

elementos comunes y ubicar el *núcleo* del término. Siguiendo esta línea es posible ver el fenómeno de forma más integral, pues el *núcleo* puede servir de hilo conductor en una agenda de investigación en la que sobresalga la propuesta de estudios de la calidad de los gobiernos.

En materia de políticas públicas, la obra contribuye al debate de los esfuerzos anticorrupción y ayuda a entender el fenómeno. Pese a la naturaleza del término corrupción como un “concepto paraguas”, se exhorta a no hacer uso indiscriminado del mismo, algo que poco contribuiría a entender el problema y a diseñar las políticas anticorrupción adecuadas. En términos generales, la obra es ampliamente recomendable para quienes tienen interés en comprender el sentido de la corrupción.

■.....

Jack Schneider, *Beyond Test Scores: A Better Way to Measure School Quality*, Cambridge: Harvard University Press, 2017, 326 pp.
doi: <http://dx.doi.org/10.29265/gypv.v28i2.633>

Por Darwin López, estudiante del Doctorado en Políticas Públicas, CIDE

El sector educativo suele estar entre los de mayor gasto del sector público. Representa 4.9 por ciento del producto interno bruto (PIB) mundial (Banco Mundial, 2018a) y supera al sector defensa en más de 2 puntos porcentuales (Banco Mundial, 2018b), ya que su participación es de casi 25 por ciento del gasto público total (Banco Mundial, 2018c). Es conocido por su incuestionable importancia en el desarrollo económico y social de los países (Vásquez *et al.*, 2013: 255). A nivel internacional existen varias pruebas que se utilizan de manera generalizada para medir el desempeño estudiantil y, por continuidad, del propio sistema educativo. Por ejemplo, el Programa Internacional para la Evaluación de Estudiantes (PISA, por sus siglas en inglés), diseñado por la Organización para la Cooperación y el Desarrollo Económicos (OCDE); Tendencias Internacionales en el Estudio de las Matemáticas y las Ciencias (TIMSS, por sus siglas en inglés) y Estudio Internacional del Progreso en Comprensión Lectora (PIRLS, por sus siglas en inglés), realizados por el International Study Center, Lynch School of Education, Boston College y la Asociación Internacional para la Evaluación del Desempeño Educativo; o el Tercer Estudio Regional Comparativo y Explicativo (TERCE), propuesto por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). De éstas, PISA es el

referente con respecto al funcionamiento de los sistemas escolares del mundo y ha servido para explicar intervenciones ya realizadas, como en el caso de Finlandia (Chung, 2016), Singapur (Deng y Gopinathan, 2016) y Hong Kong (Sui Chu Ho, 2016), o para justificar políticamente las reformas en otros países, como España (Choi y Jerrim, 2016) e Inglaterra (Adamson *et al.*, 2017).

Estados Unidos es el país con el sistema de evaluación de la educación más antiguo del mundo y además es el que gasta más en medición, pero, ¿qué relación hay entre la medición y el desempeño real de los estudiantes? ¿Hay manera de medir objetiva y completamente los resultados escolares? ¿Los exámenes estandarizados reflejan lo que se necesita saber del sistema educativo para mejorar lo que se necesita? Éstas son algunas de las preguntas que trata Jack Schneider en su libro. En seis capítulos explica su posición respecto de por qué el sistema es menos responsivo de lo que debería y, por lo tanto, es necesario diseñar una mejor manera de medir la calidad educativa.

Ante la idea equivocada de usar exámenes estandarizados, el autor presenta sus limitaciones. La aceptación de los exámenes parte de las premisas de que son más eficientes que cualquier otra forma de recolección de información; dan la apariencia de objetividad, porque no dependen de las circunstancias específicas de los individuos a los que se evalúa, y asumen que el conocimiento puede medirse de manera unívoca. Esas premisas no están necesariamente equivocadas. El problema es que son demasiado simplistas para cumplir su propósito de medir eficazmente el desempeño escolar, sobre todo porque el aprendizaje es en realidad un proceso complejo que depende de muchas variables y no puede separarse de su contexto —no si se pretende medir para hacer política pública basada en evidencias y no en intuiciones—. Algunas de las variables que intervienen en la posibilidad de respuesta de los estudiantes son el contexto familiar, que es de hecho el mayor predictor de desempeño escolar, el cual no está controlado por la escuela; la exposición lingüística, de la cual depende en buena medida la amplitud del vocabulario y la subsecuente comprensión lectora; la interpretación e intención del diseñador del examen, y la mera diversidad de los individuos. Ningún examen puede en la práctica cumplir con esas consideraciones sin volverse tan complejo como la realidad misma y eso le quitaría su ventaja principal, la eficiencia. Pero, ¿si los exámenes estandarizados son tan poco efectivos por qué se usan de manera tan generalizada? Schneider argumenta que es simplemente porque la sociedad está tan expuesta que ya le parece normal (p. 35).

Si evaluar el desempeño escolar es como ver a través de un vidrio oscuro, ¿cómo miden los padres y hacedores de políticas la calidad educativa? Schneider respon-

de: reduciendo la complejidad a algo simple, un número. Medir lo que realmente sucede en la escuela y no sólo un resultado aislado, requeriría semanas de trabajo para recolectar y procesar la información por parte de los investigadores y un tiempo proporcional para que los interesados la analizaran, convirtiéndose en un proceso económicamente inviable. El problema radica en que en la simplificación se pierde información valiosa. Así, los resultados deberían interpretarse de manera acotada, entendiendo que no es posible diseñar a la vez un instrumento sencillo, que sea válido para todos y en todos los aspectos. Padres y políticos toman este número vaciado de contenido y lo convierten en un indicador de la calidad escolar. Estos instrumentos presentan a lo sumo una fotografía de baja resolución y distorsionada de la realidad (p. 51).

En este punto puede afirmarse que lo que realmente importa es un nuevo marco de calidad educativa, que supere el reduccionismo de los exámenes estandarizados y su número único. Como tampoco es posible tratar la extensión de la realidad, ni identificar *a priori* las diferentes configuraciones ponderadas de cada actor, es necesario identificar unos elementos básicos que den cuenta de la calidad desde una perspectiva más equilibrada. ¿Cuáles serían esos elementos? Schneider propone un esquema que considera los insumos y no sólo los resultados. Argumenta que, ante la diversidad de contextos e intereses, un sistema con condiciones similares podría tener resultados también similares —a la alta—, además de obtener ganancias sociales aun con resultados académicos diferentes. Profesores profesionales, cultura escolar coherente, recursos escolares adecuados darán como resultado probable aprendizaje académico, desarrollo del carácter y bienestar de los estudiantes. Éste es un punto de partida. Para mejorar los resultados es importante refinar el modelo y aumentar el conocimiento sobre el ecosistema escolar (p. 148).

Desde esta perspectiva, entender el ecosistema es un objetivo necesario. Entonces ¿cómo obtenemos esa clase de información? El autor propone hacer uso de nuevas herramientas. Aparte de los exámenes tradicionales es necesario obtener, compilar y procesar otra clase de información, tanto de estudiantes, como de padres, profesores y directivos. Las encuestas masivas podrían ser una opción viable. Por suerte, la capacidad de cómputo y las tecnologías de la información y la comunicación permiten la distribución y procesamiento de estos instrumentos, también modificarlos y adecuarlos a cada población objetivo con costos mínimos. Así, sería posible corregir las distorsiones propias de este tipo de herramienta y hacerlo de manera continua para ver en tiempo real los cambios en la dinámica escolar. Además de las encuestas es viable recolectar toda clase de

información complementaria, hacer uso de las herramientas de la *big data*, y construir una imagen más precisa de lo que realmente sucede dentro de las escuelas (p. 187).

Esta capacidad tecnológica es como una superautopista de información; lo que sigue es hacer los datos usables. A pesar de que la tecnología computacional permite recoger y procesar la información, ésta no tiene significado por sí misma, es importante otorgárselo. Con el nuevo marco deben incluirse los mecanismos de interpretación. No es un mero asunto de tener más y mejor información, sino de que ésta tenga sentido, que promueva nuevas formas de entender la función de la escuela y el papel que desempeñan todos los sujetos. Es necesario informar y empoderar a los actores para que puedan hacer uso de estos datos. Al fin y al cabo, el objetivo es incluir a todos los interesados en todos los aspectos posibles para impulsar sus compromisos. El aporte más importante probablemente sea que la disponibilidad de información permitiría diseñar una mayor diversidad de modelos simultáneos, que experimenten, compartan información, monitoreen su progreso. En una frase: que se perfeccionen (pp. 219-220).

Para ser responsivos hay que tener una nueva rendición de cuentas. Los números deben importar. La recolección de datos tiene como finalidad ofrecer información para tomar decisiones; en el caso de los padres, para escoger escuelas según sus preferencias; a las empresas, para monitorear el perfil de los trabajadores, y a los hacedores de políticas, para definir mecanismos de intervención —los cuales suelen ser disruptivos cuando consideran que la distancia entre lo obtenido y lo esperado es mucha, y afirmativos cuando los resultados están alineados con las expectativas—. Es importante recalcar la necesidad de instrumentos variados que eviten la simulación, es decir el control de los mecanismos de medición, como puede ser que las escuelas enseñen para los exámenes nada más y olviden la complejidad del proceso educativo, o que estén midiendo más las circunstancias de los estudiantes que el aporte de la escuela. Con base en la disponibilidad de información se pueden corregir las tendencias y no sólo reaccionar al hecho. Esta rendición de cuentas sólo tiene sentido si los actores están comprometidos con la responsividad del sistema (p. 252).

Schneider ofrece cuatro conclusiones principales. Primero, que las escuelas de calidad no son recursos fijos, sino que es posible crearlas y hacerlas abundantes, para esto es necesario que todos se involucren y pongan a disposición recursos diversos. Segundo, que las decisiones que se tomen deben estar basadas en la existencia de información en cantidad y calidad. Tercero, que el objetivo de las escuelas no debe ser proveer ventajas artificiales, sino propiciar espacios de apren-

dizaje y bienestar. Por último, que si bien la educación tiene una clara función económica, no puede reducirse a eso; también debe promover el civismo, la cohesión, la justicia, la belleza, el bien común y la autorrealización.

Los exámenes estandarizados tradicionales tienen un espacio concreto de validez: sólo reflejan aspectos muy limitados de la acción escolar, por lo tanto, deben interpretarse acotados a esas limitaciones. Los números sencillos no pueden interpretarse como imágenes cabales de la calidad educativa; dicen qué tan bien responden los estudiantes a los exámenes, no por qué lo hacen. Son útiles para dar cuenta de una batería limitada de aspectos de manera eficiente, pero no pueden interpretarse como la panacea al respecto de la situación real de la educación. Si se consideran estas limitaciones, es posible aprovechar las ventajas de los exámenes sin reducir la realidad.

Es loable la posición del autor en su búsqueda de mecanismos más efectivos para medir el desempeño educativo en su complejidad. Esto implica incluir aspectos académicos diversos que superen la perspectiva matemática-lectora que prima en los modelos actuales. Es decir, incluir otros aspectos no académicos, como la seguridad e incluso la felicidad de los estudiantes en la escuela. El problema puede ser que es un modelo que no se ha probado cómo realizarlo. Es una aspiración que debe construirse, no una propuesta concreta. En ese sentido, no debe entenderse como una receta, sino como un llamado a imaginar alternativas. La dificultad es que no presenta casos donde todo eso se haya siquiera discutido, por lo que no es posible calcular los costos políticos o económicos. Tampoco se sabe qué clase de diseño organizacional se necesita o que marco interpretativo utilizar. Los exámenes parten de las simplificaciones propias de las metodologías cuantitativas que buscan generalizar, mientras que la propuesta plasmada en el libro requiere mecanismos cualitativos específicos, imposibles de extrapolar. Un sistema tan complejo como la realidad es inútil, uno tan simple como una cifra es vacío. ¿Cuál es el punto de equilibrio o cómo encontrarlo? El autor no lo dice.

El profesor Schneider invita a repensar el papel de los exámenes estandarizados como instrumentos únicos de medición del desempeño escolar. Propone complejizar la recolección de información con el uso de herramientas tecnológicas y el intercambio de experiencias sobre una base más diversificada. Es un aporte importante para los hacedores de políticas públicas en educación, en tanto su objetivo sea hacer escuelas más responsivas y sistemas más resilientes. En ese orden de ideas, es un libro recomendable para los estudiosos del tema, sean académicos o diseñadores de políticas.

REFERENCIAS

- Adamson, B., K. Forestier, P. Morris y C. Han (2017), “PISA, Policymaking and Political Pantomime: Education Policy Referencing between England and Hong Kong”, *Comparative Education*, 53(2), pp. 192-208, disponible en: <https://doi.org/10.1080/03050068.2017.1294666>
- Banco Mundial (2018a), *Gasto público en educación, total (porcentaje del PIB)*, disponible en: <https://datos.bancomundial.org/indicador/SE.XPD.TOTL.GD.ZS>
- Banco Mundial (2018b), *Gasto militar (porcentaje del PIB)*, disponible en: <https://datos.bancomundial.org/indicador/MS.MIL.XPND.GD.ZS>
- Banco Mundial (2018c), *Gasto (porcentaje del PIB)*, disponible en: <https://datos.bancomundial.org/indicador/GC.XPN.TOTL.GD.ZS?view=chart>
- Choi, Á. y Jerrim, J. (2016), “The Use (and Misuse) of PISA in Guiding Policy Reform: The Case of Spain”, *Comparative Education*, 52(2), pp. 230-245, disponible en: <https://doi.org/10.1080/03050068.2016.1142739>
- Chung, J. (2016), “The (Mis)Use of the Finnish Teacher Education Model: ‘Policy-Based Evidence-Making’? *Educational Research*, 58(2), pp. 207-219, disponible en: <https://doi.org/10.1080/00131881.2016.1167485>
- Deng, Z. y S. Gopinathan (2016), “PISA and High-Performing Education Systems: Explaining Singapore’s Education Success”, *Comparative Education*, 52(4), pp. 449-472, disponible en: <https://doi.org/10.1080/03050068.2016.1219535>
- Sui Chu Ho, E. (2016), “The Use of Large-Scale Assessment (PISA): Insights for Policy and Practice in the Case of Hong Kong”, *Research Papers in Education*, 31(5), pp. 516-528, disponible en: <https://doi.org/10.1080/02671522.2016.1225351>
- Vásquez, C., G. Girón, I. de La Luz, y A. Ayón (2013), *The Effects of Educational Reform*, Annual International Conference of the Bulgarian Comparative Education Society, Plovdiv, Bulgaria. Disponible en: <https://files.eric.ed.gov/fulltext/ED567169.pdf>