organizaciones a usar las estadísticas para simplificar las cosas. Y, según el grado en que estas cifras se transforman en elementos claves para comprender e interpretar la realidad de lo que pasa al interior de la organización, los números van adquiriendo características mágicas.

⁴ Un clásico sobre estos temas continúa siendo Harold L. Wilensky, *Organizational Intelligence: Knowledge and Policy in Government and Industry* (New York: Basic Books, 1967).

Siempre que las cifras son relevantes, siempre que las personas las toman en serio y utilizan como base para sus decisiones y acciones futuras, es que alguien tiene algún interés creado en ellas. Quienes deciden, basándose en estadísticas suministradas por otros, necesitan creer que esos números son correctos, exactos y válidos, y se asegurarán de contar con buena información. Por su parte, quienes se ven afectados por esas decisiones, preferirán cifras que impliquen resultados positivos; las estadísticas que fomentan que el jefe aumente el presupuesto son preferibles a aquellas que podrían influir en su decisión de despedirnos. A las personas les importan las cifras, y mientras más mágicas sean —mientras más significativas se les considere como fundamento para la toma de decisiones—, más importancia les darán. Además, por el hecho de que todas fueron generadas por alguien que contabilizó algo, es casi seguro que habrá esfuerzos para influir en la producción de números mágicos. Ya vimos un ejemplo de esto en los intentos de algunas universidades por ascender en el escalafón. En la mayoría de las organizaciones ocurren cosas parecidas.

En general, las cifras relativas a una organización son de dos tipos: unas para uso interno y otras para fines externos. En el primer caso, los directores o administradores de determinadas áreas deben entregar a sus superiores informes con números relacionados a gastos y productividad, quienes los usarán para decidir cuáles secciones requieren de mayor apoyo o supervisión. Inevitablemente, la complejidad —todos los factores que influyen en lo que sucede al interior de la organización— se resumirá en unos pocos números. Pero, ¿qué es lo que miden?, ¿cuándo y cómo se hace? Las respuestas a estas preguntas reflejan una elección de lo que se considera importante en la organización. Si un jefe solicita un informe a sus subordinados, supone que esos números pueden entregar una idea sobre lo que es relevante.

Dentro de una organización, la exigencia de dar a conocer estadísticas a otros, prepara el tinglado para los “juegos de cifras” burocráticos. Obviamente, un número mágico que funciona a favor de uno, es

una buena cifra; los subordinados tienen todas las razones del mundo para cooperar gustosos en la creación de esas estadísticas. Pero si, por el contrario, las cifras indican que una determinada unidad está con problemas, su impacto puede ser minimizado. Para dar la mejor impresión posible, un subalterno astuto podría manipular los números de un informe e incluso sugerir que su unidad está haciendo un buen trabajo, que es más eficiente y productiva y que, por ende, merece un mayor reconocimiento que las demás secciones. Otra posibilidad es que, cuando las cifras no pueden arreglarse, un subordinado con suficiente experiencia alegue que son imperfectas, que no logran calcular lo que realmente importa ni reconocer lo que la unidad hace bien; por lo tanto, que esa información no significa nada, es irrelevante y se deben hacer otras mediciones.

Por su parte, los supervisores perspicaces estarán conscientes del interés de los subalternos por dar la mejor impresión y tratarán de asegurarse de que las cifras que reciben son correctas. En el caso de que éstas sean poco confiables, pueden pedir informes adicionales que utilicen otros indicadores, o podrán insistir en el uso de métodos específicos y estandarizados para medir e informar. Las nuevas solicitudes invitan a los subalternos a pensar en cómo sacar ventaja de las nuevas cifras. Si los supervisores no logran ejercer control sobre una circunstancia como esa, la organización puede verse invadida por los números falsos. Por ejemplo, el caso de las estadísticas de producción agrícola de la ex Unión Soviética —generadas por subordinados más bien temerosos de ser multados por las malas cosechas, que preocupados de que sus informes falsos fueran descubiertos— destaca como un monumento a este tipo de engaño interno⁵.

En las organizaciones y empresas hay otras cifras que tienen un público externo; son números examinados —y tratados como

⁵ El sistema soviético ha sido analizado en detalle; véase, por ejemplo, Stefan Hedlund, *Crisis in Soviet Agriculture* (New York: St. Martin's, 1984).

importantes— por personas que no pertenecen a la organización. Por ejemplo, los inversionistas utilizan los resultados de los informes financieros de las empresas para decidir si una empresa es o no una buena inversión. Si una organización está consciente de que sus resultados serán examinados por gente externa, sus miembros pueden modificarlos para dar la impresión deseada. De nuevo, se vuelcan a los “juegos de cifras”, sólo que ahora no todos los jugadores forman parte de la organización. Por esta razón, algunos críticos sostienen que dado que los inversionistas actuales prestan especial atención a los informes trimestrales de ganancias de las empresas —“la Cifra”—, esto favorece políticas comerciales y prácticas de contabilidad que podrían generar utilidades que equivalgan a, o superen, las expectativas del mercado, aun cuando distintas operaciones pudieran coincidir con el interés a largo plazo de la empresa (y de los inversionistas)⁶. Igualmente, los departamentos de policía a veces clasifican los delitos de modo que las tasas delictuales de sus ciudades aparezcan disminuidas, implicando con ello que están haciendo bien su trabajo⁷. Siempre que haya agentes externos que utilizan cifras mágicas para evaluar el desempeño de una organización, ésta tiene la posibilidad de influir en los resultados.

A su vez, al saber que una organización puede manipular sus estadísticas, quienes no pertenecen a ella pueden tratar de controlarlas. Podrían insistir en que la información sea entregada de una determinada manera. Por ejemplo, cuando el FBI solicita a los departamentos de policía que completen los formularios de los *Informes Estandarizados de Delincuencia* que sirven de base para calcular las tasas delictuales, entrega instrucciones detalladas acerca de qué contabilizar y cómo hacerlo. Asimismo, la U.S. Securities and Exchange Commission define un formato general para dar a conocer las finanzas de las empresas.

⁶ Alex Berenson, *The Number: How the Drive for Quarterly Earnings Corrupted Wall Street and Corporate America* (New York: Random House, 2003).

⁷ Mark Fazlollah, Michael Matz y Craig R. McCoy, “How to Cut City’s Crime Rate: Don’t Report It”, *Philadelphia Inquirer*, November 1, 1998, p. A1.

Estos esfuerzos apuntan a que los informes de diversas organizaciones sean comparables. Los agentes externos a las organizaciones pueden incluso fijar multas —y aplicarlas— a quienes entreguen datos incorrectos. Esas medidas pueden disminuir la cantidad de informes engañosos, pero, como nos lo recuerda el escándalo de Enron, no pueden asegurarnos que los números sean correctos⁸.

El punto es que las organizaciones necesitan de las estadísticas para operar, para tomar decisiones internas y resumir sus actividades a quienes no pertenecen a ellas. Mientras mayor sea el grado en que las personas, ya sea al interior o fuera de una organización, toman en serio esas cifras y las utilizan como base para sus decisiones —es decir, mientras más mágicas sean—, más interés tendrán sus miembros en hacer calzar las estadísticas con sus objetivos. Sería ingenuo creer que las estadísticas dadas a conocer por las organizaciones reflejan toda la verdad como si fueran un espejo de la realidad.

Cuando las organizaciones utilizan estadísticas, debemos reconocer que existen ventajas y desventajas asociadas a ello. Las cifras de las organizaciones condensan la complejidad, lo que siempre trae aparejados costos y beneficios: nos permiten resumir, aclarar, captar la idea general; pero, ya que los individuos deciden qué contabilizar y cómo hacerlo, inevitablemente los resúmenes suelen simplificar demasiado las cosas. Mientras más relevantes sean los números (más mágicos), es más probable que las personas ponderen cuidadosamente las opciones y destinen trabajo para dar a conocer su punto de vista, y menor será la confianza depositada en los números como reflejos directos de la realidad.

Esta es la gran paradoja de las cifras mágicas: las necesitamos y precisamos confiar en ellas; sin embargo, mientras mayor sea nuestra necesidad, es más probable que los datos sean distorsionados y más

⁸ El escándalo de Enron amerita una extensa bibliografía. Para una introducción, véase Loren Fox, *Enron: The Rise and Fall* (Hoboken, NJ: Wiley, 2003).

cuidado tendremos al examinarlas. *Antes de confiar a ciegas en las estadísticas, debemos preguntarnos quién contó qué cosa, cómo y por qué lo hizo* —porque, como muestra el siguiente ejemplo, cuando los números mágicos se convierten en un tema muy atractivo, las probabilidades de generar confusión son muy grandes.

COLEGIOS EN TELA DE JUICIO

En las últimas décadas del siglo XX, el grado de inquietud con respecto a la calidad de la educación estadounidense llegó a niveles notables. Esto podría parecer extraño, pues el promedio de escolaridad en EEUU había crecido de manera muy significativa durante el siglo pasado. En 1900, sólo el 6% de los estadounidenses de 17 años se graduaba de la enseñanza media⁹. Al término del siglo, la mayoría de los estadounidenses continuaban su educación después de la enseñanza media, en tanto que alrededor de un tercio de los pertenecientes a las cohortes más nuevas terminaban su licenciatura. Hoy en día, EEUU tiene uno de los mayores porcentajes de ciudadanos con estudios superiores en el mundo¹⁰.

Sin embargo, eran otras las estadísticas que provocaban los problemas. Por medio de diversas investigaciones se supo que, en pruebas comparativas, los estudiantes estadounidenses obtenían puntajes menores a los de otros países, particularmente en ciencias y matemática, temas en los que las diferencias culturales debieran tener un impacto mínimo (ya que las respuestas posibles eran sólo dos: correcto o incorrecto). Esta comparación mostraba que los estudiantes estadounidenses no estaban aprendiendo suficiente, o tan bien, como los alumnos de otras partes. Además, entre mediados de la década de los 60 y 1980,

⁹ Bureau of the Census, *Historical Statistics of the United States* (Washington, DC, 1975), p. 55.

¹⁰ National Center for Education Statistics, *Digest of Education Statistics*, 2001, Table 411; postado en <http://nces.ed.gov/programs/digest/do1/dt411.asp>

cayeron los puntajes del SAT —la principal prueba de selección de las universidades—, lo que hizo pensar que el desempeño de los estudiantes en EEUU estaba empeorando. (Desde 1980 en adelante, los puntajes de matemática del SAT han mejorado significativamente, aunque los de lenguaje se mantienen bajos)¹¹.

Estas evidencias hicieron surgir dudas respecto a la calidad de la educación y los logros de los estudiantes en EEUU, al mismo tiempo que los críticos se preguntaban lo que podría significar esto para el futuro del país. Quizás hubo un tiempo en que era posible abandonar la escuela y, a pesar de ello, llevar una vida decente. Sin embargo, el mercado laboral actual exige más educación: cada vez se ofrecen menos empleos para individuos no calificados. Hoy en día, el desertor escolar estaría arriesgándose a una vida de pobreza y marginalidad. Incluso un país que deja de maximizar la educación de sus ciudadanos, se arriesga a quedar postergado ante otras naciones que educan mejor a sus jóvenes. Estas críticas hacen sonar las alarmas en relación con la falsa creencia de que basta con tener mayores tasas de alumnos que hayan completado la enseñanza media; que un mayor número de alumnos, de capacidades menores, se gradúe es sencillamente una demostración de que los colegios han descuidado los estándares académicos. Quienes plantean esto tienen una visión nostálgica del pasado educacional: antes, los alumnos se esforzaban, estudiaban mucho y aprendían; eran dedicados y decididos. ¡Pero los chicos de hoy! No saben nada, no les importa nada, no leen, sólo ven televisión; y la música que escuchan —si es que podemos llamarla música... E, inevitablemente, los críticos empezaron a sonar como sus propios padres.

¹¹ Se han entregado distintas interpretaciones sobre las tendencias de los puntajes del SAT. Dos posturas sutilmente diferentes pueden verse en Scott Menard, "Going Down, Going Up: Explaining the Turnaround in SAT Scores", *Youth and Society* 20 (1988): 3-28; y Charles Murray y R.J. Herrnstein, "What's Really Behind the SAT-Score Decline?", *The Public Interest* 106 (1992): 32-56.

La nostalgia es un lente defectuoso, si de observar los cambios se trata. Cada generación de críticos educacionales, como los que señalamos, tienden a caracterizarse por haber sido buenos alumnos en el colegio, al menos lo suficientemente exitosos como para adquirir el rótulo de críticos. Ellos recuerdan su propio desempeño escolar y el de sus amigos como bastante aceptable y suelen olvidar a sus compañeros a quienes no les fue tan bien o que debieron abandonar la escuela por diversos motivos.

El contrastar los recuerdos de los críticos con los estudiantes actuales no es necesariamente una comparación muy apropiada. Incluso los cambios en los puntajes de las pruebas estandarizadas pueden ser complicados de evaluar. Suponiendo que, en general, los alumnos más capaces permanecen más tiempo dentro del sistema educacional, entonces, al aumentar la proporción de jóvenes que siguen en la enseñanza media o la universidad, las habilidades promedio de los que se gradúan podrían disminuir, dado que hay una mayor cantidad de estudiantes de menor desempeño que siguen educándose. Por eso es que medir logros educativos a lo largo del tiempo podría estar comparando puntajes de poblaciones estudiantiles bastante diferentes.

Aun así, opinar sobre el sistema educacional parece ser algo que atrae a todo tipo de críticos. Los conservadores culpan del bajo desempeño de las escuelas a las instituciones que se han apartado del currículo académico tradicional y la disciplina estricta, y claman por recuperar esos principios fundamentales. A su vez, los progresistas pueden argumentar que la falla radica en que los colegios no cumplen su objetivo con los estudiantes menos aventajados y que el currículo debe modificarse para educar a los más vulnerables.

Sea cual sea la agenda de los críticos, la mayoría coincide en que hay que hacer algo, si bien es probable que no coincidan en qué es lo que deba hacerse. De ahí que una sociedad que predica el perfeccionismo —“Que ningún niño se quede atrás”— es más propensa a

considerar que las escuelas en realidad no están alcanzando las metas propuestas y que las críticas están bien fundadas.

Recientemente, y como un medio de responsabilizar a los colegios, esta angustia respecto a la educación ha llevado a la adopción extendida de pruebas escolares estandarizadas. Tanto los estados como el gobierno federal exigen que todos los alumnos de las escuelas públicas rindan pruebas estandarizadas y que esos resultados —en especial, los promedios de los diferentes colegios— sean de conocimiento público.

Estos exámenes tienen varias consecuencias. Los estudiantes que obtengan malos resultados deberán asistir a clases especiales para nivelarse (escuelas de verano); en algunos distritos, los estudiantes que terminen los cursos obligatorios pero que no logren un puntaje mínimo en esas pruebas, podrán recibir un diploma de enseñanza media, pero de menor categoría que sus compañeros con mejor rendimiento. Los profesores también padecen los efectos. Algunos defienden la idea de aumentar los salarios de los docentes cuyos alumnos obtengan mejores resultados en las pruebas. Junto con ello, los colegios son igualmente señalados. Los periódicos informan de los resultados, colegio por colegio, dando a entender que unos rinden mejor que otros; incluso, en algunos estados, las escuelas que muestran avances significativos en los puntajes en las pruebas, reciben bonificaciones especiales. Esto tiene implicancias fuera del ámbito educacional: los agentes inmobiliarios han descubierto que es buen negocio situarse en las cercanías de un colegio de alto rendimiento en las pruebas. En resumen, estos puntajes se han vuelto un ejemplo vívido de cifras mágicas.

Para que estas evaluaciones tengan consecuencias relevantes para los alumnos, los profesores y las escuelas, debemos suponer algunas cosas en relación con lo que las pruebas efectivamente miden. Lo más obvio es que tengamos la impresión de que éstas entregan una medida válida del aprendizaje y las habilidades de los estudiantes; que aquellos que obtienen un buen puntaje en matemática, de hecho han aprendido más matemática. Pero además tenemos que reconocer que

la enseñanza en aula es la clave de este aprendizaje. A primera vista, esto estaría fuera de discusión: “¿A qué van los niños al colegio, si no es a aprender? ¿Y no se supone que enseñar es lo que deben hacer los colegios?”.

Ahora pongamos atención al conocido rol de la pertenencia a una clase social definida en el desempeño escolar: en general, a los estudiantes de grupos de clase media-alta (en su mayoría, blancos) tiende a irles mejor que a los de clase baja (en su mayoría, negros o latinos). Las causas de esta tendencia es materia de acalorados debates. Muchas de las explicaciones enfatizan las diferencias que hay entre los estudiantes (p.ej., que la inteligencia está determinada en parte por la genética), la desigualdad del medio social de los alumnos (p.ej., si acaso los familiares, amigos y conocidos de los estudiantes aprecian y apoyan la educación) y las disparidades en la enseñanza misma (p.ej., si los niños de clase media-alta van a colegios con mejores profesores, en cursos reducidos, con una infraestructura más agradable y otras ventajas).

Las distintas explicaciones tendrán implicancias importantes para las políticas de evaluación. Por ello, si creemos que lo que sucede en el colegio es el factor determinante de cuánto aprenden los niños, entonces los puntajes en las pruebas serán un buen indicador del desempeño de los establecimientos educacionales. En cambio, si partimos de la base de que las circunstancias sociales de los estudiantes tienen efectos importantes en el aprendizaje, entonces los puntajes de las pruebas, en último término, serán una medición de la pertenencia de los alumnos a una determinada clase social.

Esta es una de las razones por las cuales la discusión sobre políticas de evaluación se ha vuelto tan cáustica. Los que defienden las pruebas estandarizadas sostienen que los colegios y los docentes deberían hacer mejor su trabajo y, como una fórmula para medir su desempeño, hay que hacerlos responsables de los puntajes obtenidos por los alumnos en esas pruebas. La suposición es que si los profesores hicieran bien

el trabajo por el cual se les remunera, sus estudiantes aprobarían los exámenes. Si los alumnos de algunos colegios rinden mal en las pruebas, se les podrá exigir asistir a las escuelas de verano, no van a calificar para recibir sus diplomas, sus profesores recibirán un menor aumento salarial, etc. En resumen, los puntajes tendrán serias consecuencias.

En la medida que esas consecuencias se tornan más evidentes, las críticas a este tipo de evaluaciones se han vuelto más elocuentes. No es de extrañar que muchos profesores y directores de colegios se opongan a los sistemas de evaluación que castigan a los educadores si sus alumnos tienen un mal rendimiento. Los sindicatos de profesores arguyen que los docentes están siendo responsabilizados por los factores sociales de sus alumnos: “Desde luego, les irá mejor a los niños que provienen de hogares de clase media-alta, con gran cantidad de libros, con padres educados que valoran la importancia de la enseñanza, y que disfrutan de otras ventajas. Naturalmente, es esperable que los niños de clase baja, que carecen de esas ventajas, tengan mayores dificultades de aprendizaje, y no podemos culpar a los profesores por cosas que ellos no pueden controlar”.

Otro grupo de críticos plantean que incluso los buenos resultados pueden ser ilusorios, porque las pruebas de alta ponderación, aquellas cuyos resultados tienen consecuencias directas para quienes las rinden, harán que los colegios terminen “enseñando para la prueba” o ejerciten a los niños en el tipo de preguntas que aparecen en los exámenes, dejando a un lado otras modalidades de aprendizaje, quizás más importantes. Esas críticas son precisamente el tipo de reacción que podemos esperar de quienes son evaluados por su desempeño, cuando las grandes instituciones adoptan números mágicos: los profesores (que están siendo evaluados por los puntajes de sus alumnos) dicen que es imposible que las pruebas puedan evaluar con precisión su desempeño, en tanto que los que no son docentes sospechan que quienes lo son pueden alterar la instrucción para maximizar los puntajes, disminuyendo la calidad de la enseñanza. Incluso hay informes relativos a

profesores que han ayudado a algunos estudiantes a mejorar sus puntajes en forma fraudulenta¹².

Además, la manera en que los puntajes son utilizados también presenta problemas. Por ejemplo, algunos estados premian a los colegios que exhiben una notoria mejoría en los puntajes obtenidos de un año a otro, una práctica probablemente equivocada. Se ha demostrado, por medio de investigaciones, que los colegios que presentan marcadas variaciones positivas año a año, tienden a tener menos estudiantes¹³, lo que sugiere que las cifras consideradas para medir la excelencia educativa pueden ser un mero artefacto estadístico. Imaginemos una pequeña escuela básica con un solo curso por nivel. Los puntajes año a año en esa escuelita son relativamente volátiles: si en tercero básico hay un par de alumnos sobresalientes, ese curso podrá obtener puntajes significativamente mayores que los del año anterior, compuesto por muchos alumnos de rendimiento regular, aun cuando el docente haya enseñado lo mismo en ambos años. Por lo tanto, ese tipo de variación año a año es menos probable en un colegio grande compuesto, por ejemplo, de cinco cursos de tercero básico; ahí, un par de alumnos brillantes tendrá menos impacto en el desempeño total del colegio, y los puntajes de las pruebas tenderán a ser bastante estables. Incluso en el caso de que los alumnos hayan respondido sus pruebas aleatoriamente, los colegios pequeños tienen mejores posibilidades que los más grandes de mostrar oscilaciones de aumentos y disminuciones en los puntajes año a año. Desde luego que, cuando los puntajes en las evaluaciones son

¹² Existe una gran bibliografía sobre las recientes reformas educacionales centradas en las pruebas. Por ejemplo, se pueden encontrar interpretaciones favorables en Williamson M. Evers y Herbert J. Walberg, eds., *School Accountability* (Stanford, CA: Hoover Institution Press, 2002). Para un análisis más escéptico, véase Laura S. Hamilton, Brian M. Stecher y Stephen P. Klein, *Making Sense of Test-Based Accountability in Education* (Santa Monica, CA: Rand, 2002). Para evidencias relativas a las trampas, véase Brian A. Jacob y Steven D. Leavitt, "To Catch a Cheat", *Education Next* 4 (Winter 2004): 69-75.

¹³ Thomas J. Kane, Douglas O. Staiger y Jeffrey Geppert, "Randomly Accountable", *Education Next* 2 (Spring 2002): 57-61.

considerados números mágicos —la medida definitiva de cuán bien se desempeña un colegio—, la posibilidad de que el azar tenga relación con el resultado de los puntajes, sus alumnos y profesores, desaparece de la discusión sobre políticas educacionales.

La evaluación educacional, con su promesa de tener bajo control a los colegios, está de moda, una tendencia que, por desgracia, promete quedarse con nosotros por un tiempo (al menos hasta que las limitaciones de estas políticas se hagan más evidentes). Ello es un claro ejemplo del poder de los números mágicos que debiera servir como precaución a todos los aspirantes a reformadores.

PRÁCTICAS POLICIALES BASADAS EN EL PERFIL RACIAL****

Otro tema polémico en los últimos años han sido las prácticas policiales discriminatorias basadas en el denominado *perfil racial*. Debemos comenzar diciendo que, según quien lo utilice, el término se refiere a cosas distintas. Me restringiré a aquellos reclamos referidos a prácticas de la policía en que los conductores son conminados a detenerse en la vía pública, en parte, debido a su raza. Muchos oficiales de policía plantean, y los tribunales así lo han confirmado, que en ocasiones la raza puede ser considerada una característica importante —aunque no la única— para detener a un vehículo. Por ejemplo, supongamos que la policía tenga motivos para creer que algunos sujetos que trabajan para una red criminal afroamericana están utilizando una ruta en particular para transportar drogas en cierto tipo de vehículos. Bajo esas circunstancias, es muy posible que una patrulla proceda a detener a un auto sospechoso, en parte porque su conductor es negro o de apariencia extraña. Así describen esta práctica discriminatoria los defensores del perfil racial.

**** Según Amnistía Internacional, el término “perfil racial” se refiere al empleo de criterios raciales en la identificación de probables sospechosos de un crimen. Véase www.amnestyusa.org/spanish/eventos/audiencias.html (N. de la T.).

En cambio, quienes la critican dicen que los conductores afroamericanos son los más afectados por esta política, simplemente por ser negros y conducir un vehículo motorizado (concepto que en inglés se denomina “*driving while black*”). Muchos afroamericanos piensan que ellos son más propensos que los conductores blancos a ser conminados a detenerse, ya que la policía sospecha que los negros tienen una mayor tendencia a involucrarse en actividades delictuales. Hay casos en que la raza del conductor puede ser el único motivo para detener a un vehículo¹⁴. Desde este punto de vista, el uso del perfil racial es una práctica racista. Hay un grupo de críticos que plantean que la raza o el aspecto jamás debería ser un motivo para dar la orden de detención a un conductor.

Tan pronto como el perfil racial se convirtió en un tema de alta visibilidad política, la opinión pública comenzó a solicitar estadísticas que pudieran determinar, de una vez por todas, la existencia y extensión de esa práctica. Es decir, estaban clamando por un número mágico: una medición de la discriminación por perfil racial.

La recolección de estadísticas se ha convertido en una medida de un pacto intermedio bastante popular en la política contemporánea; por ejemplo, para medir el alcance de los crímenes causados por odio racial, la primera ley federal relativa a este tipo de legislación exigió que el FBI los contabilizara. Esos pactos implican que las estadísticas pueden resolver polémicas mágicamente. Las estadísticas se consideran mediciones objetivas, que entregan un retrato imparcial de las prácticas policiales, de los crímenes ocasionados por odio racial, etc. Es difícil oponerse a recolectar esas estadísticas, porque se asume que recoger información no es más que establecer los hechos. Además, es muy probable que los participantes en un debate crean que las estadísticas van a apoyar su postura: en el caso del perfil racial, la policía

¹⁴ Richard J. Lundman y Robert L. Kaufman, “Driving While Black”, *Criminology* 41 (2003): 195-220.

podría anticipar que las estadísticas revelarán que se comportan de manera responsable, mientras que sus críticos pueden anticipar que las prácticas discriminatorias serán expuestas por medio de las cifras.

El problema es que medir el uso de la práctica de discriminación por perfil racial, con toda seguridad será mucho más complicado de lo que creemos¹⁵. Los estudios más sencillos sobre el tema comparan la raza de los automovilistas conminados a detenerse con la composición racial de los residentes del área. Supongamos que el 20% de los conductores detenidos por policías de tránsito de una localidad sean negros. Antes de que podamos interpretar ese hallazgo, necesitamos saber algo sobre la población que conduce por la carretera. ¿Será que el 10% de los automovilistas del área son de raza negra (lo que sugeriría que los afroamericanos son obligados a detenerse mucho más a menudo de lo que se supone)? ¿Será que el 20% de los conductores son de raza negra (lo que insinúa que la proporción de afroamericanos que son forzados a detenerse se acerca a la cifra esperada)? ¿O es que el porcentaje de personas de raza negra que conducen se acerca al 30% (lo que invita a creer que los afroamericanos son detenidos menos que otros conductores)? Sin embargo, hacer esas comparaciones es más simple en la teoría que en la práctica.

La clave está en identificar la población de conductores que servirá de base para la comparación. Un criminólogo llama a esto “la búsqueda del denominador”¹⁶. La base de comparación más sencilla es la composición racial de una población determinada (disponible en las

¹⁵ Recién ahora comienzan a aparecer estudios sobre el tema. Véase, por ejemplo, Robin Shepard Engel, Jennifer M. Cainon y Thomas J. Bernard, “Theory and Racial Profiling”, *Justice Quarterly* 19 (2002): 249-273; Albert J. Meehan y Michael C. Ponder, “Race and Place: The Ecology of Racial Profiling African American Motorists”, *Justice Quarterly* 19 (2002): 399-430; Michael R. Smith y Geoffrey P. Alpert, “Searching for Direction: Courts, Social Science, and the Adjudication of Racial Profiling Claims”, *Justice Quarterly* 19 (2002): 673-703; y Samuel Walker, “Searching for the Denominator”, *Justice Research and Policy* 3 (2001): 63-95.

¹⁶ Walker, “Searching for the Denominator”, 63-95.

estadísticas del censo). Pero si consideramos que no todos en la ciudad manejan un vehículo, es probable que debamos ajustar nuestro cálculo poblacional intentando definir la composición racial de los conductores de la ciudad que cuenten con licencia. Agreguemos que no todos los automovilistas con papeles en regla manejan la misma cantidad de kilómetros, y podríamos suponer, además, que mientras más uno maneja, mayor es el riesgo de ser detenido por la policía.

A lo anterior podemos sumar que si algunas vías sólo son transitadas por conductores locales, es probable que ellos reflejen la población de la comunidad (lo que no sucederá en las carreteras interestatales, donde hay automovilistas de todas partes). Asimismo, puede ocurrir que la población que conduce cambie entre el día y la noche, de lunes a viernes y el fin de semana, y no debiera sorprendernos que las decisiones policiales de detener a quienes manejan dependan de la hora del día. Por ejemplo, es posible que un policía detenga a alguien que maneja descuidadamente avanzada la noche, suponiendo que, a esas horas, quien lo hace se encuentra en estado de embriaguez. En resumen, lograr que las autoridades se pongan de acuerdo para recolectar información acerca de la distribución racial de los conductores que son conminados a detenerse por la policía, es sólo una parte del problema. Además, hay que acordar una base de comparación.

Si queremos medir la raza de los infractores del tránsito, se necesita un enfoque algo más sofisticado. En uno de los primeros estudios respecto al tema, un equipo de investigadores hizo el siguiente experimento. Un grupo condujo al límite de velocidad por un tramo de la carretera interestatal del estado de Maryland, tomando nota de todos los vehículos que los sobrepasaron (que iban sobre la velocidad permitida, así es que, teóricamente, cumplían con los requisitos para ser detenidos por la policía estatal que patrullaba la vía). Según el estudio, el 18% de quienes conducían a exceso de velocidad eran negros, en tanto el 28% de los automovilistas detenidos por la policía estatal de Maryland eran de raza negra.

Los resultados del estudio eran sugerentes, aunque no es difícil identificar sus falencias. La experiencia cotidiana nos indica que un porcentaje importante de conductores sobrepasa el límite de velocidad y que, en general, la policía no hará detenerse a un conductor que exceda ligeramente ese límite, digamos por unos 5 km/h. En consecuencia, un estudio que considere que todos los automovilistas que sobrepasan la velocidad límite son candidatos a recibir la orden de detenerse, no ha identificado a la población relevante. Si, por ejemplo, los blancos suelen excederse por poco en la velocidad límite al conducir y los negros tienden a manejar más rápido que eso, entonces la proporción de conductores afroamericanos que llaman la atención de la policía por exceso de velocidad será mayor que el porcentaje de los que apenas sobrepasan el límite.

Hay un sinfín de complejidades que saltan a la vista. Puede que exceder los límites de velocidad no sea lo que haga que los policías detengan a los vehículos. Quizás están tan interesados —o más— en los que conducen temerariamente. A lo mejor, se fijan en los automóviles más antiguos, los que de manera más evidente podrían estar defectuosos. Si los afroamericanos conducen coches viejos, o lo hacen con imprudencia, esto podría justificar que los obliguen a detenerse más a menudo. O quizás existan diferencias etarias entre los conductores de distintas razas. Sabemos, por ejemplo, que los jóvenes suelen verse involucrados en un mayor número de accidentes. Si la población de automovilistas negros está compuesta por un porcentaje más alto de jóvenes, podría esperarse que ése sea el factor que llama desproporcionadamente la atención de las autoridades.

El punto significativo es que ninguna de estas explicaciones es necesariamente verdadera. Más bien, hace referencia al hecho de que utilizar la raza de los conductores que sobrepasan al vehículo en que va el investigador que se desplaza en el límite de la velocidad, es un medio inadecuado de identificar a quienes la policía podría conminar a detenerse. Los hallazgos estadísticos del estudio de Maryland son

sugereentes, pero no constituyen pruebas indesmentibles respecto al alcance del uso del perfil racial como práctica discriminatoria.

En resumen, medir el perfil racial no es tan sencillo como parece. Como sea que se recolecte información al respecto, habrá quienes argumenten que las estadísticas resultantes no son válidas. El llamado a obtener datos sobre el tema parece estar fundamentado en la creencia de que estas estadísticas generalmente serán aceptadas como números mágicos, pero es poco probable que todo el mundo le asigne tal autoridad a estos números.

Sin embargo, dichos problemas no significan que el recoger información no valga la pena. Es muy probable que esa información efectivamente interese a los departamentos de policía que recaban datos acerca de la raza de los conductores que son obligados a detenerse. Si, por ejemplo, se descubriera que algunos policías —y más aún, uno en particular— obligan a detenerse a un porcentaje importante de conductores afroamericanos con mayor frecuencia que otros oficiales que patrullan las mismas vías, eso produciría un cuestionamiento legítimo a ese o esos oficiales. Habría que exigirles que explicaran las diferencias entre su patrón de detenciones y el de sus colegas. Es posible que el solo hecho de saber que esa información se está recabando y que el desempeño individual está siendo supervisado, sea suficiente estímulo para que los policías mismos comiencen a evaluar sus propias acciones, asegurándose de que sus partes de tránsito sean pertinentes y justificables. (Algunos críticos advertirán sobre otra consecuencia posible: oficiales que manipulan los registros para ocultar evidencias de acciones basadas en racismo. Nuevamente vemos cómo la decisión de un organismo de llevar registros estadísticos puede hacer que sus miembros modifiquen las cifras resultantes).

En conclusión, recolectar y revisar información sobre la raza de los conductores a quienes se les ordena detenerse mientras conducen en la vía pública, puede servir de algo. Pero datos de esa naturaleza no debieran interpretarse como una medida precisa y perfecta de la

extensión de las prácticas discriminatorias basadas en el perfil racial. Todos los intentos de medir este problema obligan a tomar decisiones que alguien va a objetar. Las cifras resultantes podrán ser de alguna utilidad, pero también contendrán defectos, y las personas tendrán motivos para cuestionar su estatus mágico.

¿POR QUÉ ES NECESARIA LA MAGIA?

Los ejemplos de este capítulo ilustran un dilema muy importante en la vida moderna. Vivimos en un mundo excesivamente complejo y necesitamos de las estadísticas para entender y reducir esa complejidad. El peligro de esto es que tendemos a aprovechar y aferrarnos a cualquier cifra que esté disponible, a tratarla como si fuera una síntesis objetiva y precisa de la realidad. En otras palabras, tomamos a las estadísticas como si tuvieran el poder mágico de resumir y aclarar la realidad y darnos una base sólida para tomar decisiones. El peligro está en que los “juegos de cifras” comienzan apenas la gente se da cuenta de que alguien está interpretando un número como mágico; y las personas tarde o temprano intentarán manipular las cifras para que la magia juegue a su favor.

Una vez entendido el proceso de generación de cifras mágicas —y, querámoslo o no, siempre nos toparemos con ellas en nuestro camino—, es importante tener presente dos razones por las cuales hay que tratar a los números mágicos con especial cuidado.

La primera es que debemos tener en cuenta las elecciones o decisiones en que se basa el cálculo de esas estadísticas. Las personas hacen un destilado de esa complejidad por medio de elecciones, destacando algunos rasgos y dejando de lado otros. Dichas elecciones son tanto inevitables como relevantes. Sin embargo, cuando nos enteramos de una cifra, solemos olvidarnos de considerar cómo esas elecciones moldearon el resultado. Si queremos evitar ser cautivados por las cifras mágicas, es esencial recordar este proceso y todo lo que implica.

La segunda razón es que debemos estar especialmente alertas ante la posibilidad de que las cifras hayan sido manipuladas por individuos con un interés creado en los resultados. Mientras más mágico sea el número, es más probable que alguien que se vea afectado intente manipularlo; y más importante es preguntarse cómo y por qué se generó esa cifra y cómo su estatus mágico afecta las formas en que las personas contabilizan.

Capítulo 6

CIFRAS CONTROVERSIALES

Cuando se trata de estadísticas, es muy fácil encontrar controversias. Las discusiones acerca de las cifras siempre llaman la atención. ¿Se habrán convertido los hispanos en la minoría étnica más grande de EEUU? ¿Se deberían modificar las normas federales respecto a los niveles aceptables de arsénico en el agua potable? ¿Es o no beneficiosa la terapia de reemplazo hormonal? Todo este tipo de situaciones acentúan los debates en torno a los datos.

La suposición extendida de que las estadísticas son capaces de reducir la complejidad llevándola a resúmenes de hechos simples, es más que una manera de atribuirles poderes mágicos a los números. También es un modo de ganar las discusiones. A veces, en los debates acerca de asuntos políticos y sociales, las personas presentan datos estadísticos como si fueran una elocuente carta de triunfo, como hechos que pueden aplastar cualquier opinión contraria. Debido a la creencia de que los números constituyen evidencias contundentes, lo habitual es que no puedan ser ignorados, pero deben ser cuestionados, ya sea mediante preguntas sobre su exactitud o con otras cifras. Como resultado, quienes presentan estadísticas para ganar sus discusiones podrán eventualmente encontrarse discutiendo acerca de números.

No todas las estadísticas inspiran una oposición intensa. Hay quienes apoyan temas de consenso. Por ejemplo, la pornografía infantil

tiene pocos defensores. Las afirmaciones estadísticas respecto a asuntos donde hay acuerdo son aceptadas sin problemas, ya que nadie las revisa acuciosamente. Pero otros temas sociales se vuelven objeto de amargos debates, pues apelan a ideologías o intereses contrapuestos. Y cuando hay fundamentos claros respecto a los cuales oponerse, las estadísticas argüidas por una facción habitualmente provocan críticas de aquellos que apoyan otra parte.

Estas controversias estadísticas —que aquí he denominado *guerras estadísticas*— adoptan diversas formas¹. La más sencilla tiene que ver con el cuestionamiento de la precisión o exactitud de una cifra. La atención de las personas se centra en un número determinado y, por algún motivo, su exactitud es criticada: “¿Será verdad que esa cifra refleja el número exacto de muertes en accidentes de tránsito vinculados al alcohol?”. Con frecuencia la discusión se centra sobre si quienes hicieron los cálculos realizaron un trabajo cuidadoso y completo, o si las definiciones de conceptos o los métodos utilizados resultaron en una cifra demasiado pequeña o demasiado inflada. En algunos casos, como cuando las estadísticas son meras estimaciones, las cifras pueden ser fácilmente cuestionadas. Es habitual que cuando se da a conocer por adelantado una cifra solitaria, ésta sea puesta en duda, porque sirve como una suerte de evidencia resumida de una aseveración determinada (“Este problema amerita ser tratado con seriedad, como se desprende por la gran cantidad de datos disponibles”). En la medida en que una cifra resulta ser clave para el argumento, los oponentes desafiarán ese número, ya sea destacando motivos para dudar de él o contraponiéndose a él con datos propios.

Los debates sobre cifras específicas tienden a darse al inicio de la historia de asuntos de interés público, cuando las personas están tratando de llamar la atención respecto a un determinado problema social, y antes de que hayan tenido tiempo para recolectar mucha información

¹ Sobre guerras estadísticas, véase Best, *Damned Lies and Statistics*, pp. 128-159.

sobre el tema. *Una de las señales de que un tema ha madurado es la proliferación de estadísticas sobre ese tema: hay más gente contabilizando de maneras distintas aspectos diversos del problema.* Lo más probable es que se genere un conjunto de investigaciones al respecto; algunos temas pueden producir cientos o miles de cifras, las que son revisadas en forma acuciosa, ya que los partidarios buscan afanosamente algún dato con poder retórico. A medida que aumenta la cantidad de datos estadísticos disponibles, van incrementándose las oportunidades de encontrar datos que sirvan para apoyar diversos puntos de vista. Es probable que los defensores de algo indaguen en ese cúmulo de datos y encuentren evidencias para apoyar cualquier posición; lo mismo puede suceder con sus contrincantes. Muy pronto las estadísticas se convierten en armas, verdaderas granadas retóricas para ser lanzadas a los oponentes.

Quienes desde ya favorecen una postura determinada en una de estas discusiones, encuentran alivio en las cifras de su bando, al mismo tiempo que los números de sus oponentes les parecerán dudosos y quizás hasta fraudulentos. Aquellos que no tenemos un interés creado en un tema —la audiencia objetiva no involucrada—, solemos convertirnos en el blanco de aseveraciones contradictorias, lo que puede llegar a ser algo muy frustrante precisamente por el flujo de cifras en apariencia contradictorias. Entonces gruñimos: “Solamente díganos qué sustancias químicas provocan cáncer”. No queremos un montón de estadísticas contradictorias, sino datos sencillos.

Pero los hechos se construyen socialmente. Lo que reconocemos como un hecho es producto del esfuerzo de las personas para darle sentido al mundo, hacer una recopilación de suficientes evidencias que sustenten un acuerdo general de que algo es efectivamente cierto. No estoy planteando que no exista un mundo real contra el cual contrastar nuestros hechos: sí lo hay. Todos sabemos que si tomamos una piedra y la soltamos, ésta caerá al suelo. Insistir en que permanecerá suspendida en el aire no evitará que suceda lo contrario. Aun así, lo que se

considera objetivo varía de época en época y de lugar en lugar: por ejemplo, las explicaciones más acreditadas acerca de las causas de las enfermedades cambian entre una sociedad y la siguiente, y a lo largo de las épocas históricas.

Lo que estimamos como objetivo o factual dependerá de la combinación de evidencias y consenso. Las evidencias importan: las afirmaciones de que las enfermedades podían ser causadas por gérmenes recibieron un gran impulso cuando los microscopios permitieron observar a los microbios. Pero también se requiere del consenso: tomó tiempo y mucha investigación antes de que la opinión médica llegara a un acuerdo sobre el valor de la teoría de los gérmenes. Si bien subsisten áreas de disputa, con el paso del tiempo los límites del consenso se han ido ampliando. Pero es la falta de consenso lo que produce nuestra molestia si, semana a semana, los reportajes periodísticos cambian su postura frente a lo que reviste o no una amenaza de cáncer. A su vez, esa ausencia de consenso puede ser un reflejo de que las evidencias disponibles son débiles o inconsistentes.

Aquí tenemos un punto importante. Los debates acerca de lo que es cierto o no, suelen polarizarse en torno a dos posturas débiles donde ninguna es capaz de eliminar a la otra. En un polo están los relativistas, los teóricos posmodernos que insinúan que la realidad es más o menos lo que uno quiere que sea, que no podemos realmente saber nada, que debiéramos estar abiertos a cualquier otra perspectiva y sospechar de cualquier supuesta autoridad. La versión extrema de esta postura justifica todo tipo de creencias paranormales, teorías conspirativas y otras ideas sustentadas en poca, o ninguna, evidencia. El polo opuesto es el ámbito de los absolutistas, que insisten en que los hechos son hechos y que no toleran los cuestionamientos al conocimiento acreditado.

Este libro aboga por una posición intermedia. Somos seres sociales. Todo lo que conocemos del mundo, cada cifra y, lo que es más importante, cada palabra que utilizamos al pensar, está moldeada por nuestra

vida social. Cualquiera que haya observado a un bebé transformarse en un niño, sabe que todos tuvimos que adquirir el lenguaje; y en ese proceso, además aprendimos el modo en que nuestra cultura organiza el mundo en categorías*. La gran contribución de la antropología clásica fue demostrar la diversidad cultural, las múltiples formas en que las personas pueden otorgar sentido a sus mundos. Cada cultura tiene sus propias nociones acerca de las causas de las enfermedades, expectativas de con cuánto recato deben comportarse las mujeres jóvenes, etc. Y cada cultura cree que sus ideas y expectativas son las correctas y verdaderas. Para comprender nuestro mundo, debemos reconocer que todo conocimiento ha sido filtrado por la cultura de su gente. En resumen, al relativismo hay que darle su lugar.

Por otra parte, la ciencia ofrece un estándar especialmente útil para evaluar algunos tipos de conocimiento respecto al mundo. La ciencia es el proceso por el cual las ideas son puestas a prueba con procedimientos en que podrían ser descartadas; las ideas que pasan esas pruebas tienen mayor probabilidad de ser consideradas ciertas y veraces. A lo largo del tiempo, este proceso produce un conocimiento en el cual depositamos gran confianza. Pero no es un proceso exento de problemas; al comienzo, los hallazgos pueden ser aceptados para luego ser retirados de circulación cuando otras evidencias los cuestionen. Las ideas pueden ser ignoradas o rechazadas, y más tarde ser aceptadas, etc. Pero estas irregularidades en la construcción del conocimiento científico no debieran considerarse evidencias de que el proceso a la larga no sirve.

Escribí el primer borrador de este párrafo en mi computador, una máquina producto de siglos de conocimiento científico creciente. Confío absolutamente en que la máquina funcionará, aunque debo

* Un dato numérico interesante: se estima que un niño, durante su aprendizaje, escucha en promedio 54.000.000 de palabras hasta llegar a hablar correctamente. Tomado de Maryanne Wolf, *Cómo Aprendemos a Leer: Historia y Ciencia del Cerebro y la Lectura* (Barcelona: Ediciones B, 2008) (N. del E.).

confesar que tengo un conocimiento muy rudimentario sobre los principios científicos mediante los que opera. Así y todo, sería torpe de mi parte plantear que la ciencia detrás de ese computador es esencialmente arbitraria, sólo una de las tantas maneras de pensar respecto al mundo, ni mejor ni peor que cualquier otra. El computador funciona. Las vacunas funcionan. El conocimiento científico no es sólo una perspectiva válida más entre muchas otras. Podemos confiar en los hallazgos científicos bien establecidos. Resumiendo: hay que darle cabida a la autoridad basada en evidencias.

No obstante, la ciencia no puede responder todas las preguntas. Puede decirnos cómo y por qué algunas personas se enferman (si bien todavía, a estas alturas, no puede explicar todas las enfermedades). Pero la ciencia no nos puede decir cuán púdica ha de ser la conducta de las mujeres jóvenes, pues ése no es un tema sujeto a evaluación científica. Precisamente debido a que tenemos grandes expectativas de la ciencia, sus limitaciones plantean un problema a la cultura. Cuando nos enfermamos, esperamos que un médico elabore un diagnóstico e indique un tratamiento, y nos frustramos si eso no sucede. Incluso utilizamos resultados de alguna investigación para documentar pautas sociales o evaluar riesgos en la recomendación de formas de conducta. Nuestra sociedad considera que los datos —las estadísticas— ofrecen, si no respuestas acabadas, por lo menos información relevante para idear las soluciones a muchos tipos de preguntas, incluyendo algunas que no necesariamente caben dentro del ámbito de la ciencia.

Cuando nos enfrentamos a las estadísticas, debemos evitar el relativismo y el absolutismo extremos. Cabe recordar que *las estadísticas son un producto social y que el proceso mediante el cual son creadas, inevitablemente moldea las cifras resultantes*. Pero también hay que tener en cuenta que la ciencia ofrece medios para ponderar las evidencias, para evaluar la exactitud de los números. Estos asuntos cobran importancia tan pronto como las estadísticas se transforman en el tema de los desacuerdos.

CIENCIA Y CIENCIA BASURA

Ciencia basura es una expresión de moda utilizada para descartar hallazgos de investigaciones dudosas. Ya que en nuestra cultura la ciencia es considerada una fuente autorizada de conocimientos, muchos individuos se autodenominan científicos, para así legitimar sus puntos de vista. Rotular algún tipo de afirmación como “científica” es una práctica moderna para otorgarse legitimidad y autoridad. Así es como algunos grupos religiosos que se oponen a la enseñanza de la teoría de la evolución, argumentan que ellos representan “la ciencia de la creación” —o creacionismo— e insisten en que las explicaciones bíblicas respecto al origen del universo, la Tierra y la vida humana deben estar en igualdad de condiciones con aquellas dadas por físicos y biólogos. Con un criterio parecido, todo tipo de parapsicólogos, sanadores psíquicos y defensores del movimiento perpetuo etiquetan sus posturas con el rótulo de “científicas”².

Pero la ciencia es más que una etiqueta: es una vocación dirigida hacia las evidencias. Los científicos deben estar preparados para poner a prueba sus ideas y estar dispuestos a que sean invalidadas por esas pruebas. No estamos hablando del sencillo proceso de comprobación de hipótesis mencionado en los textos de enseñanza media. Los científicos son humanos y, como tales, pueden dejarse llevar por sus ideas, culpando a veces al modo como se condujeron las pruebas cuando éstas fallan: algo anduvo mal con las condiciones del experimento, se requieren nuevos tests, etc. Podemos centrarnos en las debilidades de los científicos que se aferran a sus ideas ante ciertas evidencias que las cuestionan o invalidan; enfocarnos en esa conducta nos permite hacer un retrato de la ciencia que recalca sus

² Para conocer algunas de las afirmaciones recientes más discutibles en relación con el estatus científico, véase Robert Park, *Ciencia o Vudú: De la Ingenuidad al Fraude Científico* (Barcelona: Grijalbo Mondadori, 2001).

defectos³. Algunos críticos relativistas plantean que los desacuerdos al interior de la ciencia sólo aportan un nuevo punto de vista, ni más ni menos válido que otros. Quizás una salida de este embrollo sea reconocer a la ciencia como un ideal y, al mismo tiempo, asumir que los científicos en tanto individuos pueden estar muy lejos de ese ideal.

Sin embargo, a lo largo del tiempo, a medida que han ido aumentando las evidencias, la ciencia ha ido acumulando un caudal de conocimientos en los cuales hemos depositado gran confianza, basados en la fiabilidad con que son confirmadas sus predicciones. Este avance científico depende de una comunidad que en su totalidad exige un autoanálisis riguroso y continuo, examinando las ideas para determinar si se basan en evidencias reales y comprobables. Ya que todas las pruebas pueden fallar, la aplicación acumulativa de múltiples exámenes es lo que, a final de cuentas, suministra la base para que eventualmente la ciencia acepte sólo aquellas ideas que se sostienen aun bajo la más estricta revisión.

Los estudios individuales no pueden lograr esto. Cada estudio —cada prueba, cada investigación— tiene limitaciones y fallas relativas a sus métodos, lo que lo vuelve un objetivo apropiado para una crítica justificada. Los estudios debieran poder replicarse y, además, inspirar futuras investigaciones que utilicen distintas metodologías (las que, a su vez, es probable que tengan sus propias limitaciones y falencias). La confianza en un determinado hallazgo crece cuando la replicación y el uso de diversos métodos de investigación confirman los mismos resultados. La mayoría de los científicos coincide en que los resultados de un estudio único, especialmente si la investigación sugiere serios problemas metodológicos, no debieran considerarse acreditados. Sólo el

³ Para conocer estudios sobre esfuerzos científicos de defensa de posturas marginales, véase H.M. Collins, "Surviving Closure: Post-Rejection Adaptation and Plurality in Science", *American Sociological Review* 65 (2000): 824-845; Bart Simon, *Undead Science: Science Studies and the Afterlife of Cold Fusion* (New Brunswick, NJ: Rutgers University Press, 2002).

tiempo y más investigación pueden distinguir entre hallazgos erróneos o confiables.

*Desafortunadamente, la cobertura periodística de la investigación en ciencias tiende a ser más impaciente que la comunidad científica*⁴. Los medios periodísticos prefieren los temas que parezcan novedosos, inesperados y hasta dramáticos; por eso, las historias más llamativas tratan de algún hallazgo repentino, no de la replicación o confirmación de un descubrimiento anterior mediante un diseño de investigación distinto. Por lo tanto, la prensa opta por cubrir precisamente aquellos resultados que carecen de una verificación apropiada. Un estudio individual con resultados inquietantes será noticia y los medios tenderán a minimizar, o incluso ignorar, las limitaciones metodológicas de esa investigación. En consecuencia, debemos ser cautelosos con los comunicados de prensa relativos a resultados de investigaciones. Tal como vimos en el Capítulo 4, dados los esfuerzos de algunas prestigiosas revistas para ganar mayor notoriedad por medio de estos comunicados sobre los artículos que publican, hay que ser extremadamente cautelosos.

En general, el rótulo peyorativo “ciencia basura” implica una crítica metodológica y plantea que la investigación fue diseñada, o los datos fueron recogidos, de tal forma que no es posible confiar en esos resultados. Usualmente esto sugiere que la investigación fue guiada por una agenda particular donde los hallazgos fueron manipulados para apoyar una postura determinada. El uso original de la expresión “ciencia basura” se refería a las declaraciones de los testigos peritos en los juicios⁵. Los abogados interrogaban a testigos peritos con la esperanza de que su experticia convenciera a jueces y jurados de que ciertos argumentos eran objetivamente válidos y estaban respaldados por investigación científica. Cuando una parte en un juicio invita a declarar a un testigo

⁴ Dorothy Nelkin, *La Ciencia en el Escaparate* (Madrid: Fundesco, 1990).

⁵ El término se masificó gracias a Peter W. Huber, *Galileo's Revenge: Junk Science in the Courtroom* (New York: Basic Books, 1991).

perito (y le paga por ello), es razonable preguntarse si ese testimonio será completo, imparcial y representativo del consenso científico.

Por ejemplo, veamos el caso de los implantes mamarios “tóxicos”. Entre fines de la década de los 80 e inicios de los 90, el riesgo para la salud que éstos implicaban se volvió tema de gran interés público: en consecuencia, la Federal Drug Administration (FDA) prohibió los implantes rellenos de silicona. El tema tuvo una gran cobertura mediática y se presentó una multimillonaria demanda colectiva. Quienes criticaban los implantes vieron reforzada su postura gracias a médicos y científicos que presentaron evidencias de que algunas mujeres que los usaban contraían ciertas enfermedades. Aquí podemos ver un razonamiento bastante conocido: alguien se enferma e intenta comprender la causa de su mal, recuerda alguna experiencia anterior (como que se implantó silicona en las mamas) y concluye que eso debió causar la enfermedad. La lógica puede parecer perfectamente convincente para un individuo, pero no se puede considerar como una demostración científica.

La ciencia exige, entre otras cosas, datos epidemiológicos que la respalden. Sabemos que la gente contrae enfermedades, por lo que debemos esperar que algunas mujeres con implantes mamarios se enfermen, lisa y llanamente, porque son seres humanos. Pero la pregunta clave es si las mujeres que tienen implantes mamarios son más propensas a enfermar que las que no tienen prótesis. No olvidemos la discusión del Capítulo 3 acerca del riesgo: antes de concluir que los implantes mamarios son la causa probable de la enfermedad, hay que recordar que el estándar usual para esas comparaciones es que la tasa de enfermedad sea al menos un 200% mayor entre las mujeres con prótesis que en el grupo control, es decir, el grupo de mujeres sin implantes. En general, los estudios epidemiológicos no indicaban esa tasa entre las

⁶ Marcia Angell, *Science on Trial: The Clash of Medical Evidence and the Law in the Breast Implant Case* (New York: Norton, 1996).

mujeres que tenían implantes. Estos datos debieron considerarse un cuestionamiento bastante serio a los argumentos que planteaban que las prótesis eran riesgosas, pero los críticos de los implantes ganaron la batalla de las relaciones públicas (y muchos de los casos judiciales) en parte porque los resultados de los estudios epidemiológicos no se conocieron hasta muy avanzado el asunto en cuestión.

Uno de los problemas de la noción de ciencia basura es que el término está políticamente cargado: los conservadores suelen utilizarlo para descalificar las afirmaciones de ambientalistas, defensores de los consumidores y otros activistas que advierten sobre los peligros de la sociedad contemporánea⁷. En respuesta a ello, los críticos progresistas plantean que “el término ciencia basura sirve para reconciliar... el sesgo pro-empresarial con las pretensiones de superioridad científica”⁸. Cada bando plantea que los científicos del lado opuesto están sesgados y no se puede confiar en ellos para diseñar una investigación seria. Para los no especialistas es difícil evaluar estas demandas y contrademandas, aunque las diferencias en los hallazgos de la investigación sólo se deriven de suposiciones, definiciones y elecciones metodológicas irreconciliables.

Por ejemplo, los científicos que trabajan en conjunto con los ambientalistas pueden definir como un riesgo peligroso para la salud la exposición poco frecuente a una concentración muy baja de sustancias radioactivas. Por su parte, los científicos que trabajan para la industria argumentan que una exposición más frecuente a una concentración mayor de la misma sustancia no plantea un riesgo inaceptable para las

⁷ Para una declaración de principios conservadora sobre el tema, véase Steven J. Milloy, *Junk Science Judo: Self-Defense Against Health Scares and Scams* (Washington, DC: Cato Institute, 2001).

⁸ Sheldon Rampton y John Stauber, *Trust Us, We're Experts: How Industry Manipulates Science and Gambles with Your Future* (New York: Tarcher/Putnam, 2001), p. 265.

personas⁹. Ambos bandos podrán insistir en que, científicamente hablando, su posición es la más sólida y que su método de evaluación de riesgos es el más apropiado; pero dada la necesidad de sopesar argumentos de expertos en pugna, los que no son científicos quedan sumamente frustrados.

Aun cuando su vinculación con algunas posiciones ideológicas particulares impida rehabilitar la expresión “ciencia basura”, en esta discusión ronda una idea bastante útil. Todas las investigaciones tienen sus limitaciones e inevitablemente los investigadores eligen definiciones, medidas, técnicas analíticas y diseños específicos. Estas decisiones son clave, ya que moldean los resultados de cada estudio. Nunca se podrá tener tanta confianza en los resultados de un estudio único como la que se puede tener en un conjunto de investigaciones, pues las decisiones de muchos científicos han colaborado en la neutralización de las limitaciones de cada uno de los estudios por separado. Nuestra certeza de que fumar provoca cáncer pulmonar no se basa en una sola experiencia científica, sino más bien en un gran conjunto de investigaciones que han utilizado diversas metodologías y que, sobre todo, vinculan al cigarrillo con el cáncer. De seguro, algunos investigadores están sesgados de tal modo que diseñan su investigación para obtener ciertos resultados; no olvidemos que en una oportunidad, el Tobacco Institute financió estudios pensados para obtener resultados que sugieran que fumar no era particularmente dañino.

Sin embargo, el principal problema de la llamada ciencia basura no se refiere a los motivos o tendencia política de los investigadores, sino más bien a la propensión de quienes defienden alguna postura de proclamar uno o dos estudios como inmejorables. En esos casos, las preocupaciones de índole política abortan el proceso de reunir información científica.

⁹ Colin McMullan y John Eyles, “Risky Business: An Analysis of Claimsmaking in the Development of an Ontario Drinking Water Objective for Tritium”, *Social Problems* 46 (1999): 294-311.

Las discusiones sobre la ciencia basura tienen otra característica importante. Tienden a incluir desacuerdos en torno a las nociones de riesgos y compensaciones (ver Caps. 1 y 3): ¿esta sustancia química (procedimiento médico, proyecto hidroeléctrico) causará un daño irreparable o inaceptable? Desde luego que aquí las personas razonables podrán estar en desacuerdo con todo tipo de temas. ¿Cómo se debiera medir un eventual daño? ¿Cómo podríamos sopesar los perjuicios (o costos) y los beneficios proyectados (y cómo medirlos)? Si bien quienes defienden alguna postura podrían tratar de caracterizar esas discusiones como una lucha entre el bien y el mal, la evaluación de las evidencias científicas nunca es tan sencilla como eso.

MANIPULACIÓN Y SELECCIÓN CUIDADOSA

Las discusiones relativas a estadísticas sociales rara vez se inician con disputas respecto a una cifra. Más bien, casi siempre comienzan como desacuerdos en torno a la importancia de un tema social o la solución a un problema entre quienes defienden una postura u otra y traen a colación cifras para apoyarlas. El discurso político reciente denomina manipulación (*spinning*) a la práctica de entregar a los medios una interpretación de los hechos que coincida con una postura determinada; ello, con la esperanza de que los medios den a conocer —y posiblemente refrenden— ese punto de vista¹⁰. Las cifras pueden ser objeto de esta manipulación.

Consideremos las interpretaciones en pugna relativas a la información del censo 2000 que indicaba que una creciente proporción —la

¹⁰ Para revisar comentarios sobre el proceso de manipulación, véase, por ejemplo, George Pitcher, *The Death of Spin* (Hoboken, NJ: Wiley, 2003); Lynn Smith, "Putting a Spin on the Truth with Statistics and Studies", *Los Angeles Times*, June 6, 2001, p. E1.

cuarta parte— de hogares estaba compuesta por personas solas¹¹. Para los conservadores pro familia, estos datos confirmaban la decadencia de la familia estadounidense tradicional y, por ende, la necesidad de políticas sociales de promoción de la familia. Otros comentaristas más progresistas interpretaban en términos más positivos este aumento en la cantidad de hogares unipersonales: el creciente bienestar económico y una mejor salud implicaban que los jóvenes podían vivir de manera independiente, la gente ahora era capaz de terminar matrimonios insatisfactorios y las personas mayores mantenían sus propios hogares. Entonces, los datos del censo que evidenciaban el aumento de la cantidad de hogares unipersonales podían interpretarse ya sea como deterioro social o mejora en las condiciones de vida. Nótese que nadie discutió la exactitud de la cifra; los individuos pueden reconocer que una estadística es correcta sin necesariamente estar de acuerdo en su significado. El vaso podrá verse como medio lleno o medio vacío; todo depende del sentido que se le dé.

Esta manipulación es estimulada por la existencia de ideologías bien articuladas y en pugna. Estamos acostumbrados a las interpretaciones irreconciliables entre demócratas y republicanos, conservadores y progresistas, y las estadísticas ofrecen una buena oportunidad para que estos contrincantes las manipulen. Por ese motivo, la información de que un creciente porcentaje de niños estadounidenses sufre de obesidad es una invitación a las críticas de izquierda que apuntan a la industria alimentaria, pues ésta publicita productos de alto contenido calórico. La derecha, en cambio, critica que los programas de alimentación de las escuelas públicas son fuente de obesidad infantil¹². En la mayoría

¹¹ D'Vera Cohn, "Married-With-Children Still Fading; Census Finds Americans Living Alone in 25% of Households", *Washington Post*, May 15, 2001, p. A1.

¹² Compárese Eric Schlosser, *Fast Food Nation: The Dark Side of the All-American Meal* (New York: Houghton Mifflin, 2001), pp. 239-243 (*Fast Food Nation: ¿Quiere una Ración de Mentiras con su Menú?* [Barcelona: Debolsillo, 2007]); y Douglas J. Besharov, "Growing Overweight and Obesity in America: The Potential Role of

de los casos, los argumentos elegidos, los factores que se dicen responsables del problema y la naturaleza de las soluciones propuestas son predecibles para cualquiera que esté familiarizado con las ideologías.

Mientras más números haya disponibles, mayores oportunidades habrá para manipularlos. El gobierno federal recolecta amplia información sobre problemas de índole social como el consumo de drogas. Las encuestas a alumnos del último año de enseñanza media (Monitoring the Future [MTF]) son uno de los instrumentos estándar para hacer seguimiento del uso de drogas. Aplicadas por años, las MTF entregan información a partir de encuestas autoaplicadas acerca del consumo de diversos tipos de drogas por parte de estos jóvenes en distintos períodos de tiempo. Por ejemplo, en 2000, el 21,6% de los alumnos del último año de enseñanza media informó haber fumado marihuana y el 50% declaró haber bebido alcohol en los 30 días previos a la encuesta¹³. ¿Qué debemos pensar sobre estas cifras? ¿Las cosas están peor o mejor? Depende de los años y las drogas que estemos comparando. En 1990 —10 años antes—, el 14% de los alumnos informó consumir marihuana; por lo tanto, el consumo en 2000 era un 50% mayor. Sin embargo, durante ese mismo período, el consumo de alcohol disminuyó de 57,1% a 50%. No todas las tendencias son así de constantes; casi todos los informes de las MTF entregan cifras más que suficientes para que alguien lo bastante cuidadoso las tome y plantee, basándose en estadísticas, que el uso de drogas en adolescentes ha aumentado o disminuido en un determinado período de tiempo.

Esas discusiones son una forma de selección cuidadosa (*cherry-picking*) de estadísticas para apoyar una tesis específica y llamar la

Federal Nutrition Programs”, testimonio entregado a la Comisión de Agricultura, Nutrición y Silvicultura del Senado de EEUU (American Enterprise Institute, 2003).

¹³ Kathleen Maguire y Ann L. Pastore, eds., *Sourcebook of Criminal Justice Statistics*, 2000 (Washington, DC: Bureau of Justice Statistics, 2001), p. 248.

* Literalmente, se refiere a la cosecha de cerezas, en que sólo se escoge la fruta más

atención hacia esas cifras, mientras se ignoran otras que podrían llevar a una conclusión distinta. Lo que hace toda la diferencia es la cantidad de datos disponibles; entre más cifras haya para elegir, más certeza habrá de encontrar “cerezas” apropiadas, ya maduras para ser cosechadas. Esta técnica puede ser utilizada por cualquier persona. Los políticos que ostentan un cargo oficial pueden destacar con orgullo las evidencias de que su desempeño ha mejorado, mientras sus opositores argumentan que los hechos demuestran que las cosas se han deteriorado (y que van a empeorar si no eliminamos a esos pillos). Igualmente, los que proponen determinadas ideologías pueden seleccionar números que parecen confirmar sus ideas.

Un indicador de que la información ha sido seleccionada tal como mencionamos, es que el defensor de alguna postura presenta datos muy específicos para apoyar un argumento demasiado amplio. Es muy difícil detectar esta selección si los datos originales no son revisados cuidadosamente. Por ejemplo, alguien podría decir: “¡El porcentaje de alumnos del último año de enseñanza media que informó haber probado heroína en los 30 días previos a la encuesta aumentó en un 40% entre 1997 y 2000!”. Si bien esto es cierto (el consumo declarado subió de 0,5% a 0,7%, un aumento del 40%), quien lo plantea desconoce datos de las propias MTF que indican que el consumo de drogas —incluyendo marihuana, cocaína, alcohol y cigarrillos— entre estos alumnos disminuyó. Pero sólo quien preste mucha atención podría preguntarse por qué esa persona quiso centrarse en una droga específica (especialmente, en una tan poco consumida por los estudiantes de enseñanza media).

madura y sana. Un observador que sólo tenga acceso a la fruta seleccionada de esta manera, tenderá a creer, erróneamente, que casi toda, o incluso toda, la fruta está en tan buenas condiciones. Metafóricamente, se utiliza para indicar el acto de señalar alguna información específica que confirme una posición particular, en tanto se desconocen otros datos que podrían contradecir esa postura. También se denomina *data dredging*, término peyorativo que indica la búsqueda ex profeso de ciertas relaciones estadísticamente significativas (N. de la T.).

De ahí que las estadísticas pueden ser objeto de manipulación y —cuando son seleccionadas cuidadosamente— una herramienta para los manipuladores. La manipulación puede dar como resultado cifras bastante buenas, pero dado que están fuera de contexto, carecen de matices y complejidad. Entonces, las estadísticas pueden servir para promover el punto de vista de quien quiera incorporarlas en un debate.

No es mi intención resaltar la idea de que algunas cifras son correctas y que otras han sido “manipuladas”; más bien, trato de advertir que cada cifra dada a conocer públicamente puede haber sido escogida entre muchas otras porque alguien que defendía una postura determinada consideró que tenían potencial retórico. Cada vez que en una discusión sobre temas sociales las cifras parezcan apoyar una postura en particular, debemos estar alertas a los signos de manipulación o de selección cuidadosa.

CUÁNTOS JUDÍOS Y MUSULMANES HAY EN EEUU

Quienes, por defender una postura, seleccionan sus estadísticas, reconocen que las cifras pueden tener importancia simbólica en las discusiones sobre temas sociales. Incluso responder las preguntas más sencillas —¿cuántos?, ¿muchos o pocos?— tiene un valor simbólico, porque las distintas respuestas podrían aparecer apoyando un lado u otro en los conflictos sociales.

Veamos qué sucede con las actuales discusiones relativas a la población musulmana y judía en EEUU. A primera vista, la cantidad de seguidores de una religión en particular podría ser un tema poco interesante; por ejemplo, la cifra de creyentes presbiterianos no llama mucho la atención fuera de ese grupo religioso. Sin embargo, varios grupos están analizando el número de judíos y musulmanes, pues estas cifras pueden tener repercusiones en la opinión de los estadounidenses sobre el conflicto palestino-israelí, el terrorismo y otros problemas

de política internacional, así como en las perspectivas futuras de estas religiones en EEUU.

Por razones obvias, luego de los ataques terroristas del 11 de septiembre de 2001, el número de musulmanes en EEUU se convirtió en un tema candente. Hubo quienes temieron que el gobierno o la opinión pública culparían a todos los musulmanes, lo que podría llevar a una ola de crímenes por odio racial, o incluso a políticas represivas similares a la reubicación sufrida por los japoneses que vivían en EEUU durante la Segunda Guerra Mundial. Algunos defensores de los musulmanes plantearon que, para que sus intereses y preocupaciones fueran objeto de mayores garantías, debían ser reconocidos como una minoría religiosa importante en EEUU.

Los cálculos más recientes de la población musulmana estadounidense varían entre 2 y 10 millones. Puede parecer un rango demasiado amplio, pero recordemos (como vimos en el Cap. 1) que el censo —generalmente la fuente más autorizada sobre estadísticas de población— no recoge información relativa a la religión de las personas. Por eso es necesario encontrar otras formas de calcular estas cifras. Algunos analistas utilizan el origen como base de sus cálculos y suponen que los individuos cuyos antepasados (o ellos mismos) provienen de países mayoritariamente musulmanes, también lo son. Otros han intentado calcular la cantidad de mezquitas en EEUU, multiplicando esa cifra por un número promedio de asistentes, al que luego agregan otra cantidad de musulmanes que no asisten a las ceremonias en la mezquita y sobre los cuales no existen registros. Hay investigadores que obtienen sus datos de encuestas en que se consulta a las personas sobre su religión. Cada uno de estos métodos tiene sus limitaciones. Por ejemplo, la religión que profesan los inmigrantes quizás no refleje el patrón general de filiación de la población en su país de origen (así como los inmigrantes en la Massachusetts colonial probablemente eran más puritanos que la población promedio en Inglaterra). Por otra parte, también es difícil identificar todas las mezquitas, determinar un promedio

de sujetos que asisten a ellas y evaluar la cantidad de musulmanes que no lo hacen. Y, dado que no todos cooperan con las encuestas, los resultados pueden subestimar las poblaciones a medir¹⁴.

Los métodos elegidos —y las críticas de los métodos rivales— difieren según la postura del investigador dentro de la discusión mayor. Por ejemplo, en su interés por demostrar que su religión tiene muchos seguidores, varias importantes organizaciones musulmanas encargaron un estudio basado en las mezquitas. Perfectamente comprensible. El estudio concluyó que alrededor de 2 millones de personas asistían a las mezquitas y luego, suponiendo que una cantidad aún mayor de musulmanes no lo hacía, planteó que la población musulmana oscilaba entre 6 y 7 millones. Algunas organizaciones judías consideraron que esta cifra era demasiado alta y señalaron que los resultados de la encuesta daban un número algo por debajo de los 2 millones. Por su parte, los musulmanes sugerían que las encuestas suelen no contabilizar a quienes, por miedo o barreras lingüísticas, no informan de su religión. Podemos sospechar que la cantidad de judíos (estimada entre 5 y 6 millones) sirve como un importante parámetro en el debate: las organizaciones musulmanas favorecen los cálculos que proponen que la población musulmana es mayor que la judía, en tanto las organizaciones judías prefieren los números que sugieren que hay menos musulmanes.

Si bien los esfuerzos recientes por dimensionar la población musulmana atrajeron mucha atención, la discusión sobre la cantidad de judíos en EEUU, igualmente intensa, ha permanecido confinada a la comunidad hebrea. En este caso, la preocupación se refiere no sólo a que los judíos son una comunidad religiosa pequeña, sino también que la cifra estaría disminuyendo, algo que ha llevado a algunos

¹⁴ La mejor visión general de estos cálculos en disputa se encuentra en Tom W. Smith, "The Muslim Population of the United States: The Methodology of Estimates", *Public Opinion Quarterly* 66 (2002): 404-417. Para ejemplos de las críticas planteadas por quienes rebaten lo anterior, véase Bill Broadway, "Number of U.S. Muslims Depends on Who's Counting", *Washington Post*, November 26, 2001, p. A1.

comentaristas a advertir que los judíos estadounidenses están “desapareciendo”¹⁵. Esta preocupación, de larga data, se intensificó después de conocerse los resultados de algunas investigaciones recientes. En 1990 y 2000, las más importantes organizaciones judías patrocinaron las Encuestas de Población Judía Nacional (NJPS, por su sigla en inglés), estudios a gran escala diseñados para medir la población judía en EEUU. Estas encuestas revelaron que una leve mayoría de judíos se estaba casando con no-judíos (aumentando la posibilidad de que los hijos de esas parejas no fueran criados como judíos) y que las mujeres judías estaban teniendo un promedio de 1,8 hijos (cantidad insuficiente para reponer la población judía). Particularmente polémicas fueron las noticias relativas a que el cálculo de la población judía para 2000 (5,2 millones) era menor al de 1990 (5,5 millones). Los críticos argumentaron que las NJPS habían contabilizado mal, pues la cifra total correcta de judíos rondaba los 6,7 millones.

Esta discusión giraba en torno a los desacuerdos respecto a cómo definir lo que es ser judío. ¿Es un asunto de práctica religiosa? ¿O son judíos aquellos que se consideran judíos? Las NJPS consideraron ambas categorías. Pero, ¿qué sucede con quienes afirman que fueron judíos pero ya no lo son, o que viven en hogares con personas que no se consideran judías? Los sujetos en estas categorías, excluidos por las NJPS, fueron incluidos en los cálculos de los críticos. Además, éstos planteaban que el temor al antisemitismo habría llevado a algunos encuestados de las NJPS a negar que fueran judíos. Como siempre, una definición más estrecha resultará en un cálculo menor que una definición más amplia.

En cierto nivel, los debates sobre los cálculos de musulmanes y judíos en EEUU pueden considerarse un tema de métodos y

¹⁵ Calvin Goldscheider, “Are American Jews Vanishing Again?”, *Contexts* 2 (Winter 2003), pp. 18-24. Respecto a la Encuesta sobre Población Judía en EEUU en el 2000, véase Daniel J. Wakin, “Survey of U.S. Jews Sees a Dip; Others Demur”, *New York Times*, October 9, 2002, p. A23.

definiciones. Si intentamos contabilizar a los musulmanes calculando cuántas personas están afiliadas a las mezquitas y luego agregamos una cantidad referente a quienes no lo están, obtendremos una cifra (alta); si utilizamos los resultados de la encuesta, tendremos una segunda cifra (menor). Asimismo, el definir la identidad judía de manera restringida produce un cálculo menor que si optamos por una definición más amplia. En un nivel técnico, los cientistas sociales podrán argumentar respecto a las ventajas y limitaciones de los distintos métodos y definiciones. (Es probable que el consenso favorezca la utilización de encuestas para estimar la población musulmana y la definición más limitada de judío dada por las NJPS, pero, de seguro, no todos estarán de acuerdo).

Sin embargo, éstas no son meras preguntas retóricas. Estas estadísticas tienen una importancia simbólica: plantear que un grupo es grande o está creciendo puede sugerir que sus intereses son más relevantes que los de una comunidad menor o que se está reduciendo. El tomar en cuenta ese tipo de mensaje con obvias repercusiones políticas podría llegar a apasionadas defensas de cifras que de otro modo no podrían sustentarse sólo en fundamentos técnicos.

RESULTADOS DE LA REFORMA AL BIENESTAR SOCIAL

La ley federal de reforma al bienestar social de 1996 (Ley de Reconciliación de Responsabilidad Personal y Oportunidad de Empleo [PRWORA, por su sigla en inglés]) fue el resultado de décadas de ardua discusión. Por largo tiempo, los críticos del bienestar se quejaron de que el sistema alentaba la dependencia a largo plazo, que no sólo algunos beneficiados vivían de ella sino también sus hijos, los que llegarían a adultos haciendo uso del sistema, en un complejo ciclo intergeneracional. El argumento central era que la seguridad social no alentaba la independencia ni la responsabilidad personal, y que esa carencia de autonomía se perpetuaría en el tiempo. Por el contrario,

quienes la defendían insistían en que era necesaria en tanto red de seguridad vital que brindaba la protección mínima a los individuos que tenían escasos recursos para sobrevivir en una sociedad que ofrece oportunidades limitadas. Más que considerar a los beneficiarios como individuos incapaces de responsabilizarse por sí mismos, culpaban al sistema de ofrecer pocas fuentes de trabajo y ser excesivamente discriminatorio.

A la vez, tanto los críticos como quienes apoyaban las prestaciones sociales entendían que éstas se vinculaban a otros problemas sociales: los beneficiarios eran, por definición, pobres; además, tendían a tener menos años de educación y problemas de salud más serios que los que no recibían subsidios. Muchas mujeres beneficiarias tenían hijos fuera del matrimonio y los padres de éstos solían estar cesantes. Para los críticos, las prestaciones sociales disuadían de buscar empleo y contraer matrimonio; para sus defensores, la ausencia de oportunidades de empleo decente creaba las circunstancias sociales que obligaban a las personas a recurrir a esas prestaciones y permanecer en el sistema. Los defensores clamaban por más beneficios para mejorar las condiciones de vida de los usuarios, así como capacitación laboral y otros programas de mejoramiento de su calidad de vida. Por su parte, los críticos argumentaban que aumentar los beneficios y expandir los programas sólo lograría acrecentar la dependencia. La discusión se extendió por décadas.

La PRWORA —aprobada por un Congreso liderado por republicanos, pero firmada por el presidente Bill Clinton— fue diseñada para “terminar con la idea de bienestar tal como la conocemos hoy en día”. En particular, la nueva ley reemplazó el antiguo Programa de Ayuda para Familias con Hijos Dependientes (AFDC, por su sigla en inglés) por la Asistencia Temporal a Familias Necesitadas (TANF, por su sigla en inglés), consistente en fondos para los distintos estados, otorgándoles a éstos la discreción suficiente para diseñar las ayudas dirigidas a los pobres. Además, la nueva ley estableció un límite para la mayoría de los beneficiarios con un tope de 60 meses de asistencia en dinero

proveniente de los fondos federales; junto con ello, exigía que los beneficiarios estuvieran empleados a los dos años de estar recibiendo esa ayuda monetaria.

Éstos fueron considerados cambios significativos y drásticos. Quienes apoyaban la ley (en su mayoría conservadores) preveían un futuro optimista, en que los otrora beneficiarios dependientes aprenderían de responsabilidad personal, tomarían las riendas de su vida y saldrían de la pobreza. Sus críticos (mayoritariamente progresistas) advertían de una inminente catástrofe social en que la disminución del apoyo estatal forzaría a millones de personas a caer en la pobreza y aumentaría la cantidad de gente sin casa. Los defensores promovían la nueva ley como un incentivo a la independencia; los críticos insistían en que era demasiado severa y castigadora.

El quinto aniversario de la legislación sobre bienestar social —fecha importante, pues marcaba los 60 meses de límite de ayuda monetaria de la PRWORA— brindó una ocasión para evaluar el impacto de la ley¹⁶. Sus defensores dieron a conocer numerosas estadísticas que evidenciaban su éxito: la cantidad de hogares sin subsidios bajó del máximo de 1994 para el AFDC (5,1 millones) a casi 2 millones que permanecían en el sistema; la proporción de madres solteras que trabajaban había crecido y la cantidad de hijos de madres solteras —número que había aumentado rápidamente en los años previos a la reforma al bienestar social— subió muy poco. Incluso los críticos más acérrimos de la PRWORA debieron admitir que, por lo menos hasta ese momento, no se había producido la temida catástrofe.

¹⁶ Cada vez hay más literatura que interpreta las consecuencias de la reforma al bienestar social. Para revisar distintas perspectivas, véase Douglas J. Besharov y Peter Germanis, "Welfare Reform —Four Years Later", *The Public Interest* 140 (2000): 17-35; Christopher Jencks, "Liberal Lessons from Welfare Reform", *The American Prospect* (Special Supplement, Summer 2002): A9-A12; y Sanford F. Schram y Joe Soss, "Success Stories: Welfare Reform, Policy Discourse, and the Politics of Research", *Annals of the American Academy of Political and Social Science* 577 (2001): 49-65.

¿Qué había ocurrido? Los analistas concluyeron que las estadísticas generales se vieron afectadas por varios sucesos. El primero fue la fecha escogida: en los primeros años de la reforma hubo un auge económico donde el desempleo era bajo y había bastante oferta de trabajo. En segundo lugar (y probablemente más importante, aunque no fue tan publicitado como la reforma al bienestar social), el gobierno federal había ampliado las políticas para apoyar a los trabajadores de bajos ingresos, tales como la creación del Crédito Impositivo por Ingreso (EITC, por su sigla en inglés), el aumento de ayuda para el cuidado infantil y la ampliación del acceso al sistema de subsidios para la salud, Medicaid. Debido a que gran parte del rechazo a las prestaciones sociales se centraba en que los beneficiarios no trabajaban, estos programas dirigidos a personas que sí tenían empleo lograron un apoyo mucho más amplio. La extensión de los beneficios entregados por estos programas implicó que al menos algunos trabajadores de bajos ingresos que en el pasado habrían tenido que recurrir a la asistencia social para calificar para una prestación determinada (p.ej., pagar las cuentas médicas de un hijo), ahora podían seguir empleados.

Mucho se habló respecto a la cantidad de beneficiarios que salieron del registro de estas prestaciones sociales, pero en ciertos casos esto se debió a una lectura selectiva de las evidencias. La mayoría de los beneficiarios sólo recibía ayuda por un corto plazo; durante cada año de la historia del programa, hubo gente que se retiraba de los programas de subsidios. Era difícil establecer si lo que hubo fue un aumento importante en la cifra de personas que ya no estaban en el registro de estas prestaciones; quienes llevaban esos registros post-PRWORA contaron cuidadosamente esas cantidades, pero los registros de los años previos no eran tan completos. La gran diferencia pudo deberse a un tercer factor: la PRWORA, que permitió a los diferentes estados de la unión establecer los requisitos, dificultó el acceso a los beneficios. En vez de ser sumados a los registros de quienes recibían prestaciones sociales en primer lugar, los posibles beneficiarios podrían verse obligados a

postular a un trabajo o cumplir con otro requisito que los llevaría a buscar otras opciones, en vez de recurrir de inmediato a la asistencia social.

Los diferentes comentaristas ponderaron estos factores de distintas formas, todas bastante predecibles. Los críticos de la reforma al bienestar social tendían a aumentar la importancia de una saludable economía en el aparente éxito del programa (y veían con recelo cómo concluía el auge económico). También destacaron la relevancia de los diversos programas de apoyo dirigidos a los pobres que tenían empleo (generalmente más populares entre progresistas que entre conservadores). Por otro lado, los partidarios de la PRWORA planteaban que el éxito del programa demostraba que en muchos casos las prestaciones sociales no eran necesarias y pedían nuevas reformas para restringir aún más los beneficios.

Incluso hubo críticos que sugirieron que la evaluación global del éxito desestimó evidencias de que la reforma al bienestar social tuvo consecuencias negativas para algunos pobres. Por ejemplo, en 2003, el Fondo de Defensa Infantil (CDF, por su sigla en inglés) destacó que si bien la proporción de niños negros que vivían en familias oficialmente definidas como pobres cayó de manera notoria desde 1995, la cantidad de niños negros criados en la “extrema pobreza” (que el CDF definía como hogares con ingresos que no superaban la mitad de la línea de pobreza federal) aumentó abruptamente entre 2000 y 2001¹⁷. El informe generó algunas páginas editoriales empáticas con la difícil situación de “los más pobres entre los pobres”, así como algunas acusaciones de “selección cuidadosa”. (Un conservador afirmó que el CDF había “buscado con lupa algo negativo para decir, porque desde la reforma al bienestar social de 1996, la fotografía de la pobreza en EEUU era indudablemente positiva”). Es innecesario suponer que un cambio

¹⁷ Sam Dillon, “Report Finds Deep Poverty Is on the Rise”, *New York Times*, April 30, 2003, p. A18.

de políticas como el que implicó esta reforma al bienestar social, tendría el mismo efecto en todas las personas. Lo más probable es que los problemas complejos no tengan soluciones sencillas, como tampoco es probable que sus causas sean sencillas.

Políticas sociales como la PRWORA son objetos contundentes diseñados para tratar problemas que suelen implicar complejos cruzamientos de causas y consecuencias. Cuando las personas discrepan sobre si acaso es deseable o no introducir una determinada política, menos probable es llegar a un acuerdo sobre sus efectos. Una vez más, cobra relevancia la noción de compensaciones: cada política tendrá costos y beneficios y es probable que sus defensores celebren sus beneficios, aunque sus oponentes denuncien abiertamente los costos. Esta complejidad debiera hacernos tener cautela con las evaluaciones simplistas de políticas que auguran éxitos o fracasos, como aquellas basadas en una o dos cifras. La complejidad no puede resumirse en un par de cifras mágicas seleccionadas con cuidado, y debiéramos cuestionar las afirmaciones que hacen que las cosas parezcan simples.

DISPUTAS EN TORNO A LAS CIFRAS O PRESUPOSICIONES CUESTIONABLES

Los estadounidenses tienen una fe ingenua, pero muy difundida, respecto al poder de los números para resolver todo tipo de discusiones y entregar datos que puedan vencer posturas contrarias. Esta fe se basa en algunas suposiciones dudosas. La primera es la creencia de que las cifras son objetivas por naturaleza, constituyen evidencias indiscutibles. Esto desconoce una verdad aún más básica: todas las cifras son fruto de esfuerzos humanos. No podemos ignorar el hecho de que las estadísticas son construcciones sociales.

Reconocer esto significa que no podemos considerar las cifras como meros trozos de verdad; más bien, debemos ser críticos, preguntar quién contabilizó qué, cómo y por qué lo hizo. Pero eso no significa

que no podamos confiar en ninguna estadística, que debemos partir de la base de que todas carecen de valor. Hay maneras mejores y peores de contar y podemos confiar más en ciertas cifras que en otras. No toda la ciencia es ciencia basura; con un poco de esfuerzo —y paciencia para esperar por más datos— es posible distinguir entre ambas.

Una segunda suposición dudosa es la creencia de que nuestras cifras son mejores que las de nuestros oponentes, simplemente porque son nuestras. Nuestros puntos de vista, posturas, preferencias e ideologías políticas moldean la forma en que abordamos las evidencias, incluyendo las estadísticas. Tenemos una tendencia natural a aceptar aquellos números que coinciden con lo que creemos es cierto. Precisamente porque esas cifras coinciden con nuestra visión del mundo, tendemos a minimizar —o no tener conciencia de— sus debilidades. Las leemos con simpatía y las aceptamos sin cuestionar. Por el contrario, al toparnos con cifras que cuestionan lo que creemos, se pone en acción nuestra capacidad crítica y hacemos preguntas difíciles: ¿qué llevó a mostrar cifras tan evidentemente erróneas?, ¿se debió a definiciones inapropiadas?, ¿se utilizaron métodos deficientes o muestras mal diseñadas?, ¿se hizo mal el análisis? No hay como una estadística incómoda para descubrir habilidades críticas que ni siquiera sabíamos que teníamos.

Este capítulo también muestra una tercera consideración: la vida social es compleja. Mientras algunos problemas sociales son objeto de una opinión bastante compartida y consensual (p.ej., es improbable que los abusos deshonestos contra niños y los femicidios tengan defensores), otros temas derivan en discusiones entre personas con distintas creencias, suposiciones, actitudes y valores. Es usual que los antagonistas revelen estadísticas que apoyen sus posturas y, por ende, opinen que sus cifras son más convincentes, aun cuando expresen profundas reservas en relación a las del oponente. Ya sea que participemos activamente en esas discusiones o nos mantengamos más al margen, puede ser útil considerar la posibilidad de la complejidad. Al contabilizar algo,

es necesario tomar decisiones que sin duda son moldeadas por muchos factores: ciertos métodos para contabilizar son más baratos que otros, algunos se adecuan mejor a quienes hacen los cálculos, otros pueden llevar con más seguridad a los resultados que ellos desean. Debemos esperar que las decisiones que tomen las personas moldeen sus resultados.

No debiéramos suponer que la mayoría de los temas sociales son fáciles de entender. Si fueran sencillos y comprensibles, no habría tantos desacuerdos. Cuando nos topamos con desavenencias respecto a la validez de una cifra o si escuchamos a sujetos que promueven números que se contradicen, debemos considerar la posibilidad de que, bajo esa discusión, haya algo bastante complejo. Más que intentar decidir qué lado tiene la razón, debemos tratar de conciliar las cifras en disputa para comprender cómo y por qué las personas tienen distintas posturas respecto a lo que es verdad. Desde luego que podemos elegir una posición, pero siempre teniendo presente que las cosas son un poco más complejas que eso.

Capítulo 7

ALFABETIZACIÓN ESTADÍSTICA

No es inusual que nos encontremos con malas estadísticas. Lo más probable es que en el matutino hayamos detectado por lo menos una cifra que nos hará desconfiar. Reconocer las malas estadísticas no es tan difícil: más que conocimientos de matemática avanzada, se requiere pensar con claridad. Es de suponer que la mayoría de las personas estarán de acuerdo en que las malas estadísticas debieran eliminarse.

Pero aun así, las cifras malas abundan. ¿Por qué ocurre esto? ¿No debiéramos ser capaces de “alfabetizar estadísticamente” a la población, vale decir, enseñar las habilidades básicas para la interpretación crítica de las estadísticas que enfrentamos día a día? ¿Por qué este tipo de enseñanza forma parte del currículo estándar de la educación media o universitaria? Deberíamos poder preparar a los jóvenes para que puedan pensar críticamente acerca de las cifras y estar inmunizados contra las malas estadísticas.

Cada año, miles de estudiantes que están terminando la enseñanza media se inscriben en cursos avanzados en alguna materia; en este caso, de estadística*. A fines de año, estos alumnos podrán rendir los exámenes nacionales de estadística y, si obtienen buenos puntajes,

* Estos cursos se denominan Advanced Placement (AP), cuya traducción literal puede ser “posicionamiento adelantado” (N. de la T.).

muchas universidades les reconocerán haber completado un curso básico de esta materia. Al mismo tiempo, miles de estudiantes cursarán en la universidad al menos un ramo de estadística. Es de suponer que la alfabetización estadística será parte importante de esos cursos.

Pero estaríamos equivocados. Tanto los textos de estadística como el examen de AP ignoran el tipo de problemas que hemos tratado en este libro. Más bien, en la educación media y universitaria, los cursos de estadística se centran en lo que yo denomino “asuntos de *cálculo*”: la teoría y lógica detrás de mediciones estadísticas concretas; los métodos para computarizar esas mediciones; y la interpretación de los resultados. Los textos introductorios de estadística incluyen capítulos sobre teoría de las probabilidades, pruebas de significación, correlación, regresión, etc. Estos textos de estudio suponen que los alumnos que los leen podrían querer interpretar información derivada de algún tipo de investigación científica. No hay nada de malo en esto; aquellos estudiantes que se dedicarán a la investigación tendrán que saber cómo calcular esas estadísticas.

Sin embargo, ni los libros ni los cursos se refieren a cómo interpretar las estadísticas sencillas —gráficos y números— que un estudiante podría encontrar en el periódico que llega a su casa todos los días. ¿Por qué sucede esto? Si todos están de acuerdo en que alfabetizar estadísticamente es una habilidad importante, ¿por qué la interpretación crítica de las estadísticas no es parte fundamental de la enseñanza de las estadísticas?

En general, se considera que la estadística es una rama de la matemática; de ahí que se centre en el cálculo estadístico. Estoy seguro que la mayoría de los colegios consideran que las clases de estadística AP son clases de matemática y, por ende, les asignan un profesor de esta asignatura. El objetivo del curso es hacer de los estudiantes expertos calculistas en estadística; las clases no están pensadas para alfabetizarlos estadísticamente. A nadie sorprende que los profesores de matemática crean que su labor es enseñar matemática e instruir a los alumnos

para que les vaya bien en las pruebas de cálculo, tal como el examen de estadística AP.

Asimismo, los cursos de estadística dictados en la universidad —incluso los más básicos e introductorios— se centran en temas de cálculo. La masificación de los computadores y softwares estadísticos de fácil manejo ha impulsado la utilización de estadísticas muy sofisticadas. Es probable que, antes de 1970, una persona sin formación en estadística avanzada comprendiera la mayoría de los artículos de una revista como *American Sociological Review* (ASR). Pero ahora ya no es así. Hoy en día, los artículos de la ASR incluyen regresiones por estimación de mínimos cuadrados, logarítmica lineal y otras complejas técnicas estadísticas multivariadas que no pueden ser entendidas por alguien que no haya cursado al menos dos semestres de estadística de nivel universitario. Naturalmente, los profesores universitarios creen que su labor es enseñar a los alumnos a dominar estas técnicas avanzadas.

¿Será posible que para la mayoría de los autores y profesores de estadística, los temas que he planteado en este libro sean demasiado sencillos como para hacer siquiera un comentario? Quizás. Pero, más que eso, los temas que hemos considerado no tratan de asuntos de cálculo. Nos preocupa más el proceso social involucrado que los procedimientos matemáticos (cálculos). Nuestro foco se ha centrado en quién, cómo y por qué contabiliza y genera las cifras. También nos interesa saber cuáles audiencias consumen y utilizan esos datos. Más que los cálculos para llegar a las cifras, hemos intentado comprender su construcción social.

Pero, en general, las clases de estadística ignoran las formas en que se utilizan los datos como evidencias para comprender temas sociales, y las formas en que las personas contabilizan. Si es que se menciona el proceso social mediante el cual se construyen las estadísticas, lo más probable es que se relacione con la idea del sesgo; los profesores podrían advertir a los alumnos de que individuos “sesgados” pueden

generar estadísticas distorsionadas. Más allá de culpar al sesgo —considerado un tipo de contaminación generada fuera del mundo matemáticamente prístino del cálculo—, las clases de estadística rara vez exploran lo que esta distorsión podría implicar.

En resumen, aunque todos coincidan en que sería deseable que los estudiantes mejoraran su habilidad para pensar críticamente acerca de las estadísticas que aparecen en los medios, los profesores de esta materia no consideran que ésta sea parte de su labor.

¿QUIÉN DEBE ALFABETIZAR ESTADÍSTICAMENTE?

Hoy los educadores están acorralados por las exigencias. Por un lado, en la medida que surgen nuevos temas de interés público, suele haber llamados para agregar material al ya sobrecargado currículo de la enseñanza: entre otros ejemplos obvios, están la educación sexual y la educación preventiva en drogas. Un municipio puede ganarse una beca gracias a un programa contra el maltrato entre escolares. Pueden desarrollarse campañas a nivel regional o municipal para que los alumnos tengan mayor conciencia de los distintos tipos de discriminación. La lista sigue y sigue, y cambia cada año. Algunos de estos temas perduran en los currículos de los colegios, pero otros son flores de un día.

Por otro lado, muchos se quejan de que los colegios están olvidando lo esencial: las Tres Erres^{**}. Al menos en parte, el movimiento que obliga a las escuelas a rendir cuentas (*accountability movement*) exige que éstas vuelvan a centrarse en las destrezas básicas. Entonces, colegios y docentes se ven atrapados entre los llamados a dedicar más tiempo a enseñar las habilidades básicas, y la presión de añadir instrucción relativa al tema de turno que atrae la atención pública. La

^{**} Abreviatura que se refiere a las destrezas clave del currículo de enseñanza básica: *Reading* (lectura), *Writing* (escritura) y *Arithmetic* (aritmética), palabras que tienen un sonido de R fuerte al comienzo (N. de la T.).

jornada escolar tiene un número limitado de minutos y hay muchísimas personas que desean que se dedique más tiempo a otros temas que les parecen relevantes.

Una primera pregunta podría ser si la alfabetización estadística debiera considerarse una habilidad básica o un tema adicional específico. Si es promovida como un tema especial —como la educación sobre el sida o la prevención del maltrato entre escolares—, su perspectiva a largo plazo no será muy alentadora. El que sea considerado un tema especial este año, lo vuelve candidato fácil a ser eliminado el próximo, cuando aparezcan nuevas temáticas.

¿Y si decimos que la alfabetización estadística es una habilidad básica? De hecho, existe un argumento bastante atendible para considerarla como tal. Después de todo, estamos hablando de enseñar a las personas a ser más críticas, a pensar un poco más respecto a lo que leen en los periódicos o ven en las noticias por televisión, a hacerse preguntas sobre los planteamientos de científicos, políticos o activistas. Estar más capacitados para evaluar esas afirmaciones es muy valioso; incluso, podríamos argumentar que es fundamental para ser un ciudadano informado. ¿Por qué no considerar la alfabetización estadística como una habilidad básica?

Pero esto genera otra pregunta: ¿qué tipo de habilidad básica es? La respuesta es importante, porque en los colegios y universidades, la responsabilidad de enseñar ciertos temas es distribuida en varios departamentos. Una universidad cuenta con departamentos de ciencias, inglés, matemática, ciencias sociales, etc.; la mayoría de las universidades subdividen estas grandes categorías, por ejemplo, asignando a distintos departamentos la responsabilidad de enseñar biología, química y otros. En general, mientras más grande sea el establecimiento educacional, más departamentos tiene.

Los departamentos rivalizan entre sí. Si bien todos reconocen el valor de una educación completa, cada departamento tiende a asumir que el suyo es un rol especialmente importante. Y ya que el dinero

siempre escasea, compiten por los fondos disponibles para contratar docentes y comprar equipos. Es raro que un departamento no desee expandirse; en particular, muchos quisieran ofrecer una instrucción más avanzada, como los cursos AP en los colegios o los programas para graduados en las universidades.

Esta competencia significa que enseñar habilidades básicas es algo poco valorado. Por ejemplo, la mayoría de los estudiantes universitarios de primer año deben tomar un curso de redacción en inglés. Éstos deben ser cursos de pocos alumnos, pues tienen que escribir muchos ensayos, los que han de ser calificados rápida y cuidadosamente. En la mayoría de las universidades, estos cursos de redacción son dictados por estudiantes graduados o docentes de jornada parcial, no por profesores de inglés. En parte, es mucho más barato enseñar composición de esa forma; por otro lado, los profesores de inglés prefieren dar cursos de especialización (porque la materia y los alumnos son más interesantes). El punto es que enseñar esta habilidad básica no es considerada particularmente gratificante. (Los departamentos de inglés de algunas universidades han derivado esta incómoda tarea a departamentos específicos dedicados a la composición, centros de escritura u otros programas).

El ejemplo de la redacción en inglés puede ayudarnos a apreciar los problemas de la alfabetización estadística. Los profesores universitarios están conscientes de que un número considerable de estudiantes tiene dificultades para leer gráficos y tablas elementales; ni hablar de pensar críticamente acerca de ellos. Esta es una habilidad básica muy importante, pues tanto las tablas como los gráficos aparecen en muchas de las lecturas que los alumnos deberán realizar durante sus estudios universitarios. Así y todo, nadie quiere enseñarla, o al menos destinar demasiado tiempo a hacerlo. Muchos opinan que los estudiantes debieran haber adquirido esta habilidad antes de entrar a la universidad (aun cuando es evidente que no es así). Otros consideran que este asunto es demasiado simple y básico, casi una pérdida de tiempo

para los docentes que preferirían enseñar aspectos más avanzados de sus respectivas disciplinas.

Además, la amplísima difusión de los computadores personales y softwares sofisticados confirma la ilusión de que los estudiantes entienden cabalmente este asunto. Cualquiera que asista a una feria científica de alumnos de enseñanza media notará la gran cantidad de atractivos gráficos creados por medio de computadores. Mientras nadie se dé el trabajo de preguntar si esos gráficos son claros y/o útiles (en general, no son ni lo uno ni lo otro), es fácil impresionarse con el trabajo de los estudiantes. Asimismo, los alumnos aprenden que pueden encontrar casi cualquier información en internet. Es probable que no hallen buenas respuestas a sus preguntas, pero descubrirán respuestas. La experiencia de los estudiantes en la utilización de métodos de tecnología avanzada (aunque generen resultados de baja calidad) suma puntos a la justificación de que es innecesario enseñar habilidades básicas, que podemos avanzar hacia material más complejo e interesante.

Así las cosas, los profesores de estadística y matemática no se interesan en alfabetizar estadísticamente más que los profesores de inglés en enseñar redacción a los alumnos de primer año. Tampoco los otros departamentos se interesan en ello. Dicto cursos de sociología, pero sé que la mayoría de mis colegas tienden a descalificar la alfabetización estadística y no creen que sea “verdadera sociología”; es probable que los docentes de psicología y otras disciplinas piensen lo mismo. La alfabetización estadística no cabe en ninguna de las definiciones que los departamentos académicos se han asignado.

Mi pesimismo tiene precedentes. En las universidades, a fines de la década de los 80 e inicios de la de los 90, el “pensamiento crítico” se convirtió en un término de moda. Debí ser la consigna perfecta en torno a la cual congregar el apoyo a la reforma educacional. Prácticamente todos los profesores se consideran a sí mismos como pensadores críticos y muchos concordarán en que los alumnos deben aprender a pensar de manera más crítica: otra habilidad básica muy

apetecida. Pero debido a que todos esos docentes creían estar enseñando a pensar críticamente (aunque al mismo tiempo se quejaban de que muchos alumnos no lo conseguían), y dado que ningún departamento quería asumir la responsabilidad de enseñar transversalmente el tema en la facultad, el interés en este tópico llegó a su valor máximo y la fuerza de la idea en tanto lema educacional comenzó a desvanecerse.

¿Qué pasó con el pensamiento crítico? ¿Por qué una idea tan buena no se volvió parte fundamental de la educación en todos los colegios? El hecho de que no hubiera un “dueño”, algún departamento que acogiera, protegiera y alimentara el pensamiento crítico, significaba que esa habilidad era responsabilidad de todos, lo que equivalía a decir que nadie se responsabilizaría.

Este ejemplo sugiere que debiera existir un departamento específico que asuma la responsabilidad de alfabetizar estadísticamente. Como ya mencionamos, no es necesario que este asunto recaiga en un programa de matemática o estadística, por muy lógico que esto parezca a primera vista. Las ciencias sociales podrían ser un lugar apropiado. A fin de cuentas, los temas relativos a la alfabetización estadística con frecuencia surgen en torno a discusiones sobre temas sociales. Pero nos volvemos a encontrar con los planteamientos de los profesores de sociología que dicen que este asunto no es “verdadera sociología” (y es probable que otros departamentos reaccionen de la misma forma).

La organización en departamentos ofrece grandes ventajas a las instituciones educacionales, pero también tiene sus costos. Es difícil enseñar temas que no calzan a la perfección con el ámbito que un departamento considera propio de su dominio educacional. Por eso muchos graduados de educación media y universitaria se sienten incómodos cuando deben enfrentar datos estadísticos básicos; y también por eso es probable que esta situación no cambie fácilmente. Las clases de alfabetización estadística no son tan terribles; lo más complejo es encontrar a alguien que quiera dictarlas.

MOVIMIENTO POR LA ALFABETIZACIÓN ESTADÍSTICA

A pesar de estos obstáculos, ha surgido un pequeño grupo que aboga por la alfabetización estadística. Su líder es el profesor Milo Schield, director del Proyecto de Alfabetización Estadística de la Fundación W.M. Kleck en el Augsburg College de Minneapolis. Él está a cargo del sitio Statistical Literacy (www.statlit.org); para quienes se interesen en este tema en cuanto movimiento educacional, el sitio incluye una sección sobre enseñanza. Si bien éste es un avance promisorio, la campaña para promover la educación formal de la alfabetización estadística recién está comenzando.

Pero quizás este tema no deba enseñarse en las salas de clase. Han surgido varios llamados a alfabetizar estadísticamente fuera de los establecimientos educacionales. Consideremos estos recursos a modo de ejemplo:

- El Statistical Assessment Service (STATS) (www.stats.org) ha criticado el manejo periodístico de las estadísticas desde 1995. STATS publicó boletines de prensa hasta 2002, cuando comenzó a distribuir sus informes a través de su sitio. Allí existe un libro (también impreso) basado en análisis de STATS; véase David Murray, Joel Schwartz y S. Robert Lichter, *It Ain't Necessarily So: How Media Make and Unmake the Scientific Picture of Reality* (2001).
- Hay muchos sitios web de otros países que discuten en torno a las malas estadísticas. Algunos contienen sólo material original; otros se refieren a discusiones específicas. Numberwatch (www.numberwatch.co.uk) es un sitio británico; John Brignell, quien opera el sitio, es autor de *Sorry, Wrong Number! The Abuse of Measurement* (2000). El Social Issues Research Centre (www.sirc.org) es otro sitio británico que entrega análisis de temas que critican datos estadísticos. Pénombre es un sitio francés (www.pénombre.org) que

contiene material en inglés. El Canadian Statistical Assessment Service (www.canstats.org) es similar a su contraparte estadounidense, en tanto que otro sitio canadiense, Innumeracy.com (www.innumeracy.com), es básicamente un catálogo de enlaces. El sitio australiano Numeracy in the News, orientado a educadores y estudiantes, presenta artículos sobre muestreo, gráficos, etc., cada uno acompañado de preguntas y comentarios (<http://www.mercurynie.com.au/mathguys/mercury.htm>). Muchas de estas organizaciones también ofrecen enlaces a sitios más especializados, incluyendo estadísticas oficiales (numerosas agencias gubernamentales han creado sitios donde es posible acceder a sus informes estadísticos), y sitios que se dedican a ciertos temas sociales o tipos de información. Por ejemplo, Quackwatch (www.quackwatch.org) se refiere a temas médicos, Junkscience.com (www.junkscience.com) trata la cobertura periodística de noticias científicas, y el Center for Media and Democracy (www.prwatch.org) investiga y critica la manipulación de datos y las campañas de relaciones públicas de la industria y el gobierno. Como es de esperar, estos sitios difieren en sus preocupaciones e ideologías políticas, y por lo tanto, antes de aceptarlas, hay que revisar cuidadosamente sus críticas.

- En general, es bastante entretenido explorar las malas estadísticas, pero la columna de Cecil Adams, “The Straight Dope”, que aparece en semanarios alternativos, es insuperable. Su lema es: “Luchando contra la ignorancia desde 1973 (ha tardado más de lo que pensábamos)”. Cada semana, Adams aborda una o más preguntas, usualmente sobre temas que otros medios ignoran, incluyendo afirmaciones estadísticas. El sitio www.straightdope.com entrega un índice y acceso a todas las columnas. Si están interesados en temas extraños e insólitos, éste es un muy buen recurso.
- Otros comentaristas de medios también promueven la alfabetización estadística. El matemático John Allen Paulos, autor de *Innumeracy: Mathematical Illiteracy and Its Consequences* (2001) (*El*

Hombre Anumérico [2002]) y otros libros para lectores no especializados, tiene un sitio (<http://euclid.math.temple.edu/~paulos/>) que enlaza con varios trabajos suyos, incluyendo las columnas para ABC.com. La BBC tiene varios programas radiales sobre temas matemáticos, incluyendo “More or Less”, donde se analizan comentarios comunes en relación con temas estadísticos. Los programas están disponibles en http://news.bbc.co.uk/2/hi/programmes/more_or_less/default.stm

- La Asociación Estadística Estadounidense publica *Chance*, revista trimestral dedicada a diversos usos de la estadística. Algunos artículos requieren formación en estadística, pero hay otros más accesibles. Es un recurso muy útil como introducción a lo que hacen los estadísticos profesionales.
- Existen muchos libros sobre temas estadísticos, desde textos que enseñan a los estudiantes a hacer cálculos hasta otros, como éste, que se dedican a criticar el uso y abuso de las estadísticas en la sociedad contemporánea. (Varios de estos libros se mencionan en las notas al pie de los capítulos anteriores).

Estas fuentes conforman un coro de voces que promueven la alfabetización estadística. Por supuesto, existen discrepancias en el movimiento. Algunos tienen agendas ideológicas: los conservadores se centran en exponer el mal uso de las estadísticas por parte de los progresistas, y estos últimos atacan las dudosas cifras propuestas por los primeros. Hay críticos que culpan a “los medios” de publicitar irresponsablemente malas estadísticas, pero los periodistas —con razón— responden que a menudo no cuentan con buenas herramientas para evaluar las cifras entregadas por sus fuentes. Algunos estadísticos plantean la necesidad de una mejor capacitación en matemática para mejorar la comprensión de los cálculos, mientras que otros científicos sociales (como yo) discuten que es más importante contextualizar socialmente las cifras, de modo de comprender cómo son creadas y dadas a conocer.

En resumen, puede que sea cierto que “todos” están de acuerdo en que se debe mejorar la alfabetización estadística, pero no está tan claro qué entendemos por ello, qué significa mejorarla o cuáles serían las consecuencias de hacerlo.

Aunque nadie se oponga a la alfabetización estadística, los obstáculos permanecen. No hay acuerdo en torno a qué habilidades deben enseñarse y, hasta el momento, ningún grupo ha asumido la responsabilidad de diseñar la instrucción necesaria. Hay mucha información —cualquier persona interesada puede aprender formas de pensar críticamente sobre las estadísticas—, pero el movimiento por la alfabetización estadística todavía tiene que convencer a la mayoría de los profesores de que es necesario cambiar lo que el sistema educacional está haciendo.

Muchos de nosotros bromeamos con que las malas estadísticas provienen de sujetos con quienes no estamos de acuerdo, y fantaseamos que al mejorar la alfabetización estadística aumentará la cantidad de gente que concuerde con nosotros, que todos los pensadores críticos reconocerán las falencias en los argumentos de nuestros oponentes, en tanto los nuestros serán considerados más convincentes.

Yo no daría por sentado que las cosas van por ese camino. La alfabetización estadística es una herramienta y, como muchas herramientas, puede utilizarse para múltiples propósitos. Si más personas piensan críticamente las estadísticas, podrán hacerlo en relación con nuestras ideas, pero también con las de nuestros contrincantes. Cuando las cifras de todos son observadas con atención, la exigencia aumenta para todos.

Pero eso no es malo. Así como están las cosas, constantemente nos vemos expuestos a una gran cantidad de estadísticas. Algunas de esas cifras son buenas, pero muchas no lo son. Como resultado, nos preocupamos de cosas que no ameritan tanta atención e ignoramos otras que sí debiéramos considerar. Mejorar la alfabetización estadística —si logramos hacerlo— podría ayudarnos a distinguirlas y, quizás, llegar a ser más sabios.

ÍNDICE ANALÍTICO

- Absolutismo, 154, 156
 Actividad sexual de estudiantes de enseñanza media, 53-54
 Admisiones universitarias, 124-30
 Adolescentes: suicidios entre, 106-12 (*véase también* Maltrato escolar entre pares); tasas de graduación de, 136-38; y actividad sexual, 53-54; y crimen, 2, 63; y embarazo, 57-60, 63; y enseñanza de la estadística, 179-87; y uso de drogas, 165-66
 Afiliación religiosa, 14-15. *Véase también* Judíos; Musulmanes
 Afroamericanos: muertes accidentales de, 109-12; suicidios entre, 106-12; y datos de censo, 15-16; y perfil racial, 143-49; y pobreza, 175
 Agua, ingesta recomendada, 22
 Alcohol y embarazo, 55-57
 Alfabetización estadística, 179-90; movimiento por la, 187-90
 Análisis de costo-beneficio, 10-13, 88-89
 Ataque de francotirador, riesgo de, 84
 Ataques al World Trade Center: contabilizar muertes por, 112-15; huérfanos debido a, 114
 Baleos por policías, 17
 Basura gráfica, 50-51, 52
 Bienestar social, reforma al, 171-76
Bullying (maltrato escolar entre pares), 96, 98-103
 Cálculo, 179-82
 Calentamiento global, 74
 Cáncer: mamario, 78-79; pulmonar, 40-41, 76-77, 82-83, 162-63
 Categorías étnicas/raciales, 15-16
 Causalidad, 38-43
 Ciencia basura, 157, 159-63
 Ciencia, fraude en, 94, 98.
Véase también Ciencia basura; Investigaciones científicas
 Cifras: acreditadas, 93-119; controversiales, 151-78; de

- organizaciones, 130-36 (*véase también* Medidas burocráticas); engañosas, 27-61; incalculables, 8-14; legendarias, 20-25; mágicas, 121-50; oficiales, 103-17 (*véase también* Medidas burocráticas); que asustan, 63-91; que brillan por su ausencia, 1-26; que no se contabilizan, 14-19
- Clase social, 140-41
- Cocos, muertes por la caída de, 21, 23
- Comparaciones, 88-89
- Compensaciones, 163, 176
- Comunicados de prensa, 97, 103, 104, 159
- Construcción social, xv, 153, 176, 181
- Consumo de alcohol en campus universitarios, 18
- Contadores, 117-19
- Correlación, 38-43
- Crecimiento demográfico, 73-74
- Créditos para estudios, 51-52
- Crímenes, olas de, 5, 8
- Datos de censo, 14-16, 163-64
- Definiciones amplias, 9, 170
- Denuncias, 9-10
- Descuido de la probabilidad, 84
- Desempleo, 16-17, 174
- Disminuir impuestos, 27-28, 31
- Divorcio, 84-88
- Efectos colaterales, 13, 88-89
- Ejemplos, 4-8, 69-70
- Embarazo y alcohol, 55-57
- Enseñanza de la estadística, xv-xvi, 179-82
- Errores médicos, 68-70
- Escándalos estadísticos, 117-18
- Escasez de viviendas, 6
- Escenarios apocalípticos, 66-67, 73-75, 90
- Espurio, 39-41, 43, 80-81
- Estadística, enseñanza de la, xv-xvi, 179-82
- Estadounidenses, antepasados de, 15-16
- Estética y gráficos, 45, 47, 49, 50-57
- Estimar población: de judíos, 167, 169-71; de musulmanes, 167-71
- Estudiantes de enseñanza media, actividad sexual de, 53-54
- Evidencias, 154-58
- Exámenes/pruebas/evaluaciones, 24, 126, 136-43
- Expectativa de vida, 19-20
- Experimentos, 79-81
- Externalidades, 13, 88-89
- FBI (Federal Bureau of Investigation), 2, 3, 8, 10, 85, 134, 144
- Financiamiento para investigaciones, 18, 97-98
- Fumar, 40-41, 75-77, 82-83, 162
- Gráficos, 43-61
- Guerras estadísticas, 152
- Guías sobre universidades, 18, 124-30
- Hechos, 153-54
- Ideología, xiii-xiv, 105, 152, 162, 164, 165, 166, 177, 188

- Implantes mamarios, 160-61
- Índice de Precios al Consumidor (IPC), 116
- Información de censo, 14-16, 163-64
- Informes corporativos, 133-34
- Ingresar a la universidad, 124-30
- Investigaciones: científicas, 96-105, 155-59; financiamiento para, 18, 97-98
- IPC (Índice de Precios al Consumidor), 116
- Journal of the American Medical Association (JAMA)*, 96-103
- Joyas, estadísticas como, xiv-xv
- Judíos, estimar población de, 167, 169-71
- Jugar con números, 130-36, 149-50
- Maltrato escolar entre pares (*bullying*), 96, 98-103
- Mamografías, 78-79
- Manipulación (*spinning*), 163-67
- Marihuana, 41-42, 165-66
- Media/promedio, 30-31
- Mediana, 30-31
- Mediciones burocráticas, 23-25. *Véase también* Cifras de organizaciones; Cifras oficiales
- Medios de comunicación, noticias, xviii, 4-6, 64, 74-75, 81-82, 97-98, 103-4, 159-60
- Metanálisis, 83
- Mitos urbanos, 5-6. *Véase también* Cifras legendarias
- Movimiento por la alfabetización estadística, 187-90
- Muertes: accidentales entre jóvenes afroamericanos, 109-12; en colegios, 2-4; por ataques de francotiradores, 84; por caídas de cocos, 21, 23; por errores médicos, 68-70; por sarampión, 19; por suicidio, 106-12
- Musulmanes, estimar población de, 167-71
- Niños desaparecidos/raptados/secuestrados, 8-10, 47-48
- Números: jugar con, 130-36, 149-50; olvidados, 19-20
- Olas de crímenes, 5, 8
- Organización en departamentos, efectos en la educación, 183-86
- Pensamiento crítico, 184-86
- Perfil racial, 143-49
- Pesimismo, 89-90
- Porcentajes, 32-38; calcular incorrectamente, 33-36; como medidas de cambio, 36-37, 37n6
- Prensa, comunicados de, 97, 103, 104, 159
- Probabilidad, 76-79; descuido de la, 84
- Problemas sociales, 64-67, 123; tamaño de, 65, 67-70
- Profesores, evaluar calidad de, 23-25, 126, 139-41
- Progreso, 20
- Promedio/media, 30-31
- Promedios, 28-31
- Prueba de Aptitud Académica (SAT), 126, 137
- Pruebas/exámenes/evaluaciones, 24, 126, 136-43

- Psiquiatras, 107
 Puntajes en pruebas escolares,
 136-43; efecto del tamaño del
 colegio en, 142

 Raptos por familiares, 47-49
 Rebajar impuestos, 27-28, 31
 Recesión, 120-22
 Recuentos, 29
 Reforma al bienestar social, 171-76
 Registro de datos, 105-17, 117, 118
 Registros de muertes, 105-6, 114-15
 Relativismo, 154-56, 158
 Replicación, 82-83, 103, 117, 158-59
 Revisión por pares, 93-94
 Revistas académicas, 93-95,
 97-98. Véase también *Journal of the
 American Medical Association*
 Riesgos, 65-66, 75-88; medir, 79-84
 Rubios, extinción de, 21

 Sarampión, 19
 Secuestros: de niños (*véase* Niños
 desaparecidos); por familiares,
 47-49
 Selección cuidadosa (*cherry-picking*),
 163-65, 167, 175-76

 Selectividad, 57-60
 Sesgo, xvi-xvii, 181-82; editorial, 83
 Sexo, pensar sobre, 54-55
 Software, consecuencias del, 43, 47,
 49-50, 53, 57, 106, 181, 185
 Subsidios, fraude en, 7, 7n6
 Suicidios entre jóvenes
 afroamericanos, 106-12

 Tasas: de natalidad, 57-60;
 delictuales, 2, 71-72, 134
 Tendencias, 60, 66, 70-73
 Testigos expertos, 159
 Tiroteos en escuelas, 1-5
 Transporte público, 11-12
 Truncar gráficos, 45-46, 49

U.S. News & World Report. Véase
 Guías sobre universidades
 Uso de drogas, 41-42, 165

 Vacunas, 13, 19
 Ventajas y desventajas, 135
 Vida humana, valor de la, 12-13
 Viudez, edad promedio de, 21-22

 Y2K, 66, 73