

DISEÑO DE PROCESOS DE EVALUACIÓN DE COMPETENCIAS: CONSIDERACIONES ACERCA DE LOS ESTÁNDARES EN EL DOMINIO DE LAS COMPETENCIAS¹

Design of the process in the evaluation of competencies: considerations on standards in mastering skills

JESÚS M. JORNET MELIÁ, JOSÉ GONZÁLEZ SUCH,
JESÚS M. SUÁREZ RODRÍGUEZ Y M^a JESÚS PERALESS MONTOLÍO²
Universitat de València

En este artículo se analizan los conceptos de competencias y estándares de desempeño, así como una estrategia general para orientar el diseño y planificación de un sistema de evaluación de competencias. Se comentan algunos elementos conceptuales y metodológicos acerca de la determinación de estándares en el dominio de las competencias. En este punto se revisan los conceptos y roles de los criterios, estándares e indicadores de evaluación, y de la necesidad de orientar los procesos de evaluación de competencias sobre modelos de tipo cuantitativo que permitan graduar los niveles de desempeño con precisión. Finalmente, se analizan las implicaciones educativas y organizativas de la enseñanza por competencias y su evaluación, el papel de los estándares en este ámbito y las necesidades de replantear los modos de trabajo en el sistema educativo para acercarse a la enseñanza por competencias de una manera eficaz.

Palabras clave: *Diseño y planificación de la evaluación de competencias, Estándares, criterios e indicadores de evaluación del desempeño, Puntuaciones de corte, Niveles de logro.*

Introducción

La enseñanza por competencias se ha ido imponiendo como el enfoque pedagógico dominante en los últimos años. Nacido en el ámbito laboral, el concepto de competencia hace referencia a todos aquellos aspectos (habilidades, destrezas, actitudes...) que debe poseer un profesional para desarrollar su labor de una forma eficaz.

El concepto de *competencias*, si bien puede considerarse polisémico, y ha sido definido desde diversas perspectivas puede ser definido de una manera genérica que concite un amplio consenso. De este modo, nos referiremos a las competencias...

«... como un conjunto de habilidades, conocimientos, procedimientos, técnicas y actitudes, que una persona posee y que son necesarias para:

- a) Realizar las tareas que demanda una profesión en un determinado puesto de trabajo de una manera eficaz,
- b) resolver los problemas que surjan de forma autónoma, libre y creativa, y
- c) colaborar en la organización del trabajo y con su entorno sociolaboral» (Leyva y Jornet, 2006: 2).

El impacto que ha tenido este concepto en todos los niveles educativos es enorme. Si bien, también hay que decir, que se podría considerar todavía como un campo emergente³; dado que, aunque los principios teóricos que orientan esta perspectiva han sido ampliamente trabajados, aún se está desarrollando el corpus metodológico y tecnológico que permita aprovechar este enfoque de una manera eficaz. En el ámbito de la Formación Profesional, está resultando bastante claro abordar la enseñanza por competencias y su evaluación⁴. En los demás niveles (desde la Educación Primaria a la Educación Superior) —por la generalidad o por su carácter difuso e incluso complejo de las competencias a desarrollar— está resultando más complicado, y hace pensar —como posteriormente comentaremos— que los cambios a producirse para que esta aproximación cristalice de forma adecuada deben afectar más allá del aula, al modo en que se planifica el curriculum y organiza el sistema educativo. Una muestra de la brecha que se ha creado entre la orientación teórica del enfoque por competencias y sus aplicaciones es que se observa que la asimilación de las ideas que se han trasladado desde el enfoque de competencias de la formación profesional a otros niveles (como, por ejemplo, la Educación Primaria o la Superior) ha producido interpretaciones en ocasiones falaces. Por ejemplo, el énfasis en la práctica ha llevado a afirmaciones, en nuestra opinión, erróneas —«*más práctica, menos teoría*»—, cuando precisamente la dificultad del enfoque por competencias radica en que son constructos complejos en los que la integración desde la teoría a la práctica debe ser consistente y coherente (*no puede haber una buena práctica si no se*

sustenta en una buena integración cognitiva de los elementos teóricos que la sustentan); o la importancia del rol activo del estudiante (autoaprendizaje), que en ocasiones se observa que está implicando (en especial en la Educación Superior) malas interpretaciones por parte del profesorado, abdicando en parte de sus funciones de guía y acompañamiento del aprendizaje.

Estas deficientes interpretaciones también impactan de manera decisiva en los elementos a considerar en la evaluación de competencias y, en definitiva, en la definición de los estándares para llevarla a cabo. Estos procesos, necesariamente complejos (Barberá, 2005), son abordables si se estructuran de forma conveniente a nivel técnico-metodológico, pero hay que asumir precisamente su complejidad para trasladar a la práctica educativa las ventajas que puede conllevar el enfoque, ponderando su aplicabilidad según las áreas o niveles educativos en que se utilice.

Por su parte, el concepto de *estándares* tiene matices relacionados con el tipo de instrumento al que se refieran (Jornet y González Such, 2009; Jornet, González Such y Suárez, 2010). Es muy frecuente su uso en el ámbito de las pruebas estandarizadas (sistematizadas), y alude en este caso al sistema de interpretación de puntuaciones; sin embargo, también es usual en aproximaciones de tipo cualitativo, en las que actúan como un argumento de referencia para poder explicitar los criterios de evaluación que faciliten la comparación de la ejecución que muestra una persona con un referente que facilite la evaluación de la calidad del aprendizaje observado. No obstante, pese a la diversidad de acepciones que podríamos darle, también es cierto que podemos acercarnos a una definición común en los ámbitos de la medición y la evaluación de competencias.

Por estándar nos referimos a la especificación —u operativización— de los niveles de calidad que deben mostrarse como dominio de una competencia.

En definitiva, constituyen una explicitación cualitativa (o cuantitativa) de los niveles de desempeño que debe obtener una persona, para considerar que posee la competencia. Devienen, por tanto, de los criterios de evaluación que se hayan considerado como críticos para la observación del nivel de calidad del aprendizaje. Y plasman los criterios de interpretación de puntuaciones, o de evidencias, que se hayan obtenido en los diferentes instrumentos que se utilicen para la medición y/o evaluación de la competencia.

En consecuencia, los estándares sirven como nexo de unión entre un concepto de calidad (necesariamente cualitativo) acerca de las competencias y la medida (cuantitativa) o evidencia (cualitativa, cuasi-cuantitativa o cuantitativa) que se haya observado como resultado de aprendizaje (Jornet y González Such, 2009).

Las dificultades metodológicas que se dan en la determinación de estándares para la evaluación de competencias se centran en que éstos deben insertarse en un proceso de planificación y diseño de la evaluación que dé coherencia (*validez*) desde la definición del constructo (o competencia) hasta la recolección de evidencias y la interpretación de resultados, sin olvidar su adecuación a los modos y procesos de enseñanza-aprendizaje. Por ello, en este trabajo, delineamos un proceso global de planificación y diseño de la evaluación de competencias, como marco general en el que se integra la determinación de estándares. Posteriormente, nos referiremos a los conceptos implicados en este proceso. Es preciso señalar antes de adentrarnos en el planteamiento descrito que, si bien entendemos que la evaluación de competencias presentará diferencias metodológicas, organizativas y contextuales, según se trate de hacerlo en relación a un estudiante (evaluación en el aula), o en relación a una organización educativa (instituciones o sistemas) o programas, abordamos la reflexión intentando mantenernos en un plano genérico —aún a riesgo de situarnos en un enfoque poco concreto de reflexión, dado que entendemos que de manera sustantiva, desde el

punto de vista metodológico, los problemas de base son en parte similares y, con esa pretensión, se intenta aportar un *marco de referencia metodológico que pueda dar satisfacción a los usos de la evaluación de competencias* en diferentes contextos, sean para diseñar evaluaciones para su uso en el aula, o para usos más globales como la evaluación de programas, instituciones o sistemas educativos.

Unas notas previas de reflexión acerca de la planificación y diseño de la evaluación de competencias

Como señalamos en la introducción de este trabajo, la planificación y diseño del proceso global de evaluación de competencias actúa como marco de referencia en el que situamos los procesos de determinación de estándares.

- *Los procesos de planificación y diseño de la evaluación de competencias* (y posteriormente de determinación de estándares) *deben basarse en trabajos colegiados*; es decir, el procedimiento para conformar los elementos que se implican en la definición del juicio de valor, que involucra cualquier evaluación, debe sustentarse en el consenso intersubjetivo. La calidad del aprendizaje, el nivel de logro de cualquier competencia, necesariamente se apoya en apreciaciones subjetivas, por lo que su objetivación requiere de un trabajo colegiado de equipos de especialistas en currículo, investigación educativa y/o docentes. Según se trate de evaluaciones de competencias abordadas desde un centro educativo o para un propósito más general (como por ejemplo, la evaluación de sistemas), los comités de desarrollo de estándares tendrán una composición con mayor presencia de unos u otros colectivos.
- *Las competencias, como constructos complejos, suelen implicar diversos tipos de subdominios o subcompetencias* (habilidades o

destrezas, actitudes, valores, etc.), por lo que el proceso de planificación y diseño de cualquier competencia debe partir de un análisis cuidadoso de los subdominios competenciales que se integran en cada una de ellas.

- Debe tenerse en cuenta que, *en cualquier nivel educativo, se dan interrelaciones entre materias (disciplinas o asignaturas) desde las que se trabaje la misma competencia en un mismo grado educativo (interrelación horizontal), así como graduaciones de dominio (escalación vertical) entre niveles o grados educativos*. Una competencia no puede, en muchas ocasiones, desarrollarse únicamente en una disciplina o en un único nivel. Por ello, el análisis finalista relativo al posible nivel de desempeño a conseguir en cada grado educativo es fundamental para determinar adecuadamente los estándares en cada disciplina o materia.
- Por los mismos motivos que señalamos en el punto anterior, la evaluación de competencias, que requiere de una aproximación conceptual multidimensional, podrá y/o deberá estar basada en múltiples instrumentos (sean cuantitativos o cualitativos). Por lo que *cualquier estándar de calidad de una competencia se obligará a estar informado por la diversidad de instrumentos y evidencias que satisfagan las necesidades de su evaluación*.

Una estrategia metodológica para la planificación y el diseño de la evaluación de competencias

La planificación y diseño de una evaluación de competencias puede sustentarse desde diversas aproximaciones y debe estar asociada a la planificación y organización de la enseñanza. No obstante, es necesario asumir que en cualquier evaluación —y en este caso también ocurre— debe asegurarse la validez, lo cual depende del grado en que se haya realizado un proceso

riguroso en el que se compruebe que realmente lo que se evalúa representa de forma adecuada aquello que se enseña. Ambos conceptos no son disociables: planificación (diseño y desarrollo de la enseñanza) y evaluación. En cualquier caso, y pese a poderse dar diversas aproximaciones metodológicas para orientar una evaluación de competencias, puede identificarse una estrategia de trabajo que, de manera sustantiva, esté presente en cualquier aproximación evaluativa de calidad. Esta estrategia la presentamos anteriormente (Jornet, 2007), y aquí la sintetizamos, generalizamos y actualizamos. En el cuadro 1 se resume la estrategia global, de la que posteriormente comentamos brevemente sus aspectos más significativos.

Queremos destacar el papel potencial de la evaluación para clarificar e iluminar el mismo diseño curricular (Popham, 2006). Prever en los equipos de profesores o en la administración educativa un protocolo de evaluación de competencias como el que presentamos obliga a explicitar «*qué es lo que se quiere conseguir*», «*qué es considerado ser competente en...*», y dentro de ello «*cuál es la tarea específica para este nivel o grado*», e incluso cuestiona los métodos docentes, al hacernos reflexionar sobre la coherencia entre éstos, las competencias descritas para la materia y el plan de evaluación.

Nos referimos al *Dominio Competencial* como «*el conjunto de competencias que debe desarrollarse en una disciplina o titulación*». Se trata de determinar el *Mapa de Competencias (MC)* que se van a desarrollar en una asignatura —ver ejemplo, en cuadro 3-. Este análisis incluye tres elementos de clasificación:

- Clasificación de las competencias, según una tipología de referencia que puede ser la más difundida, como es el caso de la que se utiliza en el Proyecto Tuning (González y Wagenaar, 2003 y 2005) o cualquier otra adaptación como, por ejemplo, la de Leyva y Jornet, 2006, que se muestra en el cuadro 2.

- Identificación de subcompetencias, asimismo basada en un sistema de clasificación (en nuestro caso, a modo de ejemplo, diferenciamos entre conocimientos, habilidades/destrezas, actitudes y valores).
- Análisis de los componentes de aprendizaje de cada subcompetencia. También en este caso, es conveniente apoyarse en un

sistema de clasificación, lo más parsimonioso que sea posible⁵ como, por ejemplo, el derivado de la Teoría Componential de Merrill (Merrill y Twitchell, 1994) que distingue entre Niveles de Contenido (Hechos, Procedimientos y Principios) y Niveles de Rendimiento (Reconocimiento, Comprensión, Aplicación).

CUADRO 1. Esquematización de una estrategia de planificación y diseño de un proceso de evaluación de competencias (adaptado de Jornet, 2007)

Fase del proceso	¿Sobre qué se aplica?	Finalidad	Observaciones
Determinación y especificación del Dominio Competencial de la materia, que va a ser evaluado	Conjunto de la disciplina: todas las competencias implicadas	Definir todas las competencias que se trabajan en una asignatura	Es especialmente importante analizar las interrelaciones que se dan entre diversas asignaturas respecto a las competencias que se trabajan en cada una de ellas
Definición de criterios de evaluación	Cada competencia y subcompetencias implicadas	Determinar las características de calidad que se van a considerar en la evaluación	Deben orientar —y permitir— una evaluación de calidad (fiable y válida)
Especificación previa de un referente o estándar de ejecución, para determinar los criterios de evaluación		Operativizar y explicitar los criterios para evaluar las competencias y las subcompetencias	Deben estar formulados de forma clara y precisa, permitiendo una identificación de diversos niveles de competencia
Sistematización del proceso de recolección de evidencias		Se trata de identificar: <ul style="list-style-type: none"> • Las situaciones evaluativas, en las que se observará el aprendizaje, y • las tareas a realizar por parte del estudiante como expresión de su competencia 	Situaciones y tareas deben ser realistas, representando, en la medida de lo posible, funciones habituales en las que se pueda observar la competencia a evaluar
Determinación del sistema de evaluación, instrumentos, recursos y técnicas de recogida de información: indicadores		Se trata de definir qué instrumento/s se va/n a utilizar para cada competencia y/o subcompetencia. Identificar indicadores	Los instrumentos a utilizar deben aportar evidencias de calidad (<i>fiabes y válidas</i>) representativas del tipo de competencia a evaluar

CUADRO 1. Esquematización de una estrategia de planificación y diseño de un proceso de evaluación de competencias (adaptado de Jornet, 2007) (cont.)

Fase del proceso	¿Sobre qué se aplica?	Finalidad	Observaciones
Definir un sistema de comparación de evidencias con el estándar		Determinar un sistema que permita trasladar las valoraciones (cuantitativas o cualitativas) a puntuaciones	La graduación del desempeño en las competencias es un elemento clave a considerar en el sistema de evaluación. Por ello, la diferenciación entre niveles debe objetivarse de forma precisa: es necesaria una aproximación numérica que sintetice el grado de desempeño en una competencia
Determinación del proceso de formación de la decisión (por ejemplo, Competente vs. Aún no competente)	Conjunto de la disciplina: todas las competencias	Establecer las puntuaciones de corte que permitan diferenciar entre diversos niveles o grados de competencia	Un aspecto clave del sistema de evaluación de competencias es definir los elementos que diferencian entre niveles de competencia. La calidad de las puntuaciones de corte entre niveles implica directamente a la validez de la evaluación
Especificar los mecanismos de información a audiencias y/o retroalimentación para el estudiante		Definir los procesos por los que se informará al estudiante acerca de su nivel de competencia y, en todo caso, de las evidencias que pongan de manifiesto sus lagunas en el dominio global de la competencia que se señala en el estándar	Los estándares deben ser la referencia tanto para el profesorado como para el alumnado para valorar el proceso de adquisición de competencias. Así como para la sociedad para valorar el producto de sus organizaciones educativas
Diseñar los mecanismos o criterios de calidad del propio sistema de evaluación		Determinar las estrategias de evaluación para informar de la calidad del sistema de evaluación, y analizar sus posibilidades de mejora	Toda medida, y en consecuencia toda evaluación, contiene error; por lo que el sistema debe someterse a revisiones periódicas para mejorar su calidad

El término *estándar* se ha utilizado de forma prolija en la literatura de Medición, Evaluación e Investigación Psicológica y Educativa, con diversas acepciones. En este caso nos referimos a los estándares como «*la operativización de los criterios de calidad con los que juzgar las competencias*», y en consecuencia, como referente para determinar el sistema de interpretación de las puntuaciones que sintetizan el desempeño en una competencia y/o subcompetencia. Dado que destinamos el apartado siguiente de este trabajo de manera específica a los estándares, aquí únicamente vamos a señalar algunas de sus características principales, para ofrecer continuidad al discurso de revisión de la estrategia

de diseño del sistema de evaluación de competencias en el que nos situamos. De este modo, entendemos que los *estándares* (Jornet, 2007):

- Deben establecerse por consenso intersubjetivo, es decir, a través de un proceso colegiado (equipos de especialistas, docentes...).
- La calidad dependerá en gran medida de su claridad para distinguir niveles diferenciales de desempeño.
- Por lo que deben estar claramente especificados, y ser operativos para su uso como referencia en la medición y evaluación.

CUADRO 2. Estructura de análisis de competencias (adaptado de Leyva y Jornet, 2006)

Tipología de competencia	Definición de la categoría
Competencia conceptual (<i>Saber</i>)	Se refiere a la comprensión de los antecedentes, evolución y fundamentos teóricos de su ámbito profesional
Competencia metodológica (<i>Saber hacer</i>)	Se refiere al conjunto de habilidades necesarias para realizar de manera apropiada las tareas y actividades implicadas en el desempeño de las funciones laborales con un dominio o «maestría» en la metodología propia y de su aplicación en ámbitos específicos
Competencia social (<i>Saber estar</i>)	Se refiere a la comprensión del contexto político, social, económico y cultural en el que se lleva a cabo un proceso laboral, referido no solo al escenario de trabajo, sino a ambientes más amplios, tanto sociales como naturales
<ul style="list-style-type: none"> • Competencia contextual • Competencia en comunicación interpersonal 	Se refiere al uso efectivo de la comunicación oral y escrita de manera empática y responsable con los clientes, participantes, informantes y usuarios del producto laboral
Competencia personal (<i>Saber ser</i>)	Implica la capacidad de combinar la teoría y las habilidades técnicas en una práctica laboral ética y responsable.
<ul style="list-style-type: none"> • Competencia integrativa 	Requiere del desarrollo y consolidación del juicio profesional mediante no sólo la integración de la teoría y la práctica, sino de las experiencias del ejercicio profesional en los diversos ámbitos en que se practica
<ul style="list-style-type: none"> • Competencia adaptativa 	Implica la capacidad de anticiparse y adaptarse a cambios importantes en el ejercicio profesional. Requiere de: a) habilidades para ajustarse a nuevas condiciones, muchas de ellas inherentes a los cambios tecnológicos, b) una actitud de actualización permanente en todos los campos de interés de su profesión, y c) un interés por la investigación y los procesos de mejora que desarrollan los profesionales de su ámbito profesional

Los estándares actúan en este caso como el elemento inicial de referencia desde el que se deriva toda la estructura de medidas y evidencias que puedan servir como muestra de desempeño —ver cuadro 4—. Por ello, el *Sistema de recolección de evidencias* debe establecerse a partir de los estándares. En este caso, entendemos que dicho sistema hace referencia al conjunto de situaciones evaluativas y tareas en las que sea posible observar el logro en la adquisición de la competencia o subcompetencia objeto de

evaluación. De este modo, se trata de otra fase crucial en la definición del plan de evaluación que se diseñe a tal efecto. Resulta fundamental que tanto las situaciones evaluativas como las tareas que servirán de base de observación del desempeño, estén convenientemente ajustadas a las oportunidades de aprendizaje que se hayan ofrecido en el desarrollo del programa educativo (Dochy *et al.*, 2002; Wolf y Dietz, 1998; Brown y Glasner, 2003) —ver cuadros 5 a 7—.

CUADRO 3. Ejemplo de mapa competencial (Jornet, 2007).

Nivel 2 Subcompetencia	Competencia: realizar una programación de la asignatura en términos de competencias																	
	Nivel 1: Saber hacer																	
	Tipología de competencia																	
	Conocimientos						Destrezas / Habilidades						Actitudes / Valores					
Tipo de contenido			Nivel de desempeño			Tipo de contenido			Nivel de desempeño			Tipo de contenido			Nivel de desempeño			
H	Po	Pri	R	C	A	H	Po	Pri	R	C	A	H	Po	Pri	R	C	A	
C.1. Realizar una programación de la asignatura en términos de competencias																		
C.1.1. Comprender los sistemas de planificación de la enseñanza mediante competencias			X			X												
C.1.2. Realizar un análisis crítico de los componentes de una unidad didáctica de la asignatura (ej. Unidad 1)				X		X	X			X		X	X					X
C.1.3. Realizar un MC a partir de la unidad didáctica de la asignatura			X			X	X			X		X	X					X

Respecto a la *comparación de evidencias con el estándar* puede realizarse de forma cuantitativa o cualitativa. No obstante, lo cierto es que la complejidad de la medida y evaluación de competencias, aunque invita a usos de procesos cualitativos —aparentemente más sencillos de aplicar— es necesario asegurar la precisión de las decisiones. Por ello, en nuestra opinión, es conveniente que, en la medida de lo posible, se intente desarrollar un sistema cuantitativo o cuasi-cuantitativo que permita objetivar la distinción entre las graduaciones de desempeño. En este sentido, los problemas aquí implicados no son diferentes a los que se dan en el establecimiento de sistemas de interpretación de puntuaciones de las pruebas estandarizadas, si bien se hacen más complejos al

tener que integrar —de manera ponderada— informaciones de diversa índole respecto a cada competencia y/o subcompetencia. De este modo, la determinación de estándares es conveniente que anticipe los criterios explícitos contra los que deberá evaluarse el desempeño del alumnado. Si ello se ha realizado así, la traducción del sistema de comparación del nivel que el estudiante muestra en su desempeño, en relación al estándar, se simplifica convirtiéndose en un proceso de identificación de puntuaciones de síntesis de niveles individuales, y su interpretación a partir de su comparación con baremos en los que se hayan identificado puntuaciones de corte para diferenciar entre los diversos niveles de desempeño.

CUADRO 4. Ejemplo de definición de estándar (Jornet, 2007)

Subcompetencia	Estándar de Evaluación de la Competencia: C.1.3. Realizar un MC a partir de la unidad didáctica de la asignatura			
	Aún no competente	Aceptable	Competente	Muy competente
Realizar una clasificación de competencias según el sistema base (Tuning, adaptado)	• Desconoce el sistema de clasificación base	• Conoce el sistema de clasificación de competencias	• Conoce el sistema de clasificación de competencias	• Conoce y aplica de forma totalmente correcta el sistema de clasificación utilizado, asignando bien todas las competencias a las categorías de análisis
		• Define de forma adecuada al menos el 75% de competencias implicadas en la unidad • Muestra errores en la asignación de algunas competencias (nº de errores inferior al 50%)	• Define de forma adecuada todas las competencias implicadas en la unidad • Muestra errores en la asignación de algunas competencias (nº de errores inferior al 25%)	• Define y clasifica correctamente todas las competencias implicadas en la unidad

De este modo, el *proceso de formación de la decisión* acerca del nivel de desempeño observado en la/s competencia/s de la disciplina se apoya en un sistema de determinación de puntuaciones de corte que permitan distinguir entre los niveles de logro que se hayan previsto en los estándares. Por ejemplo, en una estrategia de trabajo como la que venimos presentando, si se establecen cuatro niveles de competencia (por ejemplo: Muy competente, Competente, Aceptable, Aún no competente), se deberían identificar tres puntuaciones de corte; es decir, la que diferencia entre «Aún no competente y Aceptable», la que lo hace entre «Aceptable y Competente» y la que diferencia entre «Competente y Muy competente».

Los dos elementos restantes de esta estrategia de evaluación de competencias hacen referencia a los mecanismos de información a las audiencias implicadas en la evaluación y/o retroalimentación para el estudiante, y a los de revisión del propio del sistema de evaluación.

Respecto a la *información a las audiencias implicadas y/o retroalimentación para el alumnado*, hay que tener en cuenta que tanto desde una perspectiva formativa como sumativa de la evaluación, siempre es preciso aportar información. En el primer caso, como guía para la mejora en el proceso de adquisición de las competencias, y en el segundo como elemento de acreditación y/o certificación de la cualificación o desempeño

CUADRO 5. Criterios para decidir acerca del sistema de recolección de evidencias para la evaluación de competencias (adaptado de Jornet, 2007)

Orientación de la evaluación	Decisiones implicadas	Recomendaciones
Respecto al sistema de evaluación a utilizar	<p>¿Para qué se va utilizar la evaluación (sumativa o formativa)?</p> <p>¿Utilizar métodos cuantitativos y/o cualitativos?</p>	<ul style="list-style-type: none"> Las competencias son constructos complejos que integran diferentes componentes (cognitivos y no cognitivos) Necesidad de complementariedad metodológica Partimos de que se debe dar primacía a la calidad del método o técnica para representar la competencia, frente a la orientación epistemológica o pragmática del procedimiento
Respecto a la situación evaluativa	<p>¿Qué tipos de situaciones evaluativas y tareas sobre las que observar el desempeño?</p> <p>¿Cómo ha de ser la situación evaluativa?</p> <p>¿Cuándo ha de recogerse la información?</p>	<ul style="list-style-type: none"> Considerar todas las opciones evaluativas y seleccionar entre ellas las más pertinentes para cada competencia La situación en que evaluemos a los estudiantes deberá estar adecuada al tipo y nivel de desempeño de la competencia a evaluar, y representar situaciones reales. A estos efectos, pueden diseñarse situaciones simuladas en las que se replique la realidad Identificar los momentos de recogida de información. Dos situaciones: a) momentos que informen del proceso gradual de adquisición de la competencia, y b) momento terminal en el que se hayan dado todas las oportunidades de aprendizaje para el logro de la misma

CUADRO 5. Criterios para decidir acerca del sistema de recolección de evidencias para la evaluación de competencias (adaptado de Jornet, 2007) (cont.)

Orientación de la evaluación	Decisiones implicadas	Recomendaciones
Respecto a las tareas que se plantean al estudiante para observar su desempeño	<p>¿Qué tareas debemos plantear al alumnado?</p> <p>¿Qué roles debe jugar cada participante en la evaluación?</p> <p>¿Es posible sistematizar —estandarizar— la tarea?</p>	<ul style="list-style-type: none"> Las tareas deben estar ajustadas a la situación evaluativa, por lo que las recomendaciones allí señaladas sirven en este caso Es importante sistematizar de manera pormenorizada la tarea: las instrucciones a dar al alumnado, los tiempos, materiales implicados, roles de participantes (si se trata de tareas que impliquen interacción con el evaluador o trabajos en grupo o en equipo...)
Respecto a la calidad de los instrumentos	<p>¿Qué instrumento utilizar para recabar información acerca de cada competencia o subcompetencia?</p>	<ul style="list-style-type: none"> <i>Objetividad</i>: está representada por la fiabilidad y validez que pueda otorgar el instrumento o técnica a la evaluación <i>Operatividad</i>: hace referencia a la eficacia, eficiencia, funcionalidad, y factibilidad del procedimiento <i>Ética</i>: está representado por la justicia y equidad en la evaluación, de forma que las informaciones que se aporten con el instrumento o técnica representen de forma adecuada no sólo el nivel del estudiante, sino también el que sea razonable según las oportunidades de aprendizaje dadas en la materia y que diferencie de forma adecuada entre los niveles de desempeño de los diversos estudiantes

logrado. Para usos formativos, se pueden utilizar tanto mecanismos informales de información, en el marco de la tutoría —si se trata de procesos de evaluación de estudiantes para su uso en proceso de enseñanza/aprendizaje— (a partir de entrevista, informes orales, reflexiones a partir de sus muestras de desempeño —exámenes, portafolios...—), como otros más formales (como, por ejemplo, informes individualizados o institucionales o de sistema —cuando se trata de este tipo de evaluaciones—). En el caso de usos sumativos, necesariamente estas informaciones deberían añadirse como informe individual al expediente o certificado del programa realizado por el estudiante. En cualquier caso, los estándares son la referencia ineludible para poder organizar esta información.

El proceso de evaluación de competencias de los estudiantes puede terminar aquí. Pero es imprescindible siempre prever la revisión del propio sistema de evaluación.

Por último, respecto a la estrategia global del diseño del sistema de evaluación, únicamente queremos mencionar que debe integrar mecanismos o criterios de calidad del propio sistema. Ningún sistema de evaluación (y en el caso de competencias, con mayor motivo, dada la complejidad de su evaluación) puede asumirse como un sistema infalible. Como señalamos en el cuadro 1, toda información contiene error. Ya lo puso de manifiesto Spearman al formular la Teoría Clásica de Tests, y es un principio que no conviene olvidar. La complejidad

de un sistema de estas características, junto a la gravedad de las decisiones que se tomen a partir de él, hace especialmente necesaria esta consideración. Siguiendo a Madaus (1988), este tipo de usos relativos a valorar personas son de alto impacto, pues implican decisiones acerca de los estudiantes. En evaluaciones de bajo impacto —como pueden ser

las de sistemas— es fundamental también esta consideración (recuérdese el tratamiento riguroso que se realiza de las pruebas del Proyecto PISA). Es necesario, por lo tanto, sustentarlo en procesos de validación (inicial y posterior, de forma periódica) que informen de la calidad del sistema y de los elementos a mejorar en el mismo.

CUADRO 6. Síntesis de instrumentos y técnicas evaluativas (Jornet, 2007)

Técnicas evaluativas	Observación	Directa	Registros observacionales Listas de control Anecdotario y ficha anecdótica Sistemas de signos y de categorías Escala de estimación ... <i>Sobre simulaciones, situaciones naturales o virtuales</i>
		Indirecta	Análisis de trabajos Escala de producción
		Participativa	Análisis de trabajos de grupo/equipo Análisis de elementos de interacción social
		Autoinformes	Individuales Colectivos
Instrumentos de medida	Instrumentos parcialmente estandarizados	Cuestionarios e inventarios	
		Entrevistas	
		Dirigidas a variables psicológicas o constructos educativos asimilables conceptualmente a rasgos psicopedagógicos	
	Pruebas estandarizadas normativas		
	Pruebas referidas a estándares	Dirigidas a rendimiento, logro, desempeño	
Instrumentos de integración docencia / evaluación	Portafolio Análisis biográficos	Cuasi-cuantitativos	
		Cualitativos	

CUADRO 7. Ejemplo de tabla síntesis del sistema de recolección de evidencias con expresión del sistema de ponderación (Jornet, 2007)

Indicador	Evidencias para la evaluación de la competencia: C.1.3. Realizar un MC a partir de la unidad didáctica de la asignatura			
	Prueba objetiva ev. externa (profesor)	Autoinforme autoevaluación (estudiante)	Cuestionario coevaluación (pares)	Entrevista ev. externa (profesor)
Realizar una clasificación de competencias según el sistema base (Tuning, adaptado)	Nº ítems: 10 Tipo de ítems: • Elección múltiple (5 alternativas) • Asociación Corrección: suma de aciertos	Portafolio (desarrollo del trabajo individual)	Portafolio (desarrollo de trabajo colectivo)	Portafolio (sobre trabajos: individual • colectivo)
<i>Momentos de recogida de información</i>	Final de: unidad, bloque y materia	Principio, 50% y fin del trabajo	Principio, 50% y fin del trabajo	Principio, 50% y fin del trabajo
Ponderación evaluación formativa	20	25	25	30
Ponderación evaluación sumativa	30	—	20	50

En este marco de trabajo, los *estándares* son el núcleo central de desarrollo de todo el sistema. En el apartado siguiente de este trabajo aportamos algunas consideraciones acerca de la determinación de los mismos.

Consideraciones metodológicas acerca de la determinación de estándares en el dominio de las competencias

Un primer elemento que es preciso considerar es la diferenciación conceptual entre: criterios, estándares e indicadores (De la Orden, 2000, 2007; Rodríguez Gómez, 2007). La definición de estos tres componentes se produce en interacción con la definición del objeto y finalidad

de la evaluación. Los *criterios de evaluación* hacen referencia a:

- La condición (o condiciones) que debe cumplir una determinada actuación o desempeño para ser considerada de calidad; por lo tanto, define un objetivo de calidad con relación a aquello que se pretende evaluar.
- Se apoya en un principio o axioma definido *a priori* sobre el que se puede emitir un juicio de valor.

Como elementos origen del juicio de valor final que sustentará la evaluación, deben —a su vez— cumplir unas ciertas *características de calidad*:

- a) Ser *explícitos, operativos*, que orienten hacia *elementos observables* del desempeño

de las personas en una determinada competencia.

- b) Por tanto, que cumplan con las características básicas de *evaluabilidad* y *factibilidad en cuanto a su medida*.
- c) *Que sean establecidos por consenso inter-subjetivo*, de manera participativa entre las audiencias implicadas (especializadas en el ámbito disciplinar a que afecte la competencia); deben ser *aceptados como objetivos (u objetivados)*, y por tanto aceptados como referencia inicial del sistema de evaluación.

Por su parte, los *estándares* —y tal como hemos definido anteriormente— representan de manera explícita la sistematización del concepto (o criterios) de calidad del desempeño en la competencia o subcompetencia a evaluar. Por tanto, operativizan los niveles de desempeño observables en el dominio de una competencia, por lo que representan —a su vez— metas de desempeño en la competencia o subcompetencia objeto de evaluación.

Dado que los estándares, en el uso evaluativo, orientan la decisión y juicio de valor final acerca del desempeño, su graduación dependerá del tipo de los usos evaluativos que deseen realizarse, de forma que podemos diferenciar entre estándares dicotómicos (competente/aún no competente) y policótomos (por ejemplo: muy competente, competente, aceptable, aún no competente). Obviamente, los más usuales son los segundos, mientras que los primeros serían aplicables a situaciones sumativas en las que únicamente se aspira a certificar un nivel mínimo de competencia en un programa o actividad educativa (ver cuadro 4).

Como en el caso de los criterios, los *estándares*, también deben cumplir características de calidad bien definidas en cuanto a su diseño y formulación. Las descritas anteriormente para los criterios, también son aplicables en este caso, si bien se enfatiza el hecho de su especificidad y operatividad como elementos de guía para la

toma de decisiones. Por ello, es conveniente que incluyan en su formulación referencias explícitas acerca del tipo de evidencias en las que se basará la evaluación (indicadores a utilizar, niveles de logro o desempeño, etc.), por lo que pueden incluir tanto referencias cualitativas como cuantitativas.

Entendemos los *indicadores* como «*aquellas informaciones, datos o resultados, susceptibles de una interpretación inequívoca o normalizada, que informan de la adquisición o grado de desempeño en la competencia*».

Dependiendo del tipo de competencia a evaluar, y del tipo de instrumento que se haya utilizado para recabar la evidencia de desempeño, podrán ser de tipo cuantitativo (índices, tasas, puntuaciones, ratios...), o cualitativo (argumentos, descripciones...); si bien, y siempre que sea posible —como señalamos anteriormente en este artículo— es preferible que sean cuantitativos, para facilitar la diferenciación entre niveles. Asimismo, podemos identificar indicadores simples (por ejemplo, puntuación en una prueba unidimensional de competencia) o compuestos (por ejemplo, síntesis ponderada de diversas pruebas o indicadores que miden o informan de diferentes subcompetencias de una misma competencia o un conjunto de competencias).

Como en los casos de criterios y estándares, en los *indicadores* también podemos señalar características propias de calidad. Las más genéricas son:

- a) deben estar bien descritos, y presentar garantías técnicas de calidad exigibles al instrumento del que provienen (por ejemplo, fiabilidad y validez en casos de pruebas estandarizadas);
- b) son preferibles, dadas las características graduales de las competencias, los que permitan interpretaciones sincrónicas y/o diacrónicas de la competencia evaluada, de forma que aporten información

- suficiente para usos formativos/diagnósticos en la evaluación de competencias, y
- c) es conveniente que estén relacionados técnicamente con otros indicadores afines con el nivel de desempeño en la competencia.

En términos generales, se podría señalar que *criterios* y *estándares* deberían desarrollar en esta perspectiva el papel inicial de orientación de la evaluación, por lo que se deben establecer *a priori*, siendo el origen y referencia de todo el sistema de planificación de la enseñanza y de su evaluación⁶. La selección y/o elaboración de instrumentos para la recogida de información, la identificación de indicadores, etc., encontrarían en esta línea de desarrollo lógico el sustento necesario para poder abordar la identificación de soluciones a problemas habituales en la evaluación de competencias, como:

- Que los instrumentos tengan una adecuada capacidad de discriminación entre los diversos niveles de desempeño. Por ejemplo, es frecuente encontrar, cuando se elaboran pruebas estandarizadas en las que posteriormente se establecen los estándares de interpretación de puntuaciones, que algunos de los niveles de desempeño quedan mal muestreados, de forma que su definición se apoya en pocos reactivos. Disponer de la referencia de los diversos niveles de desempeño orienta la elaboración de reactivos (en el caso de pruebas estandarizadas), o la selección de rúbricas de evaluación (para su uso en pruebas abiertas o en portafolios).
- Otro problema que se puede afrontar desde esta perspectiva es el hecho de la *escalación vertical de instrumentos*. Ciertamente, un problema habitual cuando se evalúan competencias es que a través de los diferentes grados educativos se producen disfunciones importantes. Por ejemplo, es difícil establecer la continuidad de los niveles de desempeño de un grado a otro, incluso en disciplinas bien

estructuradas como son las matemáticas. Aunque se utilizan técnicas estadístico-psicométricas para este propósito, lo cierto es que si las pruebas no están orientadas para grandes conjuntos curriculares (vgr. Matemáticas en Educación Secundaria), resulta casi imposible establecer referencias creíbles entre los diversos niveles de desempeño de cada curso. Así, sería ideal acercarnos a una interpretación global de grandes ciclos o etapas educativas, pudiendo identificar la correspondencia entre niveles de desempeño entre ellos (ver figura 1). La perspectiva que aquí proponemos constituiría la base lógica para aspirar a un acercamiento que hiciera factible abordar este tipo de problemática, con posibilidades de encontrar soluciones adecuadas.

- De igual modo, la identificación de puntuaciones de corte se facilitaría, al poderse derivar desde el punto de vista lógico, desde una graduación de la adquisición de competencias (dentro de un mismo nivel educativo, o entre ellos) previamente establecida.
- La metodología disponible para la determinación de estándares y puntuaciones de corte es amplia y diversa, y ha sido descrita de forma prolija en diversos trabajos, a los que remitimos al lector interesado: Sanders, 1994; De la Orden, 2000; Jornet y Backhoff, 2007; Cizek, Bunch y Koons, 2004; Jornet y González Such, 2009; Tourón, 2009; y Jornet, González Such y Suárez, 2010.

Con todo, el proceso necesario (más que el ideal) sería poder establecer un trabajo iterativo, de carácter circular⁷, desde la definición de los estándares, pasando por el diseño y/o selección de instrumentos o indicadores, hasta la identificación de puntuaciones de corte. De esta forma, se podría realizar un ajuste entre los estándares y las evidencias de información útiles para su evaluación, que permitiría una mejor definición del estándar, su medida y, en

FIGURA 1. Esquematación de la graduación de desempeño en evaluaciones por competencias diseñadas para grandes ciclos educativos (ND: nivel de desempeño)

Competencia X		Grado 1 Disciplina Y					
Estándar	ND1	ND 2	ND 3	ND 4			
					Grado 2 Disciplina Y		
Estándar			ND 1	ND 2	ND 3	ND 4	
						Grado 3 Disciplina Y	
Estándar				ND 1	ND 2	ND 3	ND 4
...				

consecuencia, su evaluación. Esta consideración comporta acercarnos a una perspectiva diferente de la habitualmente utilizada en la construcción de pruebas, y que identificamos en el marco de las *pruebas referidas a estándares*. Éstas, se sitúan necesariamente dentro de una perspectiva criterial, si bien el sistema de interpretación de puntuaciones no se sitúa como una consecuencia final de las cualidades psicométricas de las pruebas y su capacidad para representar niveles diferenciales en el logro del Dominio Educativo, sino que resulta el origen para elaborar las pruebas. Desde este punto de partida se aspira a que las pruebas presenten dichas cualidades psicométricas específicamente para poder medir un constructo que se ha definido de forma operativa en términos diferenciales de desempeño.

4. Discusión y conclusiones

La orientación de la enseñanza y evaluación de competencias, si bien ha irrumpido con fuerza en el panorama educativo internacional, no se ha llegado a desarrollar e integrar convenientemente en ninguno de los niveles educativos (desde la Educación Infantil y Primaria hasta la Superior). Esta afirmación la basamos en la observación acerca de que los procesos de planificación y organización de la educación, en definitiva, no han sido puestos es cuestión, sino que simplemente, y en todo caso, se intenta

trabajar desde las aulas cambiando la perspectiva metodológica que utiliza el profesorado, desde la de objetivos a la de competencias.

El problema, en nuestra opinión, es de mayor calado, y requeriría de un cambio profundo en la estructuración y forma de trabajar a través de todo el sistema educativo. Las competencias tienen un carácter gradual y se desarrollan a través de toda la vida de las personas, a partir de las influencias educativas y la experiencia laboral. Así, por ejemplo, las competencias matemáticas que requiere un ingeniero para poder diseñar un puente, no sólo se han desarrollado en su titulación superior, sino que comenzaron a desarrollarse desde el inicio de su educación matemática y continuarán desarrollándose en su vida profesional y con la formación continua. De este modo, la graduación de niveles competenciales es fundamental para poder orientar una enseñanza por competencias, y en consecuencia su evaluación.

¿Qué nivel de competencia se requiere en cada grado educativo en una determinada disciplina? ¿Las diferencias individuales dentro de cada grado educativo, qué continuidad tienen entre los grados? Son cuestiones básicas que deben clarificarse para el conjunto de un sistema educativo. Por ello, si se desea trabajar desde el enfoque de competencias, el diseño curricular debe estar vinculado de forma unívoca a estándares de desempeño en competencias, asegurando que

existe una continuidad entre grados y delimitando qué aportaciones se realizan desde cada disciplina para el desarrollo de cada competencia. En este sentido, el desarrollo de estándares para la evaluación del dominio de competencias puede actuar como un coadyuvante para la mejora del diseño curricular global del sistema, y en definitiva como una referencia necesaria para poder valorar la calidad del mismo.

La respuesta a este tipo de cuestiones pasa, desde nuestro punto de vista, por la necesidad de establecer estándares nacionales:

- para todas las disciplinas que se integran en el sistema educativo, desde la Educación Infantil a la Educación Secundaria Post-obligatoria, y
- para las titulaciones oficiales de la Educación Superior.

Como señalaba Popham (2006) en relación a las pruebas, existe un conjunto claro de decisiones en la enseñanza en las que pueden resultar un apoyo muy útil, y que entendemos que pueden generalizarse al uso de estándares: a) decisiones acerca de la naturaleza y el propósito del currículum, b) decisiones acerca del conocimiento de los alumnos, c) decisiones acerca de por cuánto tiempo debe enseñarse algo, y d) decisiones acerca de la eficacia y eficiencia de la educación. En el caso del sistema educativo preuniversitario está claro que es necesaria una revisión profunda que asegure una continuidad de niveles de desempeño en las diversas disciplinas que se extienden a través de los diversos grados. Y, en el caso de la Educación Superior, también resulta obvio que, además de la misma necesidad que acabamos de señalar, el marco de libertad interna para diseñar los currícula para las universidades españolas en que está derivando la implantación del Espacio Europeo de Educación Superior, hace necesario disponer de estándares competenciales finalistas (de egreso) evaluables para las titulaciones oficiales que sirvan de referencia a todos los implicados (administración educativa,

universidades, profesorado y alumnado) acerca de los niveles de competencia que representa poseer una titulación universitaria determinada. No estamos reclamando desde estas líneas una certificación posterior a la que realizan las universidades, sino únicamente la posibilidad de disponer de una referencia que oriente de forma más precisa su trabajo y los niveles de calidad competencial a que se aspira y se consiguen. Disponer de estándares, bien representados por instrumentos de medición y evaluación, conlleva ventajas indudables tanto para la evaluación en el aula (dado que el profesorado podría disponer de referencias externas contra las que poder valorar la evolución de sus estudiantes), como para evaluaciones de instituciones o sistemas educativos, al ofrecer una referencia más clara acerca de lo que se persigue y logra como producto educativo.

Otro aspecto no resuelto es la *evaluación de estudiantes en la enseñanza por competencias en nuestras escuelas en la actualidad*. La complejidad de las competencias como constructo a medir/evaluar, así como los aspectos que hemos comentado en las líneas anteriores, nos llevan a reflexionar sobre cómo abordar su evaluación. *¿Quién debe diseñar los sistemas de evaluación? ¿Qué rol deben jugar las evaluaciones de asignaturas? ¿Sería necesario diseñar evaluaciones terminales de ciclo o etapa educativa en las que se hubiera dado oportunidad de aprendizaje suficiente para el desarrollo de alguna competencia? ¿Evaluaciones externas y/o internas?* Son también cuestiones a responder.

En cualquier caso, la implantación de un sistema de enseñanza y evaluación de competencias reclama cambios también en las estructuras educativas. No vamos a entrar en las relativas a las que serían necesarias para hacer posible y mejorar la enseñanza. Nos referiremos únicamente, a las referidas a la evaluación de competencias. Las limitaciones de tiempo de trabajo —y en ocasiones de formación en metodología de evaluación— que tiene

el profesorado en cualquier nivel educativo, reclaman la necesidad de:

- La *creación de unidades de apoyo para el diseño y desarrollo de técnicas e instrumentos de evaluación de competencias*, que diera servicio a organizaciones educativas y al profesorado; podrían ser una ayuda para responder con nuevas formas organizativas las preguntas formuladas.
- *Conjugar en la evaluación de competencias en el aula diversas técnicas e instrumentos*. Es obvio que los estudiantes tienden a trabajar adaptándose a las pautas que nacen del modo en que son evaluados. Una competencia bien desarrollada implica generalización del aprendizaje, de forma que debería poder observarse el mismo nivel de desempeño a través de diversos sistemas de evaluación. Sin embargo, es cierto que para que la generalización se produzca, también es necesario que los estudiantes se ejerciten en su proceso de enseñanza a resolver problemas en diferentes formatos, y sean por tanto sometidos a diferentes situaciones y tareas evaluativas. Por tanto, sería necesario para mejorar también la enseñanza, diversificar los instrumentos de evaluación y para ello, conjugar las aproximaciones actualmente habituales en las aulas, con otras que representen modos de evaluar similares a las evaluaciones externas —como la mencionada—.
- *Disponer de instrumentos evaluativos estandarizados adecuados a estándares nacionales*,

que puedan servir de guía al profesorado para definir la situación de sus propios alumnos. Este elemento favorecería el desarrollo de la equidad educativa al permitir conocer de un modo más preciso cuáles son los niveles de desempeño deseables en cada disciplina en cada grado educativo. En conexión con el punto anterior, además, y desde la puesta en marcha del Proyecto PISA, algunos países —y en España, también en algunas comunidades autónomas— se han percatado de que los estudiantes estaban ofreciendo peores resultados en aquellos casos en que durante la enseñanza no se les había sometido a evaluaciones con pruebas estandarizadas —similares a las que se utilizan en PISA—. Ya disponemos de indicios acerca de que la no utilización de pruebas estandarizadas en el aula está resultando un *handicap* para mostrar mejores niveles en evaluaciones externas

En síntesis, el camino está iniciado. Los principios sobre la enseñanza y evaluación de competencias han sido presentados, pero estamos aún en un proceso en el que nos hace falta profundizar en su aplicación educativa. Ello requiere de innovaciones en las aulas, pero también en la organización global de las estructuras que hacen posible la educación a través del sistema educativo, desde la Educación Infantil a la Superior. En este proceso, los desarrollos de la evaluación, sin duda, y tal como señalan Popham (2006) o Stake (2006) pueden ayudar a mejorar la enseñanza.

Notas

¹ Trabajo desarrollado en el marco del proyecto MAVACO (EDU2009-13485: Modelo de análisis de variables de contexto para la evaluación de sistemas educativos), financiado por el MICINN, en la convocatoria de proyectos I+D (2010-2012).

² Departamento MIDE. Avda. Blasco Ibáñez, 30. 46010, Valencia. Tel.: 963864430. E-mail: jornet@uv.es, gonzalezj@uv.es, rodrigus@uv.es, perales@uv.es.

³ Son evidentes las dificultades que parte del profesorado de nuestro país aún tiene para cambiar su enfoque metodológico-didáctico, desde un planteamiento por objetivos a otro por competencias. Ello también es evidente, incluso, en los libros de texto: muchos de ellos aún mantienen su enfoque basado en objetivos e incluyen, como elemento innovador, alguna actividad al final del tema o unidad didáctica basada en competencias, o incluyen como elementos organizadores simultáneos objetivos y competencias, repitiendo prácticamente las mismas ideas.

⁴ El referente profesional es más claro y explícito en el ámbito de la formación profesional. La organización docente se orienta por el referente profesional desde el principio. Para ello, se analizaban las tareas a realizar por los profesionales, se derivaban de ello las funciones a desarrollar y en consecuencia se definían las competencias.

⁵ Algunos autores prefieren utilizar taxonomías de referencia que tengan una teoría de base mejor elaborada y en especial de carácter cognitivo (como, por ejemplo, pueden ser las taxonomías de Bloom, Gagné o, más reciente, la de Marzano). En nuestra opinión este tipo de usos es posible cuando son especialistas en diseño instruccional los que desarrollan los sistemas evaluativos estandarizados; sin embargo, aproximaciones más parsimónicas suelen ser más operativas, para abordar con éxito este tipo de enfoques en la práctica cotidiana.

⁶ Establecidos de forma explícita, deben ser conocidos por los diferentes actores del proceso educativo: profesorado (como referencias de las metas a que se aspira en una determinada materia o programa, para poder planificar y organizar el proceso de enseñanza, y valorar sus logros), y alumnado (para orientar su expectativa de logro hacia la consecución del mejor nivel de desempeño competencial), y sociedad (administradores educativos, familias..., como referencia funcional de los niveles individuales o colectivos, según del tipo de evaluación de la que se trate).

⁷ En este sentido, se parte de un proceso de juicio (definición de criterios y estándares) que origina la selección y/o diseño de instrumentos (identificación de indicadores) y que se comprueba/ajusta a partir de procesos empíricos basados en modelos de medida, retroalimentando la definición inicial.

Referencias bibliográficas

- BARBERÁ, E. (2005). La evaluación de competencias complejas. EDUCERE, vol. 31, pp. 497-504.
- BROWN, S. y GLASNER, A. (eds.) (2003). Evaluar en la Universidad. Problemas y nuevos enfoques. Madrid: Narcea.
- CIZEK, G. J.; BUNCH, M. B. y KOONS, H. (2004). Setting performance standards: Contemporary methods. *Educational Measurement: Issues and Practice*, 23(4), 31-50.
- DE LA ORDEN, A. (2000). Estándares en la evaluación educativa. Ponencia presentada en las primeras Jornadas de Medición y Evaluación (marzo, 2000). Valencia: Universidad de Valencia.
- DE LA ORDEN, A. (2007). Evaluación de la calidad de la educación. Un modelo sistémico como base para la construcción de un sistema de indicadores. *Seminario Internacional de indicadores educativos —Julio, 2007— (Memoria): Conceptos, metodologías y experiencias para la construcción de indicadores educativos*. Instituto Nacional para la Evaluación de la Educación (INEE): México.
- DELGADO, A. M. (coord.) (2006). Evaluación de las competencias en el espacio europeo de educación superior. Barcelona: J. M. Bosch Editor.
- DOCHY, F. et al. (2002). Nuevas vías de aprendizaje y sus consecuencias: una nueva era de evaluación, *Boletín de la Red Estatal de Docencia Universitaria (REDU)*, 2 (2), 1-24 (edición electrónica).
- GONZÁLEZ, J. y WAGENAAR, R. (ed.) (2003). *Tuning Educational Structures in Europe. Informe Final. Fase Uno*. Bilbao: Universidad de Deusto.
- GONZÁLEZ, J. y WAGENAAR, R. (ed.) (2005). *Tuning Educational Structures in Europe II. Universities' contribution to the Bologna process*. Bilbao: Universidad de Deusto.
- JORNET, J. M. (2007). La evaluación de los aprendizajes universitarios. Ponencia invitada en las III Jornadas de intercambio de grupos de formación del profesorado de la Universidad de Cádiz. UCA-Publicaciones: Cádiz (en prensa).

- JORNET, J. M. y BACKHOFF, E. (2007). *Modelo de determinación de Niveles de Logro de los EXCALE*. Cuadernos de Investigación, n° 30. México: INEE (disponible en <http://www.inee.edu.mx/index.php/publicaciones/resultados-de-investigacion/cuadernos-de-investigacion>, consultado 15-7-2010).
- JORNET, J. M. y GONZÁLEZ SUCH, J. (2009). Evaluación criterial: determinación de estándares de interpretación (EE) para pruebas de rendimiento educativo. *Estudios sobre Educación*, n° 16, 103-123.
- JORNET, J. M.; GONZÁLEZ SUCH, J. y SUÁREZ, J. M. (2010). Validación de los procesos de determinación de estándares de interpretación (EE) para pruebas de rendimiento educativo. *Estudios sobre Educación* (en prensa, aceptado y pendiente de publicación).
- LEYVA, Y. E. y JORNET, J. M. (2006). El perfil del evaluador educativo. *Boletín CENEVAL*, n° 7 (nov.), 2-6. México: CENEVAL.
- MADAUS, G. F. (1988). *The influence of testing on the curriculum*, en L. N. TANNER (ed.), *Critical Issues in curriculum. Eighty-seventh Yearbook on the National Society for Study of Education* (pp. 83-121). Chicago, IL: University of Chicago Press.
- MERRILL, M. D. y TWITCHELL D. G (1994). *Instructional design theory*. New Jersey: Englewood Cliffs. Educational Technology Publications.
- POPHAM, W. J. (2006). *Evaluar mejor para enseñar mejor*. Aula XXI, Santillana: México.
- RODRÍGUEZ GÓMEZ, G.; IBARRA, M. S. Y GÓMEZ, M. A. (2007). Guía para la elaboración de Catálogos de Indicadores para la Evaluación de Políticas Públicas en la Comunidad Autónoma de Andalucía. UCA-Publicaciones: Cádiz.
- SANDERS, J. R. (1994). *The Program Evaluation Standards*. The Joint Committee on Standards for Educational Evaluation. London: Sage.
- STAKE, R. (2006). *Evaluación comprensiva y evaluación basada en estándares*. Barcelona: Editorial Graó.
- TOURÓN, J. (2009). El establecimiento de estándares de rendimiento en los sistemas educativos. *Estudios sobre Educación*, 16, 127-146.
- WOLF, K. y DIETZ, M. (1998). Teaching portfolios: purposes and possibilities. *Teacher Education Quarterly*, Winter, 9-22.

Abstract

Design of the process in the evaluation of competencies: considerations on standards in mastering skills

This paper analyses the concepts of competencies and performance standards, as well as a comprehensive strategy to guide the development of the design and planning of a competency assessment system. We discuss conceptual and methodological elements on the determination of standards in mastering skills. At this point we review the concepts and roles of the criteria, standards and evaluation indicators, and the need to guide the process of competency assessment on quantitative models that allow graduate performance levels accurately. Finally, we analyse the educational and organizational implications of teaching and evaluation skills, the role of standards in this area, and needs to rethink the way to work in the education system in order to approach the teaching skills in an effective way.

Key words: *Design and planning skills assessment, Standards, criteria and indicators for performance evaluation, Cutoff scores, Achievement levels.*

Résumé

Cet article analyse les concepts de compétences et de standards de performance, aussi bien qu'une stratégie générale pour orienter la conception et la planification d'un system d'évaluation des compétences. Nous discutons quelques éléments conceptuels et méthodologiques sur la détermination

des standards dans la maîtrise des compétences. Dans ce point, nous révisons les concepts et les rôles des critères, des standards et des indicateurs d'évaluation, ainsi que la nécessité d'orienter les processus d'évaluation des compétences sur des modèles de type quantitative qui permettent de mesurer avec précision les niveaux de performance. Finalement, nous examinons les implications éducatives et organisationnelles de l'éducation par compétences et de son évaluation, le rôle des standards dans ce domaine, et les nécessités de reposer les formes de travail dans le système éducatif pour nous rapprocher à l'enseignement par compétences d'une façon efficace.

Mots clés: *Conception et planification de l'évaluation des compétences, Standards, critères et indicateurs d'évaluation de la performance, Scores de coupures, Niveaux de réussite.*

Perfil profesional de los autores

Jesús M. Jornet Meliá

Catedrático en el Departamento MIDE-UVEG. Coordinador del grupo GEM (MIDE-UVEG; www.uv.es/gem). Su trabajo se orienta en el área de Medición y Evaluación Educativas, al diseño de instrumentos para la evaluación de competencias, cuestionarios de contexto para la evaluación de sistema educativos, y la evaluación de la dimensión educativa de la Cohesión Social.

Correo electrónico de contacto: jornet@uv.es

José González Such

Profesor titular en el Departamento MIDE-UVEG. Coordinador de la Unidad InnovaMide del grupo GEM (MIDE-UVEG). Su trabajo se centra en la Medición y Evaluación Educativas, siendo sus líneas preferentes de investigación: diseño de pruebas e instrumentos de medición educativa, evaluación de la docencia y diseño y evaluación de materiales de innovación docente sustentados sobre nuevas tecnologías.

Correo electrónico de contacto: gonzalej@uv.es

Jesús M. Suárez Rodríguez

Profesor titular en el Departamento MIDE-UVEG. Coordinador de la Unidad de Tecnología Educativa (MIDE-UVEG). Su trabajo se centra en los elementos metodológicos, analíticos y de medición en los ámbitos de Ciencias Sociales y Salud. En los últimos años su línea de investigación prioritaria se centra en el impacto de las TIC en los diferentes niveles educativos.

Correo electrónico de contacto: jesus.m.rodriguez@uv.es

M^a Jesús Perales Montolío

Profesora titular en el Departamento MIDE-UVEG. Coordinadora de la Unidad de Evaluación Socio-educativa del grupo GEM (MIDE-UVEG). Sus líneas centrales de investigación se dirigen a la evaluación de programas socioeducativos, evaluación institucional, evaluación de la formación ocupacional y continua, y el diseño y desarrollo de instrumentos de evaluación de competencias.

Correo electrónico de contacto: perales@uv.es