

LOS RIESGOS DE LA EVALUACIÓN

Nadie pone hoy en duda la importancia y la necesidad de la evaluación. El debate se centra más en cómo hacerla y en cómo encaminarla a la mejora.

No se trata de hacer muchas evaluaciones, ni siquiera de hacerlas bien. Lo verdaderamente importante es saber a quién benefician y a quién perjudican. En definitiva, de saber a qué valores sirven y qué valores destruyen. Porque la evaluación es un fenómeno de naturaleza ética, más que técnica.

A pesar de sus grandes posibilidades, la evaluación entraña hoy algunos riesgos a los que me voy a referir brevemente:

1. Las trampas del lenguaje: ¿de qué evaluación hablamos?

El lenguaje es como una escalera por la que subimos a la comunicación y a la liberación y por la que bajamos a la confusión y a la dominación. El problema del lenguaje no consiste en que no nos entendamos sino en creer que nos entendemos cuando, realmente, estamos diciendo cosas no sólo distintas sino contradictorias.

Muchos debates políticos, sociales, económicos y educativos están larvados por la suposición de que estamos compartiendo códigos que no compartimos. El caso de la evaluación es paradigmático. Cuando hablamos de evaluación (en castellano utilizamos la misma palabra para referirnos a lo que los ingleses denominan *accountability*, *assessment*, *appraisal*, *self evaluation*...) es probable que cada uno tengamos en la cabeza una concepción diferente sobre el término. Hay quien utiliza el término evaluación, por ejemplo, para referirse a la simple calificación. Hay quien la entiende como un proceso de control, o de comparación o de selección. Y hay quien piensa que se trata de un fenómeno destinado a conseguir la mejora.

Me estoy refiriendo en este primer peligro a la confusión que genera el lenguaje. Considero que es muy importante hacer evaluación, pero no estoy muy interesado en potenciar la evaluación de resultados a través de pruebas estandarizadas (*assessment*) con el fin de elaborar rankings y clasificaciones.

Por eso resulta importante que nos aclaremos, que sepamos de qué estamos hablando. Cuando no compartimos los códigos semánticos podemos utilizar las mismas palabras y estar hablando de cosas muy distintas.

2. Las finalidades de la evaluación: ¿para qué evaluamos?

La pregunta fundamental de la evaluación es la de su finalidad. ¿Para qué se hace la evaluación? Sus funciones pueden ser múltiples. Enumeraré algunas:

- Clasificar
- Comprender
- Controlar
- Dialogar
- Jerarquizar
- Aprender
- Seleccionar
- Mejorar
- Discriminar
- Medir
- Comprobar
- Motivar
- Comparar

Alguna vez pregunto por la función más deseable que debería cumplir la evaluación. La respuesta, de forma casi unánime, recoge finalidades como mejorar, comprender, aprender, motivar... A continuación pregunto cuál es la función más frecuente en la práctica de la evaluación que se realiza y, en este caso, la respuesta se centra en funciones como medir, clasificar, seleccionar, controlar... La pregunta que se desprende de forma casi inevitable es la siguiente: ¿por qué no coinciden las funciones reales con las ideales?, por qué no están más presentes en la práctica las finalidades que consideramos más positivas?

No se produce un debate potente a la hora de elegir las finalidades de un tipo y otro. El problema surge a la hora de explicar el desajuste entre lo que debería ser y lo que es, entre lo deseable y lo real.

Parece producirse una obstinación, un empecinamiento en mantener aquellas funciones que son más pobres, incluso más

perniciosas. No tiene mucho sentido emprender evaluaciones que no tengan más finalidad que decir que se hacen o aquellas que pervierten el sentido educativo de la práctica, tanto individual como colectiva.

3. Los peligros de la atribución: ¿de quién es la culpa?

La evaluación tiene un componente sumamente complejo y, a la vez, inevitable. Es el componente de comprobación. Aquello que debería conseguir el programa, el Centro, el alumno/a, ¿lo han conseguido?, ¿en qué medida?, ¿a qué ritmo?, ¿con qué medios? No es fácil llegar a conclusiones fidedignas, trátase de competencias, conocimientos, actitudes o destrezas. No existen formas inequívocas de responder a estas preguntas. No existe un método o instrumento infalible de comprobación.

Pero hay otro componente de la evaluación no menos importante, que es el de atribución. Y este componente no se tiene en cuenta con la misma fuerza. No digo que no se tenga en cuenta, digo que se tiene en cuenta de manera menos explícita y, a menudo, interesada.

Pondré dos ejemplos que el lector/a comprenderá fácilmente:

Un profesor que evalúa la adquisición de conocimientos de sus alumnos/as puede atribuir el fracaso de éstos/as a que son vagos, a que son torpes, a que están mal preparados, a que ven televisión o a que la familia no les ayuda a estudiar... Estas atribuciones suelen ser gratuitas y casi siempre exculpatorias.

Un Colegio cuyos alumnos/as han obtenido malos resultados en unas pruebas explica el hecho de manera que no cuestione la filosofía de la institución o el funcionamiento de su equipo directivo y de sus profesionales.

Esta cuestión es sumamente importante porque tiene que ver con la capacidad de mejora que puede propiciar la evaluación. Si todas las causas del fracaso, se manejan de forma exculpatoria, aunque las evidencias indiquen lo contrario, la posibilidad de mejora se reduce a cero. Quienes tienen que mejorar son los demás.

4. El contexto de la evaluación: ¿dónde se realiza la evaluación?

Para entender los textos es imprescindible conocer los contextos. Es decir, para saber cómo es la evaluación, resulta imprescindible saber dónde se realiza.

Me voy a referir a dos contextos dentro de los cuales se realiza la evaluación educativa:

a. El contexto cultural

Vivimos envueltos en la cultura neoliberal. No es, a mi juicio, el mejor de los contextos para realizar una evaluación auténticamente educativa. Es decir, una educación que no sólo se ocupe de valores educativos sino que eduque al que la hace y al que la recibe.

La cultura neoliberal inspira las grandes políticas y también la macroeconomía. Pero condiciona también las actitudes, las concepciones y los comportamientos de las personas.

Pues bien, la cultura neoliberal está regida por unos presupuestos que, en su mayoría, contradicen los presupuestos de la educación: individualismo exacerbado, competitividad extrema, obsesión por la eficacia, relativismo moral, conformismo social, imperio de las leyes del mercado, privatización de bienes y servicios, olvido de los desfavorecidos, hipertrofia de la imagen, reificación del conocimiento...

En ese contexto la evaluación se encuentra con importantes condicionantes. No es de extrañar que se confiera una enorme importancia a la cuantificación, a la competitividad, a los resultados...

b. El contexto institucional

¿Qué tipo de organización es la escuela? ¿Qué normas la rigen? ¿Qué rutinas la dominan? Para comprender la evaluación que se realiza en la escuela es preciso comprender la estructura y el funcionamiento de la institución que la alberga.

Se trata de una institución heterónoma, jerárquica, de reclutamiento forzoso, con fines ambiguos e, incluso, contradictorios, con

fuerte presión social, con débil articulación, homogeneizadora, androcéntrica...

La enorme cantidad y precisión de las prescripciones dejan poco margen a la autonomía institucional y a la libertad individual.

5. Desigualdad de las evaluaciones: ¿Cómo castigar una realidad a través de la evaluación?

Para tener éxito en el sistema educativo hay que tener éxito en la evaluación. Por eso resulta tan importante reflexionar sobre ella: condiciona todo el proceso.

Existe un peligro que consiste en realizar un tipo de evaluación diferente, uno de menor importancia y rigor que otro, valorar para realidades que son igualmente importantes.

Pondré un ejemplo que es altamente clarificador. Me refiero a la diferente manera de evaluar la docencia y la investigación en la Universidad. Para tener buen resultado en la evaluación de la docencia basta haberla hecho. O peor, algunas veces ni siquiera hace falta hacer docencia. Cuando pasan cinco años, de forma automática, se considera que se ha realizado bien la tarea y se produce una mejora del sueldo.

Pero, la evaluación de la investigación tiene otras exigencias. Se constituyen comisiones nacionales que utilizan cada año criterios más exigentes. Muchos profesores no alcanzan los mínimos exigidos.

De este diferente criterio se deriva un castigo para la docencia. ¿Qué más da hacerla bien que hacerla mal? Lo que esa política de evaluación “aconseja” es dedicarse a lo que de verdad importa. Da igual cómo prepares las clases, cómo las impartas, cómo hagas la tutoría. El resultado es el mismo. Lo que tiene repercusiones en el sueldo y en la carrera docente es la estrategia de publicación. Si consigues publicar un libro o un artículo en una editorial o en una revista bien situada en el ranking, tendrás éxito y la consiguiente mejora de la remuneración. Es más, ahora la investigación no sirve sólo para mejorar el salario sino que se ha convertido en un requisito esencial para la carrera docente.

6. Las relaciones con el poder: ¿a quién sirve la evaluación?

Hay que preguntarse con insistencia a quién sirve la evaluación. ¿Sirve a la comunidad? ¿Sirve al poder?

Cuando un patrocinador encarga la evaluación a los técnicos, éstos devuelven los informes a quien los ha contratado o pagado. Pero, si el equipo de evaluación no tiene cuidado, si no procura que la comunidad conozca los informes, puede el patrocinador utilizar la evaluación a su antojo. Puede silenciar, cambiar, aplazar, seleccionar, explicar como quiera los informes.

Cuando una evaluación se realiza con dinero público es preciso garantizar la entrega de los informes a la comunidad,. Porque, cuando el poder pretenda manipular los resultados, la comunidad podrá decir: “Alto, yo tengo los informes. Eso no es así. Eso no se deduce de la evaluación”.

De ahí la importancia que tiene el hecho de que los evaluadores y evaluadoras sean personas con independencia de criterio y auténtica posición ética respecto al uso de la evaluación.

7. La cultura de los titulares: ¿qué se dice de la evaluación?

Vivimos en la cultura de los titulares. Hay quien vive intelectualmente de los titulares de prensa o de televisión. En ellos bebe y de ellos se nutre. Nunca va más allá del impacto causado por las frases que abren los periódicos o los telediaros. Como se comprenderá, esto es muy peligroso. Porque los titulares no lo pueden explicar todo. Porque los titulares constituyen una forma peculiar de filtrar la realidad. Porque los titulares suelen escribirse para causar un golpe de efecto, cuando no para servir a los intereses espurios de quien los escribe.

El problema es que los titulares crean estado de opinión. Porque, dadas las prisas, no se lee la explicación que los justifica. Porque se suele comprar el periódico o ver la cadena que está en la línea de la propia línea argumental y porque vivimos en un mundo de fragmentos informativos y de atención voluble..

Hay casos en los que la situación se agrava. Me refiero a aquellas situaciones en las que todos los periódicos coinciden en los titulares. Pondré un ejemplo. ¿Qué piensa el gran público de los resultados del Informe PISA? Lo que han dicho los titulares de prensa. No es que la gente no haya leído el Informe, es que no ha leído ni

siquiera los artículos que desarrollan los titulares. Como éstos buscan causar impacto, el resultado suele ser muy peligroso. "España es el furgón de cola de los países europeos", "La enseñanza privada aventaja a la pública" venían a decir los titulares. Y muchos convierten el titular en un dogma, en un eslogan que se repite sin argumentos y que, a base de repetirse, se convierte en un estribillo. Podría referirme a miles de ejemplos.

¿Qué peligros encierra esta configuración del estado de opinión a través de los titulares?

- En primer lugar, produce una simplificación abusiva del pensamiento. La argumentación desaparece llegando de forma repentina y gratuita a la conclusión.

- En segundo lugar, genera una tergiversación casi inevitable de la realidad. Porque es imposible encerrar en tan pocas palabras la comprensión de un fenómeno complejo.

- En tercer lugar, se origina una dogmatización de las opiniones, ya que parece que los enunciados del titular constituyen verdades indiscutibles. Como las personas no acostumbran a comprar y leer dos o más periódicos, se concluye que esa verdad es un dogma.

8. El efecto Mateo: ¿A quién beneficia la evaluación?

La expresión "efecto Mateo" proviene de la cita bíblica que se recoge en el capítulo 13, versículo 12, del Evangelio de San Mateo (que se repite en Mateo 25, 29 y en otros evangelios hasta en cinco ocasiones) y que dice así: "Porque al que tiene se le dará y tendrá en abundancia; pero al que no tiene, incluso lo que tiene se le quitará".

Autores relevantes de la psicología, como Merton o Bunge, han reflexionado sobre el efecto Mateo, entendiendo que los bienes a los que la cita hace referencia pueden ser materiales (dinero, riqueza, posesiones...) o inmateriales (prestigio, confianza, autoestima...).

Merton fue el primero que utilizó este concepto aplicándolo a la producción científica. Dice que un científico acreditado tendrá menos problemas en hacer valer sus investigaciones que un joven desconocido en el ámbito de la ciencia. Por eso funcionan tan bien los padrinzgos. Un autor novel estampa su firma al lado de un autor consagrado. Padrino y ahijado se benefician del efecto Mateo. El ahijado realiza la verá cómo su trabajo ve la luz (beneficio del

ahijado) y el padrino se apunta un mérito atribuyéndose un trabajo que no ha hecho (beneficio del padrino). Uno pone el nombre y el otro pone el esfuerzo. Los dos se benefician

Si un autor conocido lleva un texto a una Editorial o a una revista para su publicación, es probable que se lo admitan sin más preámbulos. Es más, es probable que le pidan originales sin saber qué calidad van a tener.

Cuando se hacen evaluaciones de las que se derivan clasificaciones, los situados en cabeza se benefician del efecto Mateo, mientras que los últimos de la lista volverán a ser castigados por su efecto nocivo. Por ejemplo, si en el Informe PISA, Finlandia ocupa el primer lugar, el país será premiado con el reconocimiento y el prestigio. No hace falta tomar decisiones para gratificar a los primeros. Basta que sean primeros para ser gratificados...

Este efecto tiene afinidad con otros que han sido estudiados por la psicología educativa, por ejemplo “efecto Pigmalión”, “efecto halo”, “efecto bola de nieve”, “efecto riqueza”, “efecto acumulativo”... Todos ellos tienen planteamientos análogos.

9. Los riesgos de la cuantificación: ¿por qué la obsesión por los números?

Creo que uno de los peligros más dañinos de la evaluación es considerar que las cifras son dogmas indiscutibles. Algunos piensan que cuando hay número, hay ciencia. No siempre es así. Los números están cargados de trampas.

En primer lugar porque hay que analizar cuál es la calidad y el rigor del dato numérico. Puedo decir que una persona obtiene un nueve sobre diez en una escala de motivación. El problema radica en saber con qué rigor se ha realizado la atribución.

Luego comparamos los números. Y entonces ya parece indiscutible decir que 9 es más que 8. O que el que no haya obtenido un 5 tiene una puntuación insuficiente.

El que las pruebas sean estandarizadas nos dice también algo sobre las pretensiones. Se busca efectuar una comparación. Pero, ¿se pueden comparar dos realidades incomparables? La comparación,

para ser precisa, debería contrastar realidades similares. Me temo que los resultados de estas pruebas, manipulados por la prensa y por la derecha, vayan en detrimento de la escuela pública. Los padres, obnubilados por el señuelo de los resultados acabarán diciendo que las escuelas privadas obtienen mejores resultados y que son de mejor calidad que las públicas. Sin hacer explícita la trampa de que las condiciones de partida de los alumnos y alumnas son muy diferentes.

Hay experiencias magníficas, comprometidas con aquellos a quienes Paulo Freire llamaba los "desheredados de la tierra", que quizás no queden muy bien parados en una evaluación de diagnóstico que aplica pruebas estandarizadas. Pero en esas escuelas se trabaja mucho y se trabaja bien. No me opongo a una evaluación exigente, me opongo a una evaluación tramposa. A muchas de las escuelas que ocupen los primeros lugares en el ranking habría que preguntarles dónde van a estudiar aquellos alumnos y alumnas que tienen malas condiciones de partida, escasas expectativas, fracaso continuado y ninguna ayuda en la familia. ¿Qué hacemos con ellos? ¿Quién les echa una mano?

El principal problema reside en que al haber número en el tratamiento de los resultados, se piensa que indiscutiblemente hay ciencia. Y que la ciencia es incontrovertible. No. No siempre que hay número hay rigor. Los datos, sometidos a tortura, acaban confesando lo que quiere quien los maneja. Un vendedor de hamburguesas de pollo recibe una inspección y, al analizar las carnes descubren los inspectores que hay en ellas importantes cantidades de otras carnes. Le preguntan al vendedor:

- ¿Con que carnes hace usted las hamburguesas?
- Con carne de pollo, aunque mezclo con otras carnes, contesta el vendedor.
- ¿Con qué otras carnes?, inquiera un inspector.
- Con carne de caballo, responde el dueño del negocio.
- Y, ¿en qué proporción hace las mezclas?, pregunta el inspector.
- A un cincuenta por ciento. Por cada pollo, un caballo.

En los números se pueden encerrar trampas mortales. Un individuo que no sabía nadar se ahogó al atravesar un río del que le dijeron que tenía 80 centímetros de promedio de profundidad.

Familias, profesorado, alumnado y ciudadanos en general debemos saber a qué llamamos una escuela de calidad. Se dice que los números cantan. Yo añadiría que los números desafinan. ¿Por qué dicen los padres y las madres que una escuela es de calidad? ¿Sólo por los resultados? Juntemos en una escuela a hijos de familias que saben hablar inglés, que tienen dinero para mandarlos a Inglaterra, para pagar un profesor particular, un lugar cómodo y ordenado para estudiar y, sobre todo, la convicción de la importancia de los idiomas para desenvolverse en la vida. Congreguemos en otra a los alumnos que carecen de estas condiciones. Apliquemos luego unas pruebas estandarizadas para evaluar los rendimientos del aprendizaje del inglés. ¿Sería justo decir que la primera es una mejor escuela porque sus alumnos obtienen mejores resultados? ¿Sería lógico decir que es de calidad si practica para hacer la selección el racismo, la xenofobia, la insensibilidad y el elitismo?

10. Los problemas de metaevaluación: ¿quién evalúa la evaluación?

La metaevaluación es un proceso consistente en evaluar las evaluaciones. No se trata de evaluar por evaluar. De evaluar para decir que se está evaluando. Hace falta preguntarse muchas cosas respecto a las evaluaciones que se han realizado. A saber:

- a. ¿Tiene rigor esa evaluación? ¿Es fiable? ¿Ha estado realizada de modo que podamos dar por buenos sus resultados?
- b. ¿Ha servido para algo? ¿Ha favorecido la mejora del funcionamiento de la institución o del programa?
- c. ¿Qué nos ha enseñado? ¿Qué se puede aprender del proceso y de los resultados obtenidos?
- d. ¿Se puede trasladar lo que se ha descubierto a otros contextos similares?

No es de recibo hacer evaluaciones sin ton ni son (Santos Guerra, 2003). Hay que revisar los criterios que la han impulsado, la forma en que se ha hecho y los efectos que ha producido.

Cuentan que en China se produjo una enorme invasión de ratas. La alarma se hizo mayor al saber que las ratas eran portadoras de una terrible epidemia. La proliferación fue tan grande que el gobierno

decidió tomar cartas en el asunto y preparó rápidamente un decreto con la intención de acabar cuanto antes con la plaga. En él se anunció que se premiaría con una cantidad de dinero a todos los que se presentasen en el Ayuntamiento mostrando una rata muerta. Los empleados las recogerían y las quemarían para acabar con el problema. El dinero por cada rata muerta era tan abundante que las ratas, de antes amenazadores y repungantes, se convirtieron en bienes preciados, de manera que las personas las buscaban y las sacrificaban sin descanso.

¿Qué sucedió? Que los chinos descubrieron muy pronto que la cantidad de dinero percibida por las ratas capturadas y entregadas al Ayuntamiento era tan suculenta que decidieron dejar de plantar arroz y ponerse a criar ratas. El problema no se hizo esperar. Faltaban alimentos. Tenían mucho dinero, pero era un dinero que no les permitía satisfacer sus necesidades más perentorias.

La medida parecía lógica, pero la realidad torció la intención del legislador. La pretensión de acabar con las ratas se convirtió en el principal modo de multiplicarlas. Si se hubiesen quedado tan tranquilos, sin ver cómo evolucionaba la realidad, hubieran sufrido graves consecuencias.

Hay que estar atentos, pues, a la realidad. Hay que analizar qué es lo que puede cambiar la intención de quien decide. Intereses de otras personas se interponen, a veces, en la puesta en acción de una medida cargada de bondad y de lógica. Otras veces es la aparición de nuevas e inesperadas circunstancias lo que acaba pervirtiendo la voluntad benéfica de quien decide. Quizás, en algunas ocasiones, sea un malhadado azar.

Referencias bibliográficas

SANTOS GUERRA, M.A. (1993): *La evaluación, un proceso de diálogo, comprensión y mejora*. Ed. Aljibe. Archidona.

-- -- (1998): *Evaluar es comprender*. Ed. Magisterio del Río de la Plata. Buenos Aires.

-- -- (2003): *Una flecha en la diana. La evaluación como aprendizaje*. Ed. Narcea. Madrid.

-- -- (2008): *Nieve y barro. Metaevaluación del Plan de Evaluación de Centros Escolares en Andalucía*. Ed-Eduforma. Sevilla.