



ATENCIÓN AL CONTEXTO: JUSTICIA Y EQUIDAD EN EVALUACIÓN

Felipe Martínez Rizo

Seminario Internacional

**Evaluación y diversidad: justicia y equidad en la
evaluación educativa**

INEE, septiembre 23, 2016

Planteamiento de la cuestión

¿Es aceptable usar evaluaciones estandarizadas si los contextos de los evaluados son distintos?

¿Son compatibles uniformidad, justicia y equidad?

Hay respuestas afirmativas y negativas, pero unas y otras muchas veces carecen de matices y sustento.

Intentaré una respuesta matizada y sustentada.

Partir de una definición

Evaluar es emitir un juicio de valor sobre algo.

Supone inicialmente definir con claridad lo que se pretende evaluar, y luego dos acciones:

- Medirlo con la mayor precisión, pero además:
- Definir un referente de carácter normativo.

Sólo después se podrá contrastar el resultado de la medición con el referente, y formular el juicio de valor en que consiste centralmente la evaluación.

Ingredientes de la inequidad

La justicia-equidad de una evaluación pertenece al ámbito de los juicios de valor. Los pasos que deben darse para llegar al juicio aportan algo:

- La medición puede estar sesgada, no ser igual de precisa, confiable y válida para todos los sujetos a evaluar;
- El referente normativo puede ser desigual de pertinente y apropiado para algunos sujetos.

En consecuencia, la contrastación del resultado de la medición y el referente, puede llevar a un juicio de valor justo y equitativo o no.

La medición y sus dificultades

Si la medición de una competencia da resultados distintos de sujetos de grupos diferentes en algún aspecto (sexo, nivel socioeconómico, lengua, etc.) las diferencias se pueden deber:

- ◆ Al **impacto** real de diferencias de los sujetos en su habilidad respecto a la competencia, que se reflejan en desempeños desiguales;
- ◆ Al **sesgo**, si algunos rasgos del instrumento de medición y/o su aplicación, que no tienen que ver con la habilidad, afectan diferencialmente el desempeño de algunos sujetos.

(Puede haber sesgo haya o no diferencias en los niveles de desempeño en la medición)

Los referentes y sus problemas

La medición busca identificar **cómo ES** el rasgo a evaluar, con la máxima fidelidad **empírica**.

Al definir referentes se pretende establecer **cómo DEBE SER** ese rasgo, de manera **normativa**.

Para fijar referentes se requieren consideraciones:

- ◆ Sobre su *deseabilidad teórica*, con reflexiones de tipo filosófico (ética y filosofía social) sobre lo que se considera bueno y, por ello, deseable;
- ◆ Sobre su *factibilidad práctica*, según el contexto de que se trate y sus circunstancias;
- ◆ Sobre su *viabilidad política*, según el apoyo o rechazo y las consecuencias previsibles.

Medición y referentes

Tanto la medición como la definición de referentes pueden tener deficiencias.

De cualquiera de esos dos componentes de la evaluación, o de ambos, se pueden derivar juicios de valor que den lugar a problemas de injusticia-inequidad.

El cuidado del sesgo

En toda medición hay riesgo de sesgo. El cuidado del mismo debe ser parte del desarrollo, piloteo y aplicación de todo instrumento de medición según los principios del *Diseño Universal*.

- ◆ Al desarrollar un instrumento el sesgo se debe cuidar con comités y procedimientos *ad hoc*;
- ◆ El análisis debe incluir en forma sistemática el Funcionamiento Diferencial de los Ítems (DIF);
- ◆ El DIF no basta para concluir que hay sesgo; es necesario complementarlo con análisis de ítems con protocolos verbales y otros procedimientos.

El cuidado de los referentes

El cuidado de la medición pertenece al ámbito de la *ciencia*; el de los referentes al de la *prudencia*.

Consideraciones para fijar referentes:

- ◆ En el punto de partida deberá haber visión de lo que se juzga bueno y deseable en principio.
- ◆ Perseguir lo deseable a cualquier precio puede ser destructivo; para que *utopía de referencia* sea constructiva necesita complemento de sano realismo que tenga en cuenta la *factibilidad práctica* y la *viabilidad política*.

Sobre la prudencia

Los griegos designaban conocimientos diferentes con términos especiales, como *episteme*, *techne*, *gnosis*, *doxa* o *sophía*, pero el conocimiento que sustenta la toma de decisiones para la acción, la *praxis*, no era ninguno de esos, sino otro que:

jugaba un papel central en la *ética* y *política aristotélica*, el saber iluminador y orientador de la *praxis*, la *astucia* o *sabiduría práctica* que *Aristóteles* llamó *phronêsis*, y los latinos tradujeron como *prudencia*. (Gadamer)

El sesgo en la evaluación de docentes

¿Es posible MEDIR el desempeño de los docentes sin sesgo? ¿Es adecuado utilizar LA MISMA VARA para evaluar a docentes cuyo trabajo se desarrolla en contextos muy diferentes?

EN PRINCIPIO SÍ

Aunque complejo, es posible diseñar instrumentos razonablemente libres de sesgo para medir el desempeño de docentes de contextos distintos.

La estatura depende en parte de circunstancias como buena o mala nutrición y otras relacionadas con la pobreza, y no por eso se deben usar metros distintos para niños pobres o acomodados.

Los referentes en la evaluación de docentes

Para evaluar bien no basta medir sin sesgo; es necesario también tener referentes adecuados.

Es deseable tener excelentes maestros y establecer estándares exigentes para su evaluación.

Finlandia y Singapur tienen altos resultados en las pruebas internacionales y se considera que eso se debe en parte a que tienen buenos maestros, que se forman en excelentes programas que tienen alto nivel de exigencia.

¿Sería políticamente viable y éticamente correcto establecer en México estándares como los de esos países para la evaluación de maestros?

Lo deseable y lo posible

El contexto de países como México es diferente: la profesión docente es poco atractiva, las normales tienen carencias importantes, y muchos de sus alumnos proviene de medios desfavorables.

Esto se debe tener en cuenta al definir estándares para la evaluación de docentes, reconociendo que el nivel de muchos de ellos está lejos del deseable.

De lo contrario, aunque los aspectos técnicos se cuiden, la evaluación puede tener consecuencias negativas, como el desaliento de los maestros y en el límite, el abandono de la profesión.

Hay maestros y maestros...

Hay maestros con bajo nivel académico y además incumplidos. Pero hay otras combinaciones:

Un maestro que trabaja en un escuela remota con muchas carencias y niños con gran dificultad para aprender, pero nunca falta, cuida a los niños y se esfuerzan porque aprendan lo más posible.

Hizo primaria en una escuela multigrado, pasó a una telesecundaria y luego a un tele-bachillerato; se formó en una normal llena de carencias; tiene mala ortografía, no acaba de entender quebrados; sus resultados en una prueba son muy bajos.

¿Es bueno o malo ese maestro?

Conclusión

Si se considera lo técnico, lo ético y lo político:

- Es legítimo medir con la misma vara a sujetos de contextos diferentes, siempre y cuando la vara no esté sesgada, lo que puede conseguirse con cuidados técnicos bien conocidos.
- La medición debe dar resultados consistentes y sustentar inferencias válidas, lo que implica, por ejemplo, que no se deba evaluar maestros sólo con preguntas de opción múltiple.
- Además deben definirse referentes pertinentes y exigibles, que sí deben ser diferentes y tener en cuenta las diferencias de los contextos.

De medios y fines

Toda discusión sobre evaluación debe distinguir el ámbito de los medios y el ámbito de los fines.

- ❖ La evaluación es **medio**. Hay que tener claros los **fines** a cuya consecución debe contribuir.
- ❖ Si una evaluación cumple estándares de validez confiabilidad, ausencia de sesgo y otros puede considerarse buena **técnicamente**.
- ❖ Para ser **buena a secas** debe contribuir a su propósito final. El fin último de la evaluación de docentes es el cumplimiento del **derecho de todos los niños a una educación de buena calidad**, que remite a criterios ético-políticos.



**MUCHAS GRACIAS
POR SU ATENCIÓN**