

# Hacia un marco para evaluar la competencia de los profesores

**Erik Roelofs**

Investigador principal del Centro de información e investigación psicométrica de Cito, Instituto de desarrollo de pruebas de los Países Bajos, sito en Arnhem

**Piet Sanders**

Director del Centro de información e investigación psicométrica de Cito, Instituto de desarrollo de pruebas de los Países Bajos, sito en Arnhem.

## RESUMEN

El desarrollo de instrumentos para evaluar la competencia del profesorado requiere un modelo de rendimiento competente que facilite la recopilación y valoración de pruebas en distintas situaciones de realización de tareas. De acuerdo con Kane (1992), la validación de las evaluaciones sobre la competencia de los profesores se considera una evaluación de argumentación interpretativa.

Basándose en interpretaciones contemporáneas de la formación y el aprendizaje, se denomina modelo interpretativo de rendimiento competente a aquel que, en lugar de ser preceptivo por naturaleza, admite varias formas de desempeño profesional responsable. Las consecuencias del rendimiento profesional para los alumnos/la clase/la organización constituyen la base del modelo, dado que de las consecuencias se derivan las intervenciones correctas y los procesos de toma de decisiones subyacentes, así como los componentes asociados de una base de conocimientos profesionales. A continuación se exponen las consecuencias de dichas interpretaciones para el desarrollo de ámbitos de competencia y la recopilación de pruebas.

## Palabras clave

Competence assessment, teaching personnel, teaching quality, quality of education, performance appraisal, personnel assessment

Evaluación de la competencia, profesorado, calidad de la educación, evaluación de rendimiento del personal docente

## 1. Introducción

A escala internacional, se observa un creciente interés en la evaluación de la competencia de los profesores motivado por la exigencia de garantía de calidad y de un mayor reconocimiento de la profesión de enseñante (Verloop, 1999). En los Estados Unidos existe una larga tradición de evaluación del profesorado, reflejada tanto en el volumen de artículos de investigación y libros publicados, como en los instrumentos desarrollados. En los Estados Unidos, el principio de rendición de cuentas ante los contribuyentes supuso un importante incentivo para dirigir la atención

hacia la evaluación de los profesores. Se han desarrollado varios instrumentos para evaluar a los profesores a lo largo de las diferentes etapas de su carrera profesional en el contexto de la selección, certificación y desarrollo profesional (Dwyer, 1998). Mucha de la información recopilada para este artículo se extrajo de toda esta bibliografía. Sin embargo, ello no significa que no exista también tradición de evaluación del profesorado en otros países, aparte de los Estados Unidos. Por ejemplo, Pelkmans (1998) describió las prácticas de evaluación de los profesores en Inglaterra, Gales, Alemania, Australia y Holanda. No obstante, estos países poseen menos experiencia que los Estados Unidos.

En los Países Bajos, así como en otros países (Pelkmans, 1998), se está empezando a prestar una mayor atención a la evaluación del profesorado, como consecuencia de la mayor capacidad de decisión que tienen las escuelas a la hora de elaborar sus políticas, lo que ha permitido, entre otras cosas, la diferenciación de puestos y remuneraciones (Verloop, 1999; Straetmans y Sanders, 2001). El énfasis creciente en la formación basada en la competencia también está aumentando la demanda de evaluaciones de la competencia de los profesores. Asimismo, una ley aprobada en los Países Bajos establece que los profesionales de la enseñanza deben satisfacer diversos requisitos de competencia. La fundación para la competencia en la enseñanza profesional de los Países Bajos (SBL, 2003) ha formulado una serie de requisitos en siete ámbitos de competencia que se consideran decisivos para los profesores principiantes. En el apartado 3.2.1.1 se tratarán con más detalle dichos ámbitos.

Los centros de formación del profesorado y las organizaciones escolares deben cumplir dichos requisitos. Por último, se ha facilitado la incorporación a la enseñanza de profesionales procedentes de otras profesiones, al objeto de contrarrestar la amenaza de falta de profesores en los Países Bajos (Klarus, Schuler y Ter Wee, 2000; Tillema, 2001).

Los avances descritos anteriormente exigen un enfoque coherente a la hora de evaluar la competencia de los profesores. En este artículo se presentan algunos principios básicos que permiten elaborar un marco para dicha evaluación. Empezaremos creando un modelo interpretativo para evaluar la competencia en el rendimiento (del profesorado), basado en diferentes conceptos teóricos de enseñanza de calidad. A continuación mencionaremos varios instrumentos para la evaluación de la competencia de los profesores mediante el uso de los requisitos que se derivan del concepto de validez de constructo expuesto por Messick (1996). Por último, trataremos algunos problemas para su estudio en un futuro.

## 2. Un modelo interpretativo de competencia de los profesores

No existe una definición generalmente aceptada del concepto de competencia. Últimamente, varios autores (p. ej. Bos, 1998; Mulder, 2001; Van

Merrienboer, Van der Klink y Jansen, 2002) han revisado la bibliografía existente y han elaborado definiciones globales. Una primera distinción importante que cabe señalar es la diferencia entre dos de las acepciones del término "competencia". De acuerdo con Mulder, competencia es, por un lado, un concepto global que se refiere a las habilidades o capacidades de las personas u organizaciones, mientras que, por otro lado, una competencia específica forma parte de dicho concepto global de competencia. Asimismo, la segunda acepción del término competencia (plural competencias) es un concepto más limitado y más atomístico que se utiliza para referirse a habilidades concretas (véase también McConnell, 2001). Basándose en un estudio de numerosas definiciones de competencia (p. ej. Bunk, 1994; Spencer y Spencer, 1993; Parry, 1996), Mulder (2001) propuso una definición que capta las de la mayoría de los autores importantes: "competencia es la habilidad de una persona u organización para alcanzar unos niveles de rendimiento concretos" (p. 76). Citando a diferentes autores, añade que las competencias de los individuos son:

- aptitudes integradas para la acción
- que están formadas por grupos de estructuras de conocimiento,
- aptitudes cognitivas, interactivas, emocionales y, en caso necesario, psicomotrices,
- y actitudes y valores que son necesarios para:
- realizar tareas,
- solucionar problemas,
- y, en sentido más amplio, la habilidad para actuar en:
- una profesión,
- una organización,
- un puesto
- o una función concretos.

A la hora de medir las dimensiones de la competencia, ha de tenerse en cuenta que éstas no se pueden observar directamente, sino que se manifiestan en el rendimiento en una situación específica (Spencer y Spencer, 1993). Por otra parte, es posible desarrollar competencia en un determinado nivel, por ejemplo, principiante, avanzado y experto. Mulder (2001) subraya que la competencia puede estar presente en los individuos (competencia personal) y en los sistemas (competencia de un sistema o un equipo). Finalmente, ciertos aspectos de la competencia son extrapolables de una situación a otra (Thijssen, 1998, 2001).

Al evaluar la competencia se pueden plantear varias preguntas. ¿Cómo elaborar las evaluaciones en este ámbito? ¿Qué suposiciones y conceptos teóricos subyacen a las medidas de la competencia? Para responder a estas preguntas, debe utilizarse un modelo interpretativo que sea suficientemente descriptivo y aclaratorio (Shepard, 1993).

No existe un marco científico sólido y general para determinar lo que constituye la competencia en la enseñanza del que se puedan extraer conclusiones para evaluar la competencia de los profesores (Haertel, 1991). Existen varios marcos, cuyos contenidos dependen en gran medida de la

visión subyacente del rendimiento profesional (Dwyer, 1994, 1998) y de los enfoques teóricos adoptados (véase Reynolds, 1992).

La mayoría de los desarrolladores de instrumentos de evaluación del profesorado buscan una visión compartida de la competencia en la enseñanza, obtenida a través de la interacción entre los desarrolladores y los representantes de la profesión. La visión resultante de esta interacción no es siempre la misma; puede ser una combinación de todos los tipos de visión sobre la enseñanza o tratarse de una visión muy concreta, por ejemplo, "aprendizaje orientado a un programa" o "aprendizaje orientado al desarrollo" en el contexto de la educación primaria o preescolar. Se antoja lógico adoptar el primer enfoque en la formulación de marcos de rendimiento competente aplicables a grupos numerosos de profesores, por ejemplo, requisitos de competencia a nivel nacional. El segundo enfoque es más apropiado para las organizaciones que trabajan de acuerdo con una misión específica.

Además de la visión de la enseñanza, el ángulo teórico sobre el rendimiento profesional también determina la forma que adopta un modelo interpretativo. A lo largo de la historia de la evaluación del profesorado, la bibliografía se ha centrado en distintos elementos que constituyen la competencia. Revisando la bibliografía se pueden encontrar diferentes conceptos de buenos profesores y buena enseñanza (Creemers, 1991; Verloop, 1999):

- a) que diferencian los rasgos de la personalidad que contribuyen al éxito de un profesor (Getzels y Jackson, 1963; Creemers, 1991);
- b) que describen los elementos de conocimiento que afectan al contenido de las asignaturas y a la forma de pensar de los profesores dentro de una determinada disciplina (Bruner, 1963; Tom y Valli, 1990);
- c) que describen las formas de comportamiento de los profesores que contribuyen al aprendizaje (Brophy y Good, 1986; Simon y Boyer, 1974);
- d) que describen procesos cognitivos y de toma de decisiones de los profesores (Kagan, 1990; Verloop, 1988);
- e) que describen los conocimientos prácticos que los profesores aplican en situaciones específicas (en clase o en el ámbito de la asignatura que imparten) y el modo en que formulan teorías sobre estas situaciones (Beijaard y Verloop, 1996).

Para cada uno de estos conceptos de la enseñanza de calidad se utilizaron técnicas de evaluación específicas. En la evaluación de los rasgos de la personalidad, se utilizaron cuestionarios y tests psicológicos para identificar los rasgos positivos y negativos. Este concepto de enseñanza, no se centraba tanto en la enseñanza de calidad como en las características de un buen ciudadano. Una idea que todavía sigue predominando es que los buenos profesores poseen muchos conocimientos. En un principio, esto suponía conocer hechos y elementos discretos; posteriormente, el énfasis recayó en la estructura de una disciplina (como las matemáticas o la física), y más tarde en la forma de actuar y pensar de los profesionales con respecto a una disciplina concreta. Los conocimien-

tos también están relacionados con los aspectos pedagógicos, por ejemplo, métodos de enseñanza, elaboración del currículum y agrupamiento de los alumnos. Estos elementos de conocimiento se obtienen cada vez más de la investigación en materia educativa (Bellon, Bellon y Blank, 1990). Un método utilizado con frecuencia para evaluar los conocimientos consiste en realizar pruebas de conocimientos normalizadas (p. ej. Latham, Gitomer y Ziomek, 1999).

Como reacción al énfasis puesto en las características estables de un profesor y al enfoque unilateral centrado en los conocimientos, hay quienes consideran que la enseñanza consiste en manifestar un comportamiento efectivo. Este concepto tiene en cuenta lo que los profesores muestran en el aula. Se han desarrollado numerosos instrumentos de observación que se concentran en (pequeñas) unidades de comportamiento, que parecen estar conectadas con buenos resultados de aprendizaje (Stodolsky, 1990). Sin embargo, este enfoque presta escasa atención a lo que pasa por la mente del profesor. ¿Qué piensa o decide y por qué decide del modo en que lo hace? Se han utilizado diferentes instrumentos de evaluación para desentrañar el pensamiento de los profesores, incluidos protocolos de pensamiento en voz alta mientras se resuelven problemas docentes, y entrevistas para ayudar a recordar. Durante dichas entrevistas, los profesores contemplan su actuación grabada en vídeo y responden a preguntas sobre lo que estaban pensando en un determinado momento. El concepto de enseñanza como capacidad para transmitir una amplia base de conocimientos prácticos requiere unos métodos de evaluación centrados en las situaciones específicas a las que se enfrenta un profesor. Los instrumentos utilizados no difieren en gran medida de los empleados para desentrañar los procesos del pensamiento. Sin embargo, se centran más en el contexto de trabajo concreto (p. ej. la asignatura específica en un nivel específico) en el que los profesores llevan a cabo sus actividades (Meijer, Verloop y Beijaard, 1999).

Con la reciente revolución de la visión del aprendizaje, resumida en el término "nuevo aprendizaje", ha surgido un nuevo concepto de la enseñanza: fomento de actividades que potencian el aprendizaje entre los alumnos. Según este concepto, para obtener una enseñanza de calidad, los profesores no necesitan manifestar un repertorio claramente definido de comportamientos "correctos", sino demostrar que contribuyen al aprendizaje satisfactorio de sus alumnos (Simons, 1999; Vermunt y Verschaffel, 2000).

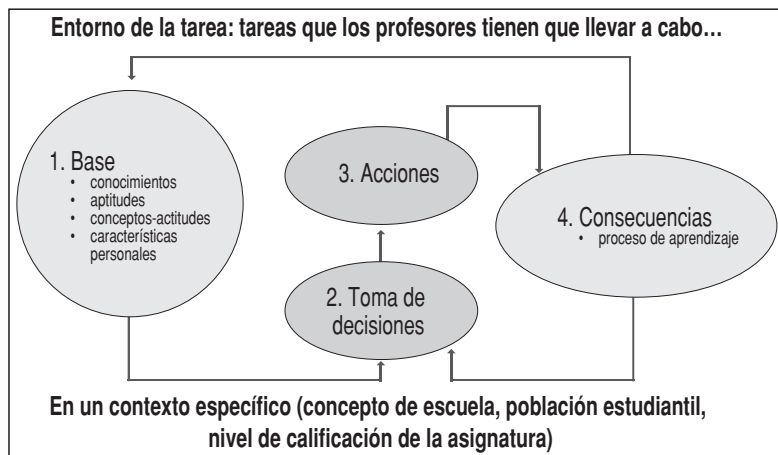
Aunque todos los conceptos de enseñanza cubren algunos aspectos de la competencia del profesorado, ninguno de ellos describe o explica detalladamente en qué consiste la competencia en la enseñanza. Volviendo a la definición descrita anteriormente, se necesita un concepto global y unificado de competencia en la enseñanza que tenga en cuenta todos los elementos que constituyen la competencia del profesorado, es decir, los rasgos personales del profesor, sus conocimientos, su comportamiento, su forma de pensar, la toma de decisiones en situaciones específicas

y las actividades de aprendizaje resultantes. Roelofs y Sanders (2003) han desarrollado un modelo de rendimiento competente para evaluar la competencia de los profesores que incluye los aspectos que se han descrito anteriormente. El modelo sigue la definición general de competencias descrita por Mulder (2001).

El punto de partida de este modelo, representado en la figura 1, es que la competencia del profesor se refleja en las consecuencias de sus acciones y, en particular, en las actividades de aprendizaje de los alumnos. Otros ejemplos de dichas consecuencias son el ambiente en el aula (tranquilo o alterado), la sensación de bienestar entre los alumnos y las buenas relaciones con los padres y los compañeros. A partir de las consecuencias pueden identificarse los demás elementos del modelo. En primer lugar, el componente "acciones" se refiere a las actividades profesionales, como la comunicación de conocimientos, los comentarios sobre el trabajo de los alumnos y la creación de una atmósfera de participación en el aula. En segundo lugar, todas las actividades del profesor se realizan en un contexto específico en el que éste debe tomar muchas decisiones, a largo plazo (planificación de futuro) o de forma inmediata ante una determinada situación en el aula (véase Doyle, 1983). Por ejemplo, los profesores tendrán que planificar la formación y adaptarla en función de las diferentes circunstancias (p. ej. distintos estilos de aprendizaje de los alumnos o diversas condiciones de organización). En tercer lugar, a la hora de tomar decisiones y llevar a cabo actividades, los profesores recurrirán a una base de conocimientos profesionales y a algunas características personales.

Cuando se evalúan diferentes ámbitos de competencia de los profesores, por ejemplo, la enseñanza y la organización del aula, ello implica que ya se han realizado inferencias interpretativas sobre los profesores (véase Kane, 1992). Cuando se combinan diferentes aspectos de la en-

**Gráfico 1:** Modelo interpretativo de competencias (basado en Roelofs y Sanders, 2003)



señanza en un modelo global de rendimiento competente existen más posibilidades de obtener inferencias válidas que cuando se utilizan modelos reduccionistas centrados en elementos aislados del proceso educativo.

### 3. Evaluación de la competencia de los profesores

Antes de describir las consecuencias generales de nuestro modelo para la recopilación de pruebas sobre la competencia de los profesores, comentaremos brevemente la importancia de la validez de constructo como concepto unificador para determinar la calidad de la evaluación. Podemos obtener numerosos requisitos de calidad específicos a partir de este marco de validez de constructo.

#### 3.1 Criterios para la validez de constructo de los instrumentos de competencia

Los requisitos de los instrumentos de competencia varían en función de la finalidad de la evaluación. Debido a las consecuencias para el candidato, se establecen requisitos más estrictos para los instrumentos que plantean grandes retos (selección, certificación) que para los instrumentos utilizados para el desarrollo profesional (Pelkmans, 1998).

Messick (1996) ha desarrollado el marco más amplio para determinar la calidad de los instrumentos. Messick indica que en cada formulario de evaluación es necesario tener en cuenta seis aspectos de validez de constructo:

- a) contenido,
- b) modelos teóricos y de procesos,
- c) estructura,
- d) generalización,
- e) aspectos externos,
- f) consecuencias.

El aspecto "contenido" se refiere a la relevancia y representatividad de la evaluación. La cuestión es: ¿dentro de qué límites se pueden extraer conclusiones de la evaluación? El aspecto "modelos teóricos y de procesos" determina hasta qué punto las tareas seleccionadas exigen acciones pertinentes por parte de un candidato y si se minimiza la influencia de los factores de constructo irrelevantes. El "aspecto estructural" indica si los criterios de rendimiento reflejan correctamente los criterios que los expertos utilizan y la precisión y coherencia con que se califica y evalúa el rendimiento.

La "generalización" indica hasta qué punto pueden extrapolarse las evaluaciones a todo el conjunto de tareas y entornos, por ejemplo.

Los aspectos externos de la validez determinan hasta qué punto los resultados de las mediciones convergen con otras medidas y constructos y hasta qué punto divergen de las mismas. El aspecto "consecuencias"

o validez consecuencial indica hasta qué punto el instrumento tiene efectos positivos o negativos y efectos indirectos sobre el aprendizaje del alumno y la enseñanza del profesor.

### 3.2 Recopilación de pruebas de competencia

Utilizando nuestro modelo global de competencia del profesorado como marco interpretativo para la evaluación y teniendo en cuenta los criterios para la validez de constructo, se pueden describir las consecuencias para la creación de un ámbito de contenido y de procedimientos con vistas a la recopilación de pruebas de competencia. Tres de los aspectos que propone Messick merecen especial atención: contenido, modelos teóricos y de proceso, y generalización.

#### 3.2.1 Desarrollo de un ámbito de competencia

De acuerdo con el aspecto "contenido" expuesto por Messick, el contenido de la evaluación debería ser pertinente y representativo de la profesión docente. Las revisiones de contenido se utilizan para establecer los límites dentro de los cuales se efectúan las inferencias sobre la competencia del profesorado. Para establecer un ámbito de competencia suelen adoptarse varios procedimientos complementarios: análisis empíricos de cómo actúan los profesores, consultas a profesores excelentes, investigación empírica sobre las variables que contribuyen a mejorar el aprendizaje y consultas a los comités de profesionales en activo (véase Verloop, Beijaard y Van Driel, 1998). La combinación de perspectivas científicas y prácticas contribuye a la aceptación y utilidad práctica de los instrumentos (Beijaard y Verloop, 1996; Duke y Stiggins, 1990; Uhlenbeck, 2002).

La base de toda evaluación debe contener una visión general de los aspectos de la competencia, las situaciones en las que éstos deben demostrarse y el grado deseado de dominio. Hay que responder a tres preguntas:

- a) ¿Cuál es el contenido crucial de la competencia?
- b) ¿Cómo se definen los criterios de rendimiento?
- c) ¿De qué modo pueden evaluarse los niveles de competencia?

##### 3.2.1.1 Selección del contenido

Para demarcar los ámbitos de competencia, se antoja necesario seleccionar lo que es característico de una actuación profesional adecuada y lo que es esencial para dicha actuación. Tomando como punto de partida nuestro modelo de competencia, las preguntas importantes son las siguientes:

- ¿qué se espera de los profesores y en qué situaciones de realización de tareas?
- ¿a qué grado de dificultad en la realización de sus tareas deben ser capaces de hacer frente los profesores?
- ¿que resultados de los alumnos ("consecuencias") cabe esperar de las actividades de los profesores?



- ¿mediante qué actuaciones y procesos de toma de decisiones podrían contribuir los profesores a los resultados de los alumnos?

Se han desarrollado distintas descripciones de ámbitos para la competencia de los profesores tanto en los Estados Unidos como en los Países Bajos. Danielson y McGreal (2000) distinguen cuatro áreas de actividad profesional amplias y pertinentes: planificación y preparación, enseñanza, ambiente en el aula y responsabilidades profesionales. Por lo que respecta al ambiente en el aula, afirman: “[...] dichas actividades y tareas establecen un ambiente relajado y respetuoso en el aula, que estimula una cultura de aprendizaje y crea un lugar seguro para asumir riesgos” (op. cit., p. 31). Sobre la enseñanza, escriben: “[...] Los profesores que sobresalen en el ámbito 3 [enseñanza] crean una atmósfera de entusiasmo sobre la importancia del aprendizaje y el significado de los contenidos” (op. cit., p. 32). Estas afirmaciones ilustran el enfoque centrado en las consecuencias de las acciones, más que en las acciones en sí mismas.

La descripción proporcionada por Danielson y McGreal también constituyó la base de los instrumentos de evaluación Praxis III, integrados en la serie denominada Praxis, desarrollada por *Educational Testing Service*, que medían, entre otras cosas, las prácticas de los profesores en el aula en cada una de las cuatro áreas mencionadas. En la evaluación Praxis III, las observaciones realizadas por el profesor en el aula y el comportamiento de los alumnos se combinan con entrevistas previas y posteriores a la observación (Dwyer, 1998), éstas últimas destinadas a analizar el proceso de toma de decisiones de los profesores.

El *American National Board for Professional Teaching Standards* (Comité nacional estadounidense de normas de enseñanza profesional) reduce la multiplicidad de tareas incluidas en la evaluación para la certificación de los profesores (de cursos avanzados) a las dimensiones de la experiencia en el campo de la enseñanza, como la improvisación, los retos y la pasión por enseñar y aprender (Bond et al., 2000).

En los Países Bajos, la fundación para la competencia en la enseñanza profesional (SBL), desarrolló recientemente, a petición del Ministerio de Educación, un conjunto de requisitos iniciales de eficiencia para los profesores de educación primaria y secundaria basados en siete ámbitos de competencia amplios. En la descripción de los requisitos, SBL empieza por el funcionamiento de las clases y la actuación individual de los alumnos de los profesores competentes, ilustrando un enfoque coherente de la enseñanza. Se han desarrollado los siguientes ámbitos:

- a) competencia interpersonal: aptitudes para crear un ambiente amistoso y de cooperación y una comunicación abierta;
- b) competencia pedagógica: aptitudes para crear un entorno de aprendizaje psicológicamente seguro para los alumnos, contribuyendo a su bienestar;
- c) competencia didáctica y conocimiento de la asignatura: aptitudes para guiar a los alumnos en la adquisición de los principios básicos de

las asignaturas impartidas en la escuela y el modo en que se pueden utilizar estos conocimientos en la vida diaria y en el trabajo;

- d) competencia organizativa: aptitudes para crear un ambiente basado en la disciplina de trabajo en las clases;
- e) competencia en la cooperación con los compañeros: aptitudes para engranar el trabajo propio con el trabajo de los compañeros y para contribuir a la organización de la escuela en general;
- f) competencia en la cooperación con el ambiente escolar: aptitudes para contribuir a la cooperación con las personas (padres) y organizaciones dentro del contexto de la escuela;
- g) competencia en la reflexión y el desarrollo: aptitudes para reflexionar sobre la propia competencia y para evolucionar de acuerdo con las nuevas exigencias y los avances dentro de la profesión.

Este conjunto de requisitos constituirá la base para muchos instrumentos de evaluación en los centros de formación del profesorado.

### 3.2.1.2 Criterios de rendimiento

Una vez descritos los ámbitos de competencia, una cuestión importante es cómo formular criterios que sirvan de referencia a la hora de juzgar el rendimiento de los profesores. De acuerdo con nuestro modelo, debería adoptarse un enfoque global en la formulación de criterios para evitar la tendencia a basarse excesivamente en las actividades aisladas de los profesores, en aspectos concretos de los conocimientos o en los resultados de los alumnos. En lugar de ello, estos elementos del rendimiento competente deben combinarse en los descriptores verbales de los criterios. Resulta fácil encontrar ejemplos de criterios unilaterales y, en ocasiones, tautológicos: "el profesor se explica con claridad", "elige material adecuado". De acuerdo con nuestro modelo de competencia (véase la figura 1), los criterios de rendimiento parten de actividades y resultados de aprendizaje adecuados entre los alumnos, de los cuales es posible inferir las acciones y decisiones correctas de los profesores. La aceptabilidad de las decisiones del profesor depende de la calidad de la base de conocimientos profesionales con respecto a situaciones de enseñanza específicas. Frederiksen et al. (1998) hablan de "criterios funcionales". Roelofs y Van den Berg (2005) presentan un ejemplo de criterios funcionales, tomado de la competencia docente de los profesores de guardería, cuya finalidad es la adquisición de conceptos por parte de los niños pequeños: "por medio de actividades de formación (preguntas, explicaciones, ejecución de tareas, debates) el profesor actúa satisfactoriamente cuando los niños realizan actividades que contribuyen a un entendimiento más profundo de un conjunto elegido de conceptos (p. ej. otoño)".

### 3.2.1.3 Niveles de rendimiento

En los casos en los que los criterios pueden considerarse definiciones del rendimiento competente en las distintas situaciones de realización de tareas, los estándares se centran en la calidad de las acciones y sus resultados. En lugar de plantear un debate sobre la elaboración de criterios de rendimiento, que no entraría dentro del ámbito del presente artículo, destacamos la importancia de disponer de un modelo interpretativo capaz de describir y explicar las diferencias en los niveles de rendimiento. Un modelo que tenga en cuenta las diferencias entre principiantes y expertos dentro de una profesión podría incorporarse a la validez de constructo de la evaluación de los profesores. El trabajo de Berliner (2001) en el ámbito del desarrollo de la experiencia resulta especialmente significativo. Berliner explica en qué se diferencian los profesionales expertos de los principiantes. Los expertos:

- a) destacan en su ámbito de especialidad y en contextos específicos;
- b) desarrollan una capacidad de automatización de las acciones que realizan con frecuencia;
- c) aprovechan mejor las oportunidades y son más flexibles;
- d) responden mejor a los requisitos de las tareas y a las situaciones cuando se trata de resolver problemas;
- e) representan los problemas de formas cualitativamente diferentes (más ricas);
- f) reconocen las pautas en las situaciones de trabajo con más rapidez y precisión;
- g) observan las pautas más significativas dentro del ámbito en el que tienen experiencia;
- h) dedican más tiempo a prepararse para resolver problemas y utilizan fuentes de información más ricas y más personales.

Las características que Berliner atribuye a los profesores expertos se utilizaron recientemente como modelo para investigar la validez de constructo del sistema de certificación NBTPS para el profesorado de enseñanza avanzada (Bond et al., 2000). Los resultados mostraron que, en la mayoría de las dimensiones, los profesores certificados superaron a los profesores no certificados.

### 3.2.2 Fuentes de las pruebas de competencia

En el desarrollo de instrumentos es preciso buscar la forma de obtener las mejores pruebas posibles de competencia de un candidato. Teniendo en cuenta los requisitos de validez relativos a la representación de los contenidos, a los modelos teóricos y de procesos subyacentes y a la generalización, "las mejores posibles" se refiere a la representatividad de las tareas y las situaciones de realización de tareas y al grado en el que se representan correctamente en la evaluación los supuestos procesos y efectos del rendimiento competente.

La primera elección en la recopilación de pruebas de competencia se refiere a la naturaleza de las mismas. Los tipos de pruebas básicos son:

documentación de clase, observación de la clase (en directo o grabada, y centrada en las acciones del profesor o en las actividades de los alumnos), registros del profesor (centrados en las acciones), entrevista de reflexión (centrada en los procesos de toma de decisiones), informe de reflexión (centrado en los procesos de toma de decisiones), pruebas de los alumnos (centradas en los resultados), prueba escrita del profesor (centrada en la base de conocimientos o en los procesos de toma de decisiones), pruebas multimedia del profesor (centradas en la base de conocimientos o en los procesos de toma de decisiones).

De acuerdo con nuestro modelo de competencia, todas las pruebas deberían registrarse e interpretarse en el contexto de situaciones de enseñanza específicas. Los instrumentos de competencia difieren en gran medida a este respecto.

Las observaciones de las clases pueden ser desde evaluaciones sin contexto basadas en visitas a las clases (“el profesor explica con claridad”) hasta informes narrativos de episodios de clase o secuencias de vídeo sin filtrar. Esto también se aplica a la recopilación de documentación de clase, que a veces abarca materiales de series completas de clases y otras veces consiste en extractos de lo que los profesores consideran su mejor trabajo. Los contenidos de la documentación también varían considerablemente. Pueden referirse a la planificación de la clase por parte del profesor, a (ejemplos de) los comentarios que éste hace a los alumnos o a (ejemplos de) lo que los alumnos asimilan de los comentarios de los profesores. En relación con todo lo anterior, se debe determinar quién está en la mejor posición para proporcionar las pruebas de competencia necesarias: los propios profesores, sus compañeros, los alumnos, los padres, los directores, expertos externos u otros. Según nuestro modelo, todos los participantes en la evaluación deben estar en posición de proporcionar pruebas representativas y convincentes de competencia sobre las consecuencias, las acciones de los profesores y los procesos de toma de decisiones. Peterson (2002) describe, para diferentes fuentes de datos, las ventajas y los inconvenientes de involucrar a cada uno de estos participantes.

La segunda elección en la recopilación de pruebas se refiere a la selección de un conjunto de tareas y situaciones de realización de tareas que pueda considerarse representativo, tanto desde el punto de vista cuantitativo como cualitativo, en el ámbito de competencia objeto del estudio. Es necesario responder a las siguientes preguntas:

- a) ¿Se ha elegido una situación o un conjunto de situaciones representativas de las acciones que se realizan en la situación profesional?
- b) ¿La tarea o tareas y la situación o situaciones de trabajo elegidas son pertinentes o decisivas para demostrar la competencia?
- c) ¿Cuál es la dificultad o complicación de la tarea o la situación de realización de la tarea?
- d) ¿Tiene el candidato la oportunidad de proporcionar las pruebas de competencia necesarias?

- e) ¿Cubre el conjunto elegido de tareas y situaciones de trabajo toda la variedad de tareas y situaciones de trabajo?
- f) ¿Es posible extrapolar las evaluaciones de las situaciones que se están midiendo a la situación de trabajo?

Las tareas y situaciones de realización de tareas difieren en cuanto al grado de autenticidad. Pueden darse situaciones de evaluación reales, reales simplificadas, simuladas y simbólicas. En las evaluaciones en situaciones reales, el candidato lleva a cabo las tareas cotidianas que realiza en la realidad, sin intervención alguna. Un ejemplo de situación real es impartir una clase al propio grupo de alumnos o supervisar las tareas individuales que realizan los propios alumnos. En esta situación pueden presentarse todas las tareas docentes, las cuales deben llevarse a cabo en el acto. El éxito o el fracaso en la realización de las tareas tiene consecuencias directas para los alumnos.

En una situación real simplificada, el candidato realiza una tarea real, aunque menos complicada de lo que es en la realidad (por ejemplo, impartir una "miniclase" a un grupo reducido de alumnos). En situaciones simuladas, las consecuencias directas para los alumnos y la posibilidad de "detener" la situación hacen que la evaluación también sea incompleta. La autenticidad de la prueba de competencia en las situaciones reales simplificadas o simuladas es menor que en las situaciones de trabajo real. Sin embargo, la ventaja es que resulta posible presentar tareas que no surgen a menudo en condiciones normales de trabajo. También existe la posibilidad de evaluar la competencia en situaciones simbólicas, que no se producen normalmente, y en las que no existe la presión del tiempo ni la inmediatez de la situación en clase. Hay que prestar especial atención a la recopilación de pruebas sobre las decisiones en distintas situaciones de trabajo. Aunque el nivel de autenticidad sea bajo, las situaciones descritas permiten abarcar una gran parte de las tareas y las situaciones de trabajo.

Por regla general, los desarrolladores de instrumentos de evaluación evitan las deliberaciones sobre la naturaleza y el alcance de las pruebas requeridas y pasan directamente al diseño de los instrumentos. Debido a ello, la importancia de las distintas fuentes de pruebas no resulta a veces suficientemente clara. Este problema se plantea a la hora de evaluar los expedientes (sin estructurar), dado que, en este caso, la naturaleza de la recopilación de las pruebas puede variar drásticamente. Los expedientes pueden contener pruebas directas, como material realizado en clase, logros de los alumnos o informes de reflexión, pero también productos que en sí mismos son resultados de evaluaciones, como resultados de exámenes escritos, cartas de recomendación y evaluaciones de los compañeros. Los expedientes resultan a veces muy útiles para agrupar varias pruebas de competencia. Sin embargo, la posibilidad de evaluar depende en gran medida de la estructura del expediente y de la admisibilidad y la posibilidad de analizar y puntuar las pruebas recogidas (Heller, Sheingold y Myford, 1998).

## 4. Análisis

Este artículo presenta un marco global para la evaluación de la competencia de los profesores que puede servir de base para llevar a cabo evaluaciones válidas. Este apartado comienza extrayendo algunas conclusiones, a continuación se analizan las posibles ventajas del modelo cuando se utiliza para elaborar argumentos interpretativos y concluye examinando las aplicaciones recientes del modelo en evaluaciones de expedientes en vídeo.

La primera parte del artículo está dedicada al desarrollo del modelo. Basándose en la revisión de la bibliografía, se llegaba a la conclusión de que no existe un marco científico sólido y general para determinar lo que constituye una enseñanza competente. Se han destacado algunos elementos relativos a la competencia del profesorado a lo largo de la historia de la evaluación de este colectivo profesional: rasgos de la personalidad que contribuyen al éxito de un profesor, elementos de conocimiento esenciales relacionados con los contenidos de las asignaturas, manera de pensar con respecto a una determinada disciplina, formas de comportamiento que contribuyen a mejorar el aprendizaje, conocimientos prácticos y teorías subjetivas de los profesores que determinan las acciones de éstos en situaciones de enseñanza específicas, y enseñanza entendida como el fomento de actividades que potencian el aprendizaje entre los alumnos.

Los distintos elementos de la enseñanza cubren algunos aspectos de la competencia de los profesores, pero ninguno de ellos describe o explica detalladamente en qué consiste la competencia en la enseñanza. Por tanto, se ha presentado un concepto global y unificado de competencia en la enseñanza que tiene en cuenta los diferentes elementos de la competencia de los profesores.

En resumen, el modelo establece que la competencia del profesorado se refleja en las consecuencias de sus acciones, y en particular en las actividades de aprendizaje de los alumnos. A partir de las consecuencias se han identificado los demás componentes del modelo. El componente "acciones" se refiere a las actividades profesionales que fomentan, entre otras cosas, el aprendizaje de los alumnos. El componente "toma de decisiones" quiere decir que un profesor debe tomar decisiones en el aula, a largo plazo o de forma inmediata, p. ej. decidir emprender o no determinadas actividades. Además, se ha señalado que la toma de decisiones, las acciones y las consecuencias tienen lugar en el contexto específico en el que los profesores desempeñan sus tareas profesionales. Finalmente, a la hora de tomar decisiones y llevar a cabo actividades, los profesores deberán recurrir a una base de conocimientos profesionales y a una serie de características personales.

En la segunda parte del artículo se describe la forma en la que el modelo puede ayudar a desarrollar ámbitos de evaluación y criterios de rendimiento y a recopilar pruebas de competencia. Asimismo se analiza

cómo se podrían cumplir tres de los criterios de Messick para la validez de constructo: contenido, modelos teóricos y de proceso y generalización.

Según nuestro modelo de rendimiento competente, los criterios deberían partir de actividades y resultados de aprendizaje adecuados entre los alumnos, de los cuales puedan inferirse actuaciones y decisiones adecuadas de los profesores. A la hora de definir los ámbitos de competencia se aplica cada vez más este enfoque, es decir, en lugar de centrarse en una línea de acción concreta, describir categorías amplias de actividades y desarrollar actividades de aprendizaje deseables. Asimismo, para distinguir los niveles de rendimiento deseables, es posible realizar comparaciones sistemáticas entre principiantes y expertos en la profesión.

En el desarrollo de instrumentos es preciso buscar la forma de obtener las mejores pruebas posibles de competencia de un candidato. Teniendo en cuenta los requisitos de validez relativos a la representación de los contenidos, a los modelos teóricos y de procesos subyacentes y a la generalización, "las mejores posibles" se refiere a la representatividad de las tareas y las situaciones de realización de tareas y el grado en el que se representan correctamente en la evaluación los supuestos procesos y efectos del rendimiento competente.

Al seleccionar las fuentes de las pruebas, debería tenerse en cuenta en qué medida está representado el proceso completo de rendimiento competente, es decir, las decisiones y actuaciones del profesor y las actuaciones de los alumnos. Por último, debe elegirse un conjunto de tareas y situaciones de trabajo que pueda considerarse representativo, desde el punto de vista cuantitativo y cualitativo, del ámbito de competencia objeto del estudio.

De acuerdo con Kane (1992), es posible utilizar el modelo para elaborar argumentos interpretativos que apoyen los juicios sobre la competencia del profesorado. Estamos de acuerdo con Kane en que la competencia casi nunca se puede demostrar y que un argumento interpretativo sobre la competencia del profesorado será, como mucho, creíble. Si los evaluadores son capaces de interpretar los resultados de la evaluación según los procesos planteados en nuestro modelo, se sostendrá el argumento interpretativo. Por ejemplo, al evaluar la calidad de la enseñanza, los evaluadores pueden interpretar los resultados de los alumnos basándose en el modo en que los profesores toman decisiones cuando desempeñan su actividad docente, así como en sus actuaciones y en las consecuencias que éstas tienen para los alumnos en un determinado entorno de enseñanza.

Una ventaja de este modelo es que las variaciones o los cambios en el concepto de aprendizaje no afectan a su estructura. Las diferentes visiones del aprendizaje como, por ejemplo, la formación programada frente al aprendizaje por el descubrimiento, se verán reflejadas en algún tipo de actividades de aprendizaje deseables, en un repertorio de actuaciones adecuadas y en los procesos inherentes de decisión por parte del profesor. Sea cual sea la visión sobre la enseñanza, a la hora de tomar sus de-

cisiones relativas a la recopilación de datos, los desarrolladores de las evaluaciones podrán basarse, de forma más consciente, en los procesos que les gustaría presentar en sus evaluaciones.

El modelo presentado podría constituir la base para el desarrollo profesional de los profesores, puesto que describe los procesos en los que éstos se ven involucrados, procesos que los profesores podrían modificar y adaptar a partir de los resultados obtenidos y como consecuencia de sus propias actividades de reflexión. De esta forma el modelo contribuiría a mejorar la calidad de los procesos de formación y de aprendizaje de los profesores.

Si se parte de un modelo global de competencia de los profesores, cabe preguntarse hasta qué punto todos los aspectos del rendimiento competente deben estar cubiertos en una tarea de evaluación. Es decir, si se utilizan diferentes tareas y diferentes fuentes de pruebas, ¿cómo se combinan todos para llegar a una única conclusión? Por último, deseáramos mencionar los resultados de un estudio reciente de la Universidad de Leiden y el Instituto Nacional para la Evaluación en materia de Educación de los Países Bajos (CITO) en el que se empleó nuestro modelo como base para la recopilación de pruebas de competencia. Utilizando expedientes grabados en vídeo, se recogieron varias pruebas de rendimiento competente en un mismo grupo de situaciones de aprendizaje. Se grabaron las actuaciones de los profesores (en el vídeo), los procesos de toma de decisiones (mediante entrevistas), las consecuencias observadas en los alumnos (en vídeo) y el contexto de la clase (situación) en el que se demostraron las acciones de aprendizaje (utilizando la documentación de clase). Se creó un sistema de puntuación que permitió a los evaluadores experimentados emitir juicios generales sobre la competencia en la formación. Los primeros resultados de un estudio piloto (Roelofs y Van den Berg, 2004, 2005) muestran que los expedientes constituyen una forma coherente y evaluable de recopilar pruebas, si bien los evaluadores no emplean todas las pruebas para extraer conclusiones. Los evaluadores tenían pruebas de la toma de decisiones de los profesores y del contexto de las tareas, pero no las utilizaron para emitir sus dictámenes. No obstante, al informar de sus conclusiones a los profesores, examinaron estas fuentes de pruebas para proporcionarles unos comentarios más completos. Así es como se utilizó el modelo general de competencia para ofrecer argumentos interpretativos.

## Bibliografía

- Beijaard, D.; Verloop, N. Assessing teachers' practical knowledge. *Studies in Educational Evaluation*, Vol. 22, nº 3, 1996, p. 275-286.
- Bellon, J. J.; Bellon, E. C.; Blank, M.A. *Teaching from a research knowledge base*. Nueva York: MacMillan, 1992.



- Berliner, D. C. Learning about and learning from expert teachers. *International Journal of Educational Research*, Vol. 35, 2001, p. 463-482.
- Bond, L. et al. *The certification system of the National Board for Professional Teaching Standards: a construct and consequential validity study*. Greensboro, NC: University of North Carolina, Center for Educational Research and Evaluation, 2000.
- Brophy, J.; Good, T. L. Teacher behaviour and student achievement. In Wittrock, M.C. (ed.). *Handbook of research on teaching*, 3rd ed., p. 328-375. Nueva York: MacMillan, 1986.
- Bruner, J. S. *The process of education*. Cambridge: Harvard University Press, 1963.
- Creemers, H. P. M. *Effectieve instructie: een empirische bijdrage aan de verbetering van het onderwijs in de klas*. 's-Gravenhage: Instituut voor Onderzoek van het Onderwijs, 1992.
- Danielson, Ch.; McGreal, T.L. *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision & Curriculum Development, 2000.
- Doyle, W. Academic work. *Review of Educational Research*, Vol. 53, 1983, p. 159-199.
- Duke, D. L.; Stiggins, R.J. Beyond minimum competence: evaluation for professional development. In Millmann J.; Darling-Hammond L. (eds). *The new handbook of teacher evaluation. Assessing elementary and secondary teachers*, p. 116-132. Newbury Park, CA: Sage, 1990.
- Dwyer, C. A. Criteria for performance-based teacher assessments: validity, standards, and issues. *Journal of Personnel Evaluation in Education*, Vol. 8, n° 2, 1994, p. 135-150.
- Dwyer, C. A. Psychometrics of Praxis III: Classroom performance assessments. *Journal of Personnel Evaluation in Education*, Vol. 12, n° 2, 1998, p. 163-187.
- Frederiksen, J. R. et al. Video portfolio assessment of teaching. *Educational Assessment*, Vol. 5, n° 4, 1998, p. 225-298.
- Getzels J. W.; Jackson, P.W. The teacher's personality and characteristics. In Gage, N.L. (ed.). *Handbook of research on education*. Chicago: Rand McNally, 1963.
- Haertel, E. H. New forms of teacher assessment. In Grant, G. (ed.). *Review of research in education* (p. 3-29). Washington, DC: American Educational Research Association, 1991.
- Heller, J. I.; Sheingold, K.; Myford, C.M. Reasoning about evidence in portfolios: cognitive foundations for valid and reliable assessment. *Educational Assessment*, Vol. 5, n° 1, 1998, p. 5-40.
- Kagan, D. M. Ways of evaluating teacher cognition: inferences concerning the Goldilocks principle. *Review of Educational Research*, Vol. 60, n° 3, 1990, p. 419-469.
- Kane, M.T. An argument-based approach to validity. *Psychological Bulletin*, Vol. 112, n° 3, 1992, p. 527-535.

- Klarus, R.; Schuler, Y.; Ter Wee, E. *Bewezen geschikt: een assessment-procedure voor aspirant leerkrachten (zij-instromers) in het PO en VO*. Wageningen: Stoas Onderzoek, 2000.
- Latham, A.S.; Gitomer, D.H.; Ziomek, R. What the tests tell us about new teachers. *Educational Leadership*, Vol. 56, nº 8, 1999, p. 23-26.
- McConnell, E. A. Competence vs. competency. *Nursing Management*, Vol. 32, nº 5, 2001, p. 14.
- Meijer, P. C.; Verloop, N.; Beijaard, D. Exploring language teachers' practical knowledge about teaching reading comprehension. *Teaching and Teacher Education*, Vol. 15, nº 1, 1999, p. 59-84.
- Merriënboer, J. J. G. van,; Klink, M. R. van der; Hendriks, M. *Competenties: van complicaties tot compromis*. La Haya: Onderwijsraad (Consejo de Educación de los Países Bajos), 2002.
- Messick, S. Validity in performance assessments. In Phillips, G.W. (ed.). *Technical issues in large-scale performance assessments*. Washington, DC: Department of Education, Office of Educational Research and Improvement, Estados Unidos, 1996.
- Mulder, M. *Competentieontwikkeling in organisaties: perspectieven en praktijk*. La Haya: Elsevier bedrijfsinformatie, 2001.
- Parry, S.B. The quest for competencies. *Training*, julio de 1996, p. 48-56.
- Pelkmans, A. *Evaluatie van leraren en schoolleiders*. Wageningen: Leers-toelgroep Agrarische Onderwijskunde, Landbouwwuniversiteit Wageningen, 1998.
- Peterson, K.D. *Teacher evaluation: a comprehensive guide to new directions and practices*. Thousand Oaks, CA: Corwin Press, 2002.
- Reynolds, M. What is competent beginning teaching? A review of the literature. *Review of Educational Research*, Vol. 62, nº 1, 1992, p. 1-35.
- Roelofs, E. C.; Van den Berg, E. *Videodossiers: leerkrachtbekwaamheid in beeld*. Artículo presentado en el seminario sobre investigación educativa de 2004 (onderwijsresearchdagen), celebrado en Utrecht, 2004.
- Roelofs, E. C.; Sanders, P. Beoordeling van docentcompetenties. In Mulder, M. et al. (eds). *Bevoegd, maar ook bekwaam?* Groningen: Wolters-Noordhoff, 2003, p. 277-299.
- Roelofs, E. C.; Van den Berg, E. *Images of teacher competence. Design characteristics of a multimedia design portfolio*. Artículo presentado en la reunión anual de la American Educational Research Association, Montreal, 2005.
- Shepard, L. A. Evaluating test validity. In Darling-Hammond, L. (ed.). *Review of research in education*. Washington, DC: American Educational Research Association, 1993, p. 405-450.
- Simon, A.; Boyer, E. G. *Mirrors for behaviour: an anthology of classroom observation instruments*. Wyncote, PA: Communication Material Center, 1974.
- Simons, P. R. Competentiegerichte leeromgevingen in organisaties en hoger beroepsonderwijs. In Schlusmans, K. et al. (eds). *Competentiegerichte leeromgevingen*. Utrecht: Lemma, 1999.

- Spencer, L.M.; Spencer, S.M. *Competence at work: models for superior performance*. Nueva York: Wiley, 1993.
- Stichting Beroepskwaliteit leraren en ander personeel. *Bekwaamheidseisen leraren*. La Haya: Stichting Beroepskwaliteit Leraren en ander onderwijspersoneel, 2003.
- Stodolsky, S.S. Classroom observation. In Darling-Hammond, J.; Millmann, L. (eds). *The new handbook of teacher evaluation: assessing elementary and secondary school teachers*. Newbury Park, CA: Sage, 1990, p. 175-190.
- Straetmans, G.J.J.M.; Sanders, P.F. *Beoordelen van competenties van docenten*. Utrecht: Programmamanagement Educatief Partnerschap, 2001.
- Tijssen, J.G.L. Hindernissen voor competentie management. *Opleiding & Ontwikkeling*, Vol. 11, n° 10, 1998, p. 17-23.
- Tijssen, J.G.L. Personele flexibiliteit in strategisch perspectief: de betekenis van employability voor organisaties en individu. In Poel, R.F.; Kessels, J.W.M. (eds). *Human resource development: Organiseren van het leren*, p. 103-116. Alphen aan den Rijn: Samsom, 2001.
- Tillema, H. *Assessment van competenties: van beoordelen naar ontwikkelen*. Alphen aan den Rijn: Kluwer, 2001.
- Tom, A.R.; Valli, L. Professional knowledge for teachers. In Houston R.W. (ed.). *Handbook of research on teacher education*. Nueva York: MacMillan, 1990, p. 372-392.
- Uhlenbeck, A. *The development of an assessment procedure for beginning teachers of English as a foreign language*. Leiden: ICLON Graduate School of Education, 2002.
- Verloop, N. Investigating teacher cognitions. *Journal of Curriculum Studies*, Vol. 20, 1988, p. 81-86.
- Verloop, N. *De leraar: reviewstudie uitgevoerd in opdracht van PROO*. Den Haag: NWO/PROO, 1999.
- Verloop, N.; Beijaard, D.; Van Driel, J.H. Beoordeling van docenten. *Pedagogische Studiën*, Vol. 75, n° 6, 1998, p. 171-187.
- Vermunt, J.; Verschaffel, L. Process oriented teaching. In Simons, P.R.J.; van der Linden, J.; Duffy, T. *New learning*. Dordrecht: Kluwer, 2000, p. 209-225.