



ISSN: 1989-0397

LA VENTAJA DE LOS CRITERIOS, LA ESENCIALIDAD DEL JUICIO

Robert E. Stake

Revista Iberoamericana de Evaluación Educativa 2008 - Volumen 1, Número 3 (e)

http://www.rinace.net/riee/numeros/vol1-num3_e/art2.pdf

Estamos aquí, el día de hoy, para hablar sobre la evaluación de la docencia en la universidad. La evaluación es un asunto complicado; la enseñanza, también es un asunto complicado. En una institución educativa un departamento puede preocuparse por enseñar cuestiones muy simples y, a la vez, si sus indicadores de calidad de la enseñanza son demasiado simples; quizás nosotros deberíamos comenzar hoy con la promesa de honrar la complejidad de la enseñanza.

La evaluación de la docencia también es compleja. He tratado de representar la complejidad de nuestra tarea en el mapa conceptual que se presenta en el anexo 1, y que se puede utilizar como índice del contenido de mi presentación.

1. LAS PALABRAS DE LOREN EISELEY

Para capturar la complejidad, buscamos tanto criterios como imágenes icónicas. A continuación un ejemplo de una imagen icónica. Se trata de un pasaje del ensayo, "La naturaleza de hombre" del paleontólogo poeta, Loren Eiseley:

Ahora, después de muchos años de haber sido un educador, a menudo siento la necesidad de buscar una banca en un parque tranquilo para inspeccionar mentalmente el río enorme y anónimo de los estudiantes que ha fluído bajo mis manos. En el dolor he meditado: "Este hombre está muerto –un suicidio. ¿Era yo, quien inconscientemente dirigió en alguna hora negra su mano sobre el arma? Este hombre es un mentiroso y un tramposo. ¿Dónde se equivocó mi golpe?" O allí viene a la memoria del hombre quién, después de mucho tiempo y esfuerzos, regresó felizmente a la granja de la que había venido. ¿Lo serví bien, aunque no coincida con la opinión de los demás? ¿O es que el joven poeta ricamente dotado, a quien protegí de la ira de su padre –lo estaba mimando o defendiendo– en el momento correcto o incorrecto de su vida? Contingencia, contingencia, y cada día por la palabra u obra del cincel de la enseñanza, esculpiendo de forma correcta o a ciegas el futuro de algún niño o niña.

La nuestra es una profesión mal pagada y tenemos nuestra parte de tontos, al igual que la generación anterior a la nuestra, somos los resquebrajados, los aporreados, los productos mal formados de cinceles más remotos que forman la sustancia más obstinada en el universo –la sustancia del hombre. Alguien tiene que hacerlo, pero quizás podría hacerse de forma más amable, con más precisión, al grado que somos deliberadamente conscientes de lo que hacemos –incluso si el pensamiento a veces congela nuestros corazones de terror. ¿O, si estuviésemos más conscientes de nuestra tarea, temblarían nuestras manos o permanecerían inmóviles sobre el cincel?

No lo sé. Sólo sé que en estos últimos años, pusilánimes algunas veces, detengo mi mano sobre la perilla de la puerta antes de entrar al aula.

Los investigadores cualitativos buscan imágenes como las siguientes para ayudarse a describir la enseñanza de calidad: "Ella da clases como Sócrates", "José no explica tan bien como Luis". La gente funciona como estándares útiles o iconos de comparación. Los investigadores cuantitativos prefieren indicadores basados en criterios designados por adelantado, como aquellos indicados en el anexo en el borde derecho. Criterios e iconos. Podemos usar ambos. Necesitamos el pensamiento analítico y necesitamos el juicio subjetivo.

2. EVITANDO LA REPRESENTACIÓN ERRÓNEA

Fui entrenado en la psicometría, pero actualmente vivo en Estados Unidos, como opositor de la psicometría, más que como un defensor de la misma.

Eso se debe al movimiento de rendición de cuentas en la educación preuniversitaria, una demanda brutal para los exámenes estandarizados de alto impacto. Protesto, en parte, porque la forma actual de medir los logros de los estudiantes no es más que un indicador de aptitud escolar y porque los usos políticos de las pruebas dañan más a la enseñanza y al currículo de lo que ayudan (Stake, 1995).

El problema puede ser caracterizado por el lema de algunas personas que se dedican a la aplicación de dichas pruebas y que reconocen su invalidez, pero que dicen: "Bueno, esto es mejor que nada". A veces, una aproximación es mejor que ninguna medición, pero hay veces que aún la medida más exacta causa daño. Lo sé, lo sé, en este coloquio no estamos hablando de la aplicación de pruebas, pero nosotros también nos arriesgamos a dañar a la enseñanza.

Menciono este punto con la esperanza de que podemos prometer, no sólo honrar lo complejo que es la enseñanza sino también asegurarnos que, a través de la investigación y el diálogo profesional, evitaremos dañar a la enseñanza con nuestras propias formas de evaluarla. Se le puede dar demasiada importancia a las rúbricas de observación. También se le puede dar demasiada importancia a las evaluaciones que hacen los alumnos o demasiado énfasis a las evaluaciones entre compañeros. La enseñanza fácilmente puede estar equivocadamente representada por datos casuales. Como Eiseley dijo, a veces la enseñanza es delicada y vulnerable.

3. LA VALORACIÓN DEL INSTRUCTOR

Nuestro interés principal se enfoca en la evaluación de la calidad de enseñanza de profesores individuales. En otro congreso podríamos hablar sobre la evaluación de la calidad de enseñanza para un departamento entero o la evaluación de la docencia de una nación. Pero nuestro enfoque hoy es cómo mejorar la evaluación de la calidad de profesores individuales.

Lo primero que nos llega a la mente es la diversidad de criterios sobre la efectividad de la docencia. Muchos de ustedes tienen gran experiencia examinando cómo los estudiantes o administradores pueden evaluar la efectividad de la enseñanza en clase. Preferirían tener una medición de datos duros sobre lo que está siendo aprendido por los estudiantes, datos con una validación más objetiva que las pruebas que el instructor o el departamento hacen. Pero tal evaluación es lejana y demasiado costosa. Nosotros podríamos hacerlo pero esto costaría más que la enseñanza en sí, e implicaría costos políticos altos.

Dentro del presupuesto, hay dos opciones disponibles: 1) hacer que los estudiantes evalúen la calidad de enseñanza en el aula o 2) tener personas muy familiarizadas con el contenido de los cursos y observar las clases para evaluar la calidad de la enseñanza. La validez de los datos puede ser alta y ambos pueden honrar la complejidad de la enseñanza; a su vez ambas opciones proporcionan datos subjetivos, pero podemos vivir con esto, y aún a veces estar contentos con el juicio humano implicado.

En el caso de los instrumentos que son respondidos por los estudiantes, el conjunto de reactivos generalmente enfatiza el estilo de enseñanza y la relación alumno-maestro. Puede ser informativo, pero

rara vez diagnóstico, y pocas veces conduce a la mejora de la enseñanza. Algunos estudiantes contestan todos los reactivos en términos de cómo se llevan con el maestro. Y, con algunos procedimientos, un solo estudiante puede arruinar la posición de un instructor con puntajes bajos. Pero estos defectos no incapacitan.

Para una observación en el aula, un administrador o evaluador pasa un tiempo en el salón durante la clase y relata lo bueno y lo malo, según su experiencia personal o, a veces, utilizando un juego de preguntas u hoja de evaluación. Al observador, por lo general, le falta capacitación y es susceptible a las opiniones generalizadas en el departamento sobre la capacidad del profesor. A veces la opinión que tiene del maestro no tiene nada que ver con sus clases y está basada más en sus relaciones personales. Es fácil criticar y sugerir mejoras a estos procedimientos, pero aun cuando son llevados a cabo por personas con poco conocimiento de medición educativa, pueden proporcionar información importante.

4. LOS PROPÓSITOS DE LA EVALUACIÓN

Hemos reconocido hasta ahora que la evaluación de la docencia puede ser muy informal o muy formal y puede ser enfocada a la mejora de la enseñanza o utilizada como prueba de calificación. Hemos reconocido los dos objetivos principales de la evaluación: proporcionar la información para la mejora de la enseñanza o para un cambio de categoría del individuo: la promoción, el premio o el aumento de salario y, raras veces pero posiblemente, para la rescisión de contrato. Estos deben facilitar el cambio de nivel, y no deben afectar otras decisiones de personal (Cronbach y Glaser, 1965) como la selección para un puesto vacante. Otro objetivo de la evaluación de calidad de enseñanza es satisfacer los requisitos del departamento aún sin una intención real de revelar los resultados de la calidad de la docencia o emprender programas para mejorar. Esto es una parte común de la burocracia universitaria, y vivimos con ello.

Para un posible cambio de categoría de un profesor, basamos nuestras recomendaciones en nociones generales de calidad, relaciones interpersonales, su lealtad a la institución, su colaboración con colegas, cumplimiento con las políticas, su madurez, si es actualizada, etc. Prestamos poca atención a los objetivos del curso.

Nuestros esfuerzos para ayudar al maestro a enseñar mejor están basados en sentimientos y opiniones. Y eso que ya somos mucho más analíticos, diagnósticos y teóricos en nuestras formas de pensar. ¿Qué dice la investigación? ¿Qué dicen los viejos expertos?

5. ESTUDIO DE CASO

Les leeré uno de los casos que Edith, mi colega, recabó cuando trabajamos juntos en Illinois. Ella estudió el caso de un profesor que fue informado por su jefe y por sus alumnos que no era buen maestro y él mismo lo sabía.

El profesor George Alderman había recibido muchas quejas sobre sus clases; pertenecía al departamento de ciencias con 20 años de experiencia. Recibía evaluaciones de sus alumnos que oscilaban entre adecuado y pobre. George tenía poco dando un curso de matemáticas que era requisito para todos los estudiantes en la facultad. Por lo general los estudiantes evalúan más bajo a los maestros en cursos que son requisitos que los de los cursos que ellos mismos elijen.

Anteriormente, George estaba impartiendo un curso cada semestre a grupos muy grandes y sus evaluaciones eran bajas. Esta vez, el director del Departamento dividió a los estudiantes entre tres profesores, en grupos de 45 cada uno. Las clases eran dos veces por semana. Después de un semestre uno de los profesores recibía muy buenas evaluaciones de sus alumnos. George y el otro recibieron malas evaluaciones. George pidió ayuda al departamento de Desarrollo de la Enseñanza. Los otros dos maestros decidieron trabajar juntos.

El siguiente semestre, sólo George tuvo evaluaciones bajas; eran peores, en parte, porque siguió el consejo de requerir más trabajo en equipo de sus estudiantes. Puso a sus alumnos a calificar entre ellos sus trabajos. Uno de los estudiantes de George le acusó formalmente de calificar en forma caprichosa. El director del Departamento le envió a George una carta donde le exigía mejoría. En entrevistas, el director habló de una necesidad general de mejorar la enseñanza y reconoció que el departamento no había dado retroalimentación directa a los profesores sobre la calidad de sus clases.

Un colega de George, Frank Edwards dijo: "Las evaluaciones de estudiantes no miden la efectividad de la enseñanza. Yo apoyo a la observación por compañeros, no sólo ayudaría al profesor nuevo a desarrollar sus habilidades de enseñanza más rápido, sino también recibiríamos más retroalimentación sobre la calidad de nuestras clases. Sin embargo, las observaciones toman mucho tiempo. Los profesores no son tan bien pagados como los investigadores, entonces algunos piensan que no les conviene observar clases". El director del departamento de George estuvo de acuerdo, y dijo: "Nos gustaría usar varias fuentes como ex-alumnos o compañeros para evaluar las clases, además de estudiantes actuales, pero esto no es posible. Usamos a estudiantes porque ellos son accesibles".

George describió así al especialista enviado por la División de Desarrollo de Enseñanza para ayudarlo: "La considero la única fuente disponible de retroalimentación significativa para mejorar mi enseñanza".

En otras palabras, no había ninguna "comunidad de práctica" para ayudar a George a dar mejores clases. Él y sus colegas están de acuerdo en que la evaluación entre compañeros podría ser una fuente valiosa de retroalimentación pero insisten que la ayuda debe ser individual en vez de colectiva. Bill Wilson, un amigo de George dijo: "En, al menos, cuatro ocasiones he pedido a un colega sentarse en mi clase y proporcionar retroalimentación. A veces la retroalimentación es útil; a veces no. En mi experiencia la retroalimentación de un compañero es más útil cuando la proporcionan en privado (lo cual significa que el director no fue informado del problema o el esfuerzo para solucionarlo).

Frank dijo: "En este departamento, las tutorías formales no existen. Es bastante común, sin embargo, tener una tutoría informal, donde un miembro de la facultad ayuda a otro". George dijo: "La retroalimentación de un colega puede funcionar si varios profesores enseñan la misma materia, se observan uno al otro, y luego tienen una plática informal. Pero esto tiene que ser como una consulta y no implicar que los profesores escriban evaluaciones formales para la administración".

El caso de George es un ejemplo de un departamento que expresamente comunicó sus intenciones de mejorar la enseñanza. Pero sin el apoyo de la dirección, George y los otros profesores sufrían de falta de apoyo. Muchos querían retroalimentación. Sus esfuerzos para mejorar la enseñanza no incluyeron una decisión sobre la responsabilidad de cada profesor de ayudar a los demás a mejorarse. Esto se vio como una responsabilidad individual.

A pesar de que en la dirección de la facultad no se habían dado cuenta de la necesidad de una comunidad de práctica, era claro que había disponibilidad de algo de ayuda entre compañeros. Muchos

de nosotros consideramos que es responsabilidad de la dirección cambiar los sistemas y ya no depender de un sistema estelar (*star system*) tanto en la investigación como en la enseñanza, para estimular la colaboración.

Situaciones como ésta son comunes en Estados Unidos, incluyendo a los estudiantes de posgrados que dan clases. Hay muy poca ayuda disponible para que mejoren sus clases.

6. EL JUICIO HOLÍSTICO DE LA ENSEÑANZA

Como ya saben, la práctica más común hoy en día es hablar analíticamente de la enseñanza de calidad en términos de criterios. En un análisis común de la enseñanza, es valioso hablar de la persona en su totalidad, usando tipologías y descripciones de incidentes específicos. Pruebas anecdóticas de éxito y fracaso pueden ser un problema pero, por lo general, esto es un componente importante de una evaluación comprensiva.

Se han realizado esfuerzos para clasificar la enseñanza y describirla más holísticamente. Harry Broudy identificó categorías de enseñanza en su libro, *Ejemplos de enseñanza*. Distinguió entre las metas de la enseñanza como: didácticas, heurísticas, y filéticas y que los profesores difieren en cuanto a los usos del conocimiento que enfatizan, como: replicativo, asociativo, aplicativo e interpretativo. Estas categorías pueden ser útiles para reducir la dependencia que tenemos en el análisis de rasgos personales, como a veces hacemos.

7. LOS DEBERES DE LOS INSTRUCTORES

Podríamos diseñar una evaluación más comprensivamente si incluimos en ella a todos los deberes del profesor, que rara vez son explícitos, más bien son de costumbre, pero se asume su cumplimiento como si fueran obligaciones de contrato. Cuando un profesor no recibe una planta o una promoción, por lo general se debe a falta de empatía, o una violación ética, o trabajos incompletos, pero casi nunca se debe a una falla en un deber explícito. Sin embargo, tal cumplimiento de deberes visible a los ojos de la administración, compañeros y estudiantes, puede ser la forma más lógica para una evaluación formal de desempeño. Es fácil encontrar una lista de deberes. El experto en evaluación Michael Scriven (1988, 1995) desarrolló un método que llamó la Evaluación Basada en los Deberes del Profesor (DBTE); En la tabla 1 se desarrolla cómo los identificó.

Con una Evaluación Basada en los Deberes del Profesor, Scriven impulsó a la administración de las universidades a explicar las labores de los miembros de la facultad. Propone el trato tanto numérico como comparativo. Su meta fundamental era reducir el sesgo de la evaluación. Cuando los deberes son expresados claramente, hay menos posibilidades que el sesgo del evaluador determine los resultados. El sesgo tiene que ser controlado; pero, en mi opinión, muchas veces es imposible distinguir entre el sesgo y un entendimiento profundo. Enfocar en la utilización de palabras claras y concisas puede disminuir la atención a relaciones importantes complejas. Esto puede cambiar el significado de mérito. En lo siguiente, propongo más énfasis en el alcance de juicios complejos de la enseñanza de valores y menos énfasis en la representación de mérito numéricamente.

TABLA 1. LAS TAREAS PROFESIONALES DE UN INSTRUCTOR

Conocimiento de deberes	Incluye el conocimiento de la ley y regulaciones que aplican a su universidad o institución así como las expectativas de su escuela en particular. Esto supone el conocimiento de las exigencias del plan de estudios y deberes en general; la obligación de seguir órdenes y regulaciones, pero esto jamás justificaría un trato poco profesional o una violación de ética.
Conocimiento de escuela y comunidad	Incluye el conocimiento de cualquier característica especial, fondo o ideología de la escuela, su personal y estudiantes y de su ambiente en particular.
Conocimiento de materia	Supone el conocimiento de: <i>a)</i> en el campo (s) de capacidad especial; <i>b)</i> de cursos adicionales como inglés, hábitos de estudio, conocimiento personal/profesional, estudios de computación, etc.
Diseño educativo	Consta de: <i>a)</i> diseño del curso; <i>b)</i> selección y creación de materiales; <i>c)</i> empleo adecuado de recursos disponibles; <i>d)</i> evaluación del curso y plan de estudios; <i>e)</i> conocimiento para adecuar a las necesidades de grupos especiales; y <i>f)</i> empleo de recursos humanos especializados cuando los haya (especialistas en planes de estudios, expertos en audiovisuales y métodos, etc.)
Recopilar información del aprendizaje del alumno	Incluye: <i>a)</i> diseño de exámenes; <i>b)</i> habilidad para calificar (calificación, corrección, rangos, etc.); <i>c)</i> procedimientos claros para calificaciones; y <i>d)</i> asignación de grado.
Proporcionar información del aprendizaje del estudiante	Supone: <i>a)</i> a cada alumno; <i>b)</i> después de cada prueba; <i>c)</i> a la administración; y <i>d)</i> a padres, tutores u otras autoridades apropiadas
Habilidades en clase	Incluye: <i>a)</i> habilidades de comunicación; <i>b)</i> habilidades de dirigir bajo el estándar (disciplina, cumplimiento del contenido diseñado con el nivel apropiado de entendimiento), y planes para emergencias (fuego, inundación, tornado/tifón, terremoto, erupción volcánica, etc.).
Características personales	Supone: actitud profesional y desarrollo profesional (buscar sistemáticamente la mejoría en la enseñanza, materiales de clase y contenido, utilizar autoevaluación y planes de desarrollo cuando sea posible).
Servicio a la profesión	Incluye: <i>a)</i> conocimiento de asuntos profesionales; <i>b)</i> conocimiento y funcionamiento que adhiere a la ética profesional; <i>c)</i> ayuda a principiantes y colegas; <i>d)</i> trabajo en proyectos para otros profesionales; y <i>e)</i> contribuciones a la investigación lo cual enriquece la profesión.
FUENTE: Adaptado de Scriven, 1994, pp. 25-32. Usado con permiso del autor.	

8. ¿EN QUIÉN ESTÁN PENSANDO LOS ESTUDIANTES?

Uno de los primeros en desarrollar un instrumento de evaluación de los profesores de parte de los estudiantes era Richard Spencer quien alguna vez me dijo que encontraba las posiciones bastante unidimensionales, que lo más importante era si el alumno quiere al maestro o no y precisamente eso era lo que evaluaban los alumnos independientemente de las preguntas en la evaluación. Es decir, una pregunta puede hablar sobre la ética del profesor pero el estudiante responde: "él es realmente agradable".

Dick Jaeger y E. A. Barnes (1981) eran dos investigadores que hicieron un análisis factorial de evaluaciones de estudiantes, encontraron que el primer factor tenía un peso ponderado más grande que los otros tres. Llamaron al factor mayor "la eficacia total" pero ello no pareció ser uno interpersonal. Su evaluación no incluyó artículos interpersonales. Sus otros factores eran: esfuerzo, clasificación clara e imparcialidad.

Hace mucho, cuando estaba investigando la evaluación de enseñanza, me preocupaban las valoraciones de alumnos basadas en un grupo de referencia no especificada. Los alumnos tenían que evaluar el comportamiento y características de sus profesores pero, a la vez, debían considerar a otros profesores. Propuse una escala que trataba de hacer las evaluaciones más normativas, pidiendo a los alumnos una comparación de su profesor con un grupo de referencia explícito (los mejores profesores que tuvieron).

El anexo 2 muestra la escala que desarrollé. Las hojas de valoraciones verdaderas tenían una línea perforada –pueden imaginárselo– para separar la hoja de trabajo de las valoraciones. Cada estudiante debía identificar a ocho profesores, de su elección, escribiendo los ocho nombres en la hoja de trabajo. Entonces para cada característica, *1-*8, el estudiante debía clasificar al profesor y llenar la caja con el número de profesores mejores que el profesor que están evaluando. No estoy recomendando la hoja, lo menciono para recordarles la falta de claridad de pensamiento que tienen los estudiantes.

9. ACTIVIDADES ACTUALES AL NORTE DE LA FRONTERA

Mario Rueda me pidió decir unas palabras sobre los avances actuales en tema de evaluación de la enseñanza a nivel superior en Estados Unidos. Hablé con mi amigo John, mucho más actualizado que yo, y les puedo decir que no ha habido muchos cambios. Hay súplicas fuertes para evaluaciones de enseñanza que utilizan métodos múltiples. Se sigue usando evaluaciones hechas por estudiantes y hay poca averiguación administrativa sobre la calidad de las clases de los profesores, probablemente lo mismo que hace veinte o tal vez hace cuarenta años.

El cambio más notable, quizá, es en la evaluación del profesor por el alumno hecho en línea. El escáner óptico está siendo sustituido por los resultados computarizados. En mi campus tenemos tres semestres usando este método y los resultados han sido buenos. La participación de alumnos de forma voluntaria y anónima es prácticamente igual en línea que con el método anterior. Me daba miedo que con el hecho de que la evaluación no se hacía en el salón, tal vez los alumnos que no tenían queja, se irían sin hacerla. No parece ser el caso ya que la distribución de participantes es prácticamente la misma.

De acuerdo con la ética de responsabilidad en la educación, aquellos responsables de evaluar el campus para objetivos de acreditación todavía apoyan a la medición de respuestas de aprendizaje y, a veces, apoyan al concepto de “aprendizaje con valor agregado”. Además de la evaluación de aprendizaje, algunas agencias de acreditación regional, como la Comisión para Estudios Superiores de la Región Norte Central, están poniendo énfasis en la disponibilidad de servicios de desarrollo de enseñanza ubicados en el campus. El ofrecer ayuda a los que reconocen que la necesitan es una diferencia chica pero importante para una estrategia de “mejoría basada en diagnósticos medidos”.

Todavía hay una defensa de las observaciones de compañeros y el uso de portafolios pero los costos son prohibitivos con la excepción de programas voluntarios. Se me hace interesante que los directivos de las universidades de Estados Unidos apoyen fuertemente a evaluaciones de desempeño basadas en autoevaluación. El desempeño es, a veces, compañero de la buena enseñanza pero un indicador cuestionable de ello –como mencioné antes (Stake, 1995).

10. EVALUACIÓN COMPRENSIVA

A veces nos sentimos presionados a desarrollar una medida más válida para evaluar competencias. En mi opinión, el mejor instrumento que tenemos es el portafolio, o tres portafolios coincidentes. Un volumen debería ser compartido por el profesor y la dirección, guardado por el profesor para mostrar lo mejor de él. El segundo, también guardado y controlado por el profesor, debería ser usado para la resolución de problemas personales y la mejoría en sus clases. El tercero es para los récords oficiales que deben ser

confidenciales. El segundo y el tercero deben estar disponibles para decisiones de personal hechas por un panel de evaluadores, lo cual deber tener acceso a cualquier información adicional que requiera.

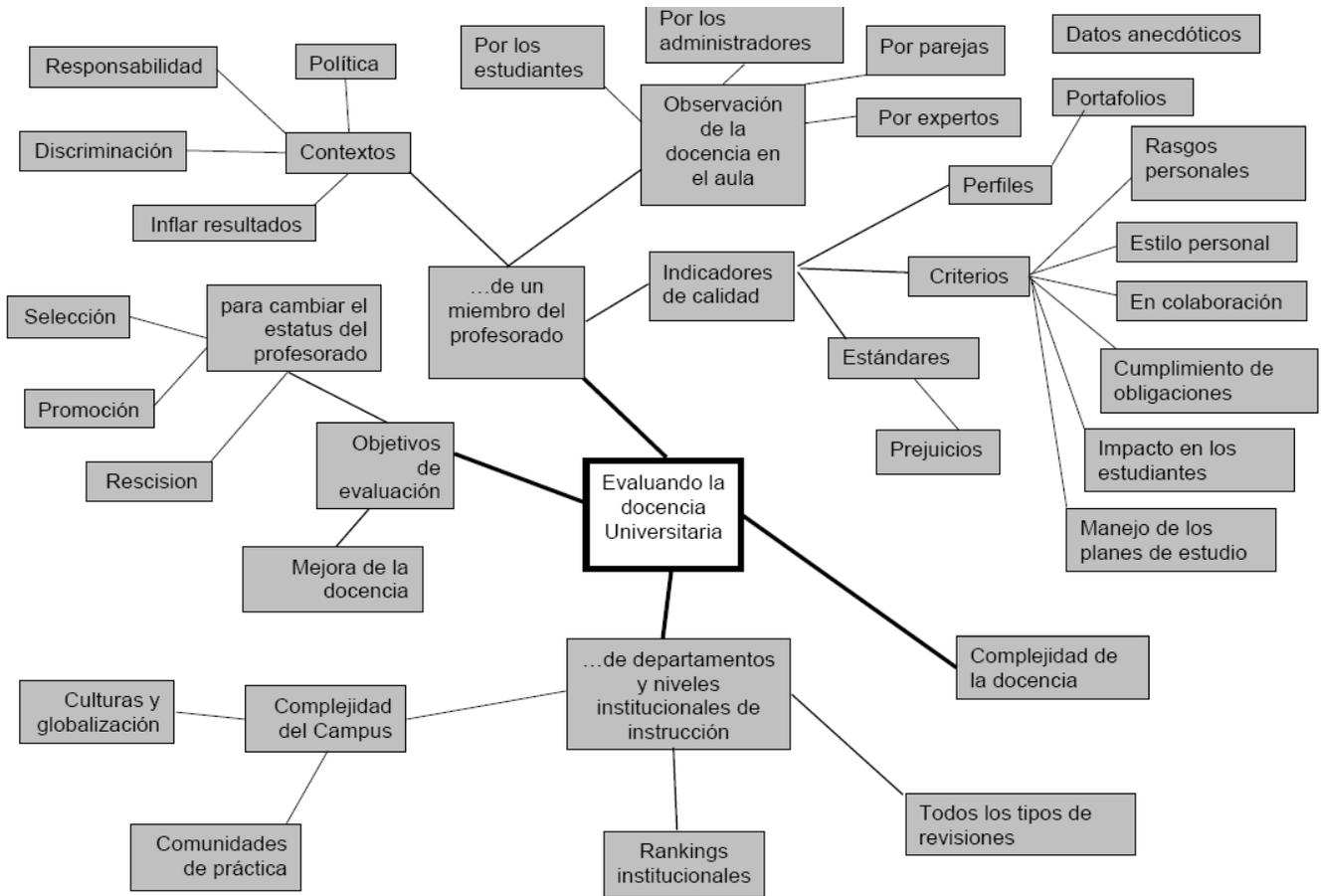
Para las revisiones departamentales, algunos criterios deberían ser fijados por adelantado, pero las decisiones tendrían que ser críticas y no por satisfacer estándares cuantitativos anteriores. Las circunstancias deben estar inscritas en registros; son comunes las confirmaciones independientes.

Como nos ilustra Loren Eiseley, la enseñanza es sumamente compleja, mucho más que los instrumentos que hemos diseñados para medir su calidad. Hay múltiples realidades. Se deben buscar opiniones diversas de gente pertinente. Los indicadores de la efectividad de la enseñanza son simplistas, con un lugar en un portafolio, pero con las reservas debidas. Decisiones importantes para universidades o profesores no deben ser basadas un solo indicador o anécdota. Por ética estamos obligados a usar múltiples fuentes de información para tomar decisiones sobre sus futuros. Los portafolios y evaluaciones por panel dan una indicación válida de la calidad de la enseñanza. Un criterio ayuda a proteger la complejidad y ayuda al individuo a mejorar pero tiene que haber un proceso para evaluar la información departamental.

REFERENCIAS BIBLIOGRÁFICAS

- Cronbach, L.J. y Glaser, G.C. (1965). *Psychological tests and personnel decisions*, Second edition. Urbana: University of Illinois Press.
- Jaeger, R.M. y Barnes E.A. (1981). "Methodological influences on the dimensions of instructor performance". Texto presentado para la reunión anual de la North Carolina Association for Research in Education, Greensboro, noviembre. Un sumario aparece en *Jaeger's Statistics, A Spectator Sport*, 2 ed. Sage, page 352-359.
- Scriven, M. (1988). The state of the art in tertiary teacher evaluation. *Research and Development in Higher Education*, 10, pp. 2-27.
- Scriven, M. (1994). *Evaluating teachers as professionals*. Perth: University of Western Australia.
- Scriven, M. (1995). A unified theory approach to teacher evaluation. *Studies in Educational Evaluation*, 21, pp. 111-129.
- Stake, R. (1995). The invalidity of standardized testing for measuring mathematics achievement. En Thomas A. Romberg, editor, *Reform in school mathematics and authentic assessment*. New York: State University of New York Press.
- Stake, R. (2004). *Standards-Based and Responsive Evaluation*. Thousand Oaks, CA: Sage. En español (2006) *Evaluación comprensiva y evaluación basada en estándares*. Barcelona: Grao.
- Stake, R. y E. Cisneros-Cohernour (en prensa). "Situational evaluation of teaching on campus", en Katherine Ryan, editor, *Evaluation of teaching in higher education: The new millenium*. San Francisco: Jossey Bass.

ANEXO 1



ANEXO 2

Stake Form 22X

Este lado de la línea perforada es su HOJA DE TRABAJO. Usted removerá la hoja antes de entregar sus opiniones.

Nombre de su Instructor



II. Escriba abajo los...



1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____

IV. Después de que usted ha respondido a todas las preguntas, revise sus puntajes. Recuerde que 7 u 8 es en el cuadro un puntaje bajo para **Tu Instructor** y que 0 o 1 en la caja es un puntaje alto para **Tu Instructor**

V. Doble, corte la hoja de trabajo y entregue sus puntajes.

---HOJA DE TRABAJO---

OPINIONES DE LA INSTRUCCIÓN

Tu Instructor desea que tú evalúes su enseñanza en esta clase este semestre con el mismo cuidado con el cual esperas que él evalúe tu trabajo. **Tu instructor** se da cuenta que él tiene considerable oportunidad de decirte lo que él sabe y lo que piensa, pero hay poca oportunidad de que **tú** compartas con él lo que piensas

Hoy puedes expresar algunos de tus pensamientos acerca de la instrucción en su clase como se vincula con tu propio desarrollo educativo. Esta forma es sencilla y fácil de usar. Por favor úsala con el espíritu con el cuál se ha desarrollado, con respeto a tu propia integridad y con respeto a la integridad de tu instructor. Aún cuando permaneces anónimo, por supuesto, considera ésta una comunicación personal con tu instructor. Se intenta que sea solo para él, para ayudarle a formarse mejor idea de como puede contribuir mejor al bienestar educativo de sus estudiantes.

Por favor sigue las Etapas 1 a V cuidadosamente.

↔ I. Escribe el nombre de **Tu Instructor** en la línea de la izquierda.

...nombres de al menos OTROS ocho de los últimos (más recientes) instructores que has tenido en esta Universidad. Si no haz tenido ocho instructores aquí en esta Universidad, escribe los nombres de solo cuatro de los instructores y escribe el número cuatro dentro del círculo.



III. Ahora, compara a **Tu Instructor** con los Instructores listados a la izquierda, uno por uno, en cada una de las siguientes características. El número que tú escribas en cada cuadro debe indicar a cuantos de los otros que haz calificado como MEJORES que tu Instructor en ese aspecto.

1. PREOCUPACIÓN PERSONAL. Cuántos de los instructores de la izquierda son más sensibles y apreciativos de las circunstancias personales de los estudiantes de manera individual que Tu Instructor? Escribe un número aquí →

2. HABILIDADES DE COMUNICACIÓN. Cuántos de los instructores de la izquierda son mejores que Tu **Instructor** al hablar de forma efectiva, estableciendo las ideas, y en lograr que se les entienda? Escribe un número aquí. →

3. ORGANIZACIÓN. Cuántos de los instructores de la izquierda son mejores que **Tu Instructor** en guiar las actividades de la clase hacia los propósitos del curso? Escribe un número aquí →

4. MOTIVACIÓN. Cuántos instructores de la izquierda son mejores que **Tu Instructor** en hacer el periodo de clase interesante y significativo, aún cuando el tema es difícil o aburrido? Escribe un número aquí. →

5. CONOCIMIENTO. Cuántos de los instructores de la izquierda parece conocer el contenido de lo que está enseñando mejor de lo que Tu **Instructor** conoce su contenido. Escribe un número aquí. →

6. EFECTIVO AL CALIFICAR. ¿Cuántos de los instructores de la izquierda son más hábiles al evaluar tu aprovechamiento, en calificar de más forma justa y sensible que Tu **Instructor**? Escribe un número aquí. →

7. ETICA DEL TRABAJO. ¿Cuántos de los instructores de la izquierda trabajan más duro en hacer su curso un buen curso, incluyendo prepararse para la clase, leer tus tareas y responder a tus preguntas? Escribe un número aquí. →

8. CONTRIBUCIÓN GENERAL A TU EDUCACIÓN. Comparado con **Tu Instructor**, ¿cuántos de los instructores de la izquierda han contribuido más a tu propia vida y avance profesional? Escribe un número aquí. →