

LAS MUCHAS FUNCIONES DE LA EVALUACIÓN DE PROGRAMAS EDUCATIVOS

D.C. Phillips, Universidad de Stanford

- “Aquellos que olviden el pasado están destinados a repetirlo”. (Santayana)

JOYAS DEL PASADO

- Un rápido bosquejo histórico
- Evaluación: decisiones o juicios de valor
- Principales efectos, y efectos involuntarios
- Las muchas funciones de las evaluaciones
- La distinción formativa–sumativa
- Más sobre lo sumativo
- Más sobre lo formativo
- La pregunta perfecta para un examen

UN RÁPIDO BOSQUEJO HISTÓRICO

- Ralph Tyler, 1930s—1940s
- Desarrollo de la evaluación y de la investigación de la misma como campo semiindependiente dentro de la comunidad de investigación educativa, 1960s and 1970s
- Pero trabajo mal hecho, y con frecuencia ignorado. La evaluación se había convertido en una especie de rito
- Lee Cronbach y el Consorcio de Evaluación de Stanford, uno entre varios centros.

- Cronbach y sus socios escribieron, en 1981:
“la evaluación tiene una tarea vital que cumplir, y sin embargo sus instituciones y las concepciones que la gobiernan son inadecuadasAdemás, los evaluadores terminan por quedar fascinados con la técnica”.

(Toward Reform of Program Evaluation, p.1)

Evaluación: DECISIONES versus JUICIOS DE VALOR

- Cronbach & Suppes: investigación orientada a la decisión versus investigación orientada hacia la conclusión
- El ensayo de Cronbach de 1963:
podemos definir la evaluación en líneas generales como la *recolección y uso de información para tomar decisiones sobre un programa educativo....* Se deben tomar muchos tipos de decisiones, y son útiles muchos tipos de información. Se vuelve inmediatamente aparente que la evaluación es una actividad diversificada.... (Course Improvement Through Evaluation, TCR, 1963, p. 672)

➤ Michael Scriven: el término “evaluador” se basa en la palabra “valor”:

“El sentido clave del término [evaluación] se refiere al proceso de determinar el mérito, estimación, o valor de algo....El proceso de evaluación normalmente implica ... algo de integración o síntesis de los resultados para lograr una evaluación o serie de evaluaciones generales”. (Scriven, *Evaluation Thesaurus*, 1991, p.139)

- Según el punto de vista de Cronbach, en una democracia el poder para decidir entre prioridades de valor que compiten entre sí no radica en un evaluador no elegido; depende del proceso político.
- La tarea del evaluador consiste en proporcionar a todos los interesados información relevante para que puedan participar en este proceso.

Efectos involuntarios

- Scriven (y LC, y Karl Popper, etc.) se dieron cuenta de que las intervenciones con frecuencia no tienen el efecto deseado, pero las consecuencias involuntarias (positivas y negativas) eran a menudo muy importantes.
- **Ejemplo de un nuevo plan de estudios de matemáticas**
- Scriven ideó la “evaluación sin metas” para tratar con este problema.

Las muchas funciones de la evaluación

- Scriven comentó los varios tipos de evaluación y sus diferencias: evaluación de productos, evaluación de personal, evaluación de programas, evaluación de costo-beneficio, evaluación de componentes, evaluación de dimensiones, etc.. (véase su *Evaluation Thesaurus*)
- Pero enfatizó consistentemente el hacer juicios de valor

- Cronbach y sus socios escribieron:
“Las evaluaciones se inician por muchos propósitos, a veces conflictivos: elegir la mejor posibilidad entre varias líneas de acción propuestas, afinar un programa que ya está en operación, mantener control de calidad, obligar a los subordinados a cumplir con instrucciones, documentar que la agencia de uno merece su presupuesto, crear apoyo para un proyecto favorito, despertar sospechas sobre una política que favorecen los oponentes políticos... (*Toward Reform of Program Evaluation*, 1980, p.13)

- A lo que se puede agregar: lograr una demora importante en la toma de decisión sobre terminar con un programa (hasta que se complete la nueva evaluación), proporcionar un mecanismo en el que las disputas acaloradas entre grupos interesados podrían enfriarse, recabar pruebas a favor o en contra del despido de un administrador o funcionario de programa de alto nivel, establecer crédito o responsabilidad cuando un nuevo administrador o una camarilla política se han apoderado del programa, determinar si el programa realmente se está llevando a cabo de la manera que habían previsto sus diseñadores.

- **Ejemplo de la física del Comité de Estudios de Ciencias Físicas (PSSC)**
- **Ejemplo de una evaluación de una biblioteca académica grande**

- Para que la evaluación sea relevante, se debe determinar su función haciendo preguntas como las siguientes desde el principio:
- “¿Qué individuos o grupos tienen interés en el programa, es decir, quiénes se beneficiarían o perderían con el éxito o el fracaso del programa?”
- “¿Qué querrán saber del programa estos grupos interesados?” “¿Existe una decisión pendiente sobre el programa? ¿Qué opciones se están considerando, cuándo es probable que se tome la decisión y quién la tomará?”
- También se puede nombrar una Junta de Asesores con representantes de los grupos interesados.

La distinción formative – sumativa

- En su artículo de 1963 Cronbach señaló varios aspectos de un programa/plan de estudios sobre el cual debían tomarse algunas decisiones, pero la que enfatizó fue:
- “Mejoramiento del curso: decidir qué materiales y métodos de enseñanza son satisfactorios y dónde hacen falta cambios”

- Este énfasis sobre el mejoramiento del curso indica que Cronbach fue uno de los primeros defensores de la **evaluación formativa**.
- Scriven hizo explícita la distinción entre evaluaciones formativas y sumativas
- “Cuando el cocinero prueba la sopa, ésta es una evaluación formativa; cuando el cliente prueba la sopa, ésta es una evaluación sumativa” (Stake)

Más sobre las evaluaciones sumativas

- Cuando un programa maduro se está volviendo obsoleto, o requiere de una revisión considerable, etc., conviene intentar resumir sus características buenas y malas; o a punto de tomar una decisión de “va/no va”; etc.
- Los evaluadores sumativos han respaldado la prueba aleatoria controlada (RCT) como el diseño ideal
- Este modelo sumativo es cautivador, y por lo tanto es posible que se use cuando no sea conveniente... por ejemplo, cuando se hace caso omiso de la historia del programa (desarrollo de etapas)

- **Ejemplo de evaluación de un programa de alfabetismo en un accidentado entorno rural**
- Los evaluadores con frecuencia piensan que las evaluaciones sumativas son “más sexis” que las formativas
- Cronbach y sus socios advirtieron a los evaluadores de no adoptar una lealtad irreflexiva hacia los “mandamientos” casi religiosos:

- Pondrás a prueba el valor de un programa cuyas metas sean definitivas y cuyos planes se hayan desarrollado por completo. De lo contrario, no evalúes.
- Compararás. Compara el programa que sea de interés central con casi cualquier otra cosa, pero ¡compara!
- Asignarás. De preferencia, distribuye sujetos o instituciones de manera que sean equivalentes los grupos a comparar.
- Medirás el logro de metas...
- Tus herramientas deberán ser objetivas.
- Tus procedimientos deberán ser objetivos.
- Juzgarás. Dile al cliente lo bueno o malo que sea el remedio.
(*Toward Reform...*, p. 215)

- El grupo de Cronbach ofreció algunos consejos aquí:
- Mucho de lo que se escribe sobre la evaluación recomienda algún plan “científicamente riguroso” único. Las evaluaciones, sin embargo, deben tomar muchas formas, y los enfoques menos rigurosos tienen valor en muchas circunstancias. (*Toward Reform....*, p.7)

Y:

- “Es mejor para una indagación evaluativa lanzar una flotilla de estudios que poner todos sus recursos en un enfoque único”. (Ibid.)
- ***Ejemplo: relatos de la evaluación de los contratos de rendimiento. Lo que posiblemente podría fallar, fallará. También ilustra el valor de la “valoración de la evaluabilidad”.***
-

- 6 empresas con fines de lucro firmaron contratos para proporcionar instrucción en matemáticas y lectura a alumnos con bajo aprovechamiento
- Cada firma tenía su propia “filosofía” de instrucción.
- Cada firma tenía tres sitios, distribuidos en forma aleatoria por todo Estados Unidos
- Cada sitio tenía dos escuelas, con alumnos seleccionados como sujetos del experimento o sujetos de control
- Los alumnos del experimento recibieron 2 horas de instrucción en salones separados con el personal de la firma

¿Qué falló?

- Problemas logísticos graves: diseño y cumplimiento antes y después de la prueba, nombrar personal y capacitarlo
- Barras de chocolate
- Sindicatos de maestros
- Equipo de alta tecnología
- Huracán
- **Las lecciones de esto:** valoración de la evaluabilidad, o formativa, pero no sumativa. Y cualquier cosa que pueda fallar, ¡FALLARÁ!

Más sobre evaluaciones formativas

- ¿Son las evaluaciones formativas un tipo diferente o una función diferente?
- Las evaluaciones sumativas se pueden usar de manera formativa, pero con frecuencia las formativas son distintas:
- Más pequeñas en escala
- Presupuesto más pequeño
- Cronograma más corto
- Menos formal en su diseño (¡rigor no es sinónimo de prueba aleatoria controlada !)
- EN GENERAL: “se le saca más partido al dinero”

- **Ejemplo de una evaluación de un texto sobre competencia informática**
- **Ejemplo de evaluaciones formativas de “Plaza Sesamo”**

Evaluación formativa de “Plaza Sésamo”

- Cada programa de 1 hr. consiste de una cantidad de escenas cortas, que se pueden reemplazar fácilmente
- Cada programa lo “patrocinan” unas letras y números, por ejemplo, “las letras A y P, y el número 7” **Pero, ¿cuántos “patrocinadores”, y cuántas repeticiones durante la hora? ¿Cuántas apariciones en programas posteriores para refrescar la memoria?**
- En cada programa destacan seres humanos y títeres de animales, etc. **¿Cuántos en pantalla al mismo tiempo antes de que sea una distracción? ¿Cuáles voces no distraen?**
- Los programas deben gustar a los padres de familia al igual que a los niños.
- **¿Cómo armar escenas para crear un programa de una hora de duración que sea efectivo?**

Por último: una gran pregunta para un examen

- ¿Cuál es la diferencia entre **evaluación** e **investigación científica**?