

La autenticidad de la evaluación¹

Dr. Carles Monereo Font

Introducción

En la actualidad ningún investigador educativo se atrevería a discutir la decisiva influencia que tiene la evaluación sobre los procesos de enseñanza y de aprendizaje. Fredericksen (1984), y más recientemente Prodromou (1995), han demostrado que la evaluación tiene una influencia decisiva en el contenido de lo que enseñan los docentes, incluso en relación a apartados del temario que el profesor considera menos importantes pero cuya presencia en las pruebas de evaluación los convierten en centrales.

En el caso del aprendizaje ocurre algo similar. La evaluación, o mejor dicho, la percepción que tiene el estudiante de las características de esa evaluación, resulta determinante en el modo de enfocar el estudio de la materia (Gibbs, 1992; Wolf et al., 1991). Gulikers, Bastiaens y Kirschner (2004) lo resumen de forma contundente: *“Aprendizaje y evaluación son dos caras de la misma moneda, e influyen fuertemente el uno en la otra. Para cambiar el aprendizaje del alumno en la dirección del desarrollo de competencias auténticas es necesaria una enseñanza basada en competencias auténticas, alienada con una evaluación basada también en competencias auténticas”* (p. 68). Modificando la manera en que evaluamos aquello que aprenden nuestros alumnos tenemos la posibilidad de modificar lo que realmente aprenden y, consecutivamente, también tenemos la oportunidad de modificar el modo en que se enseña lo que aprenden.

Las dos razones más importantes para aproximarse a una *“authentic competency-based assessment”* son, en palabras de Gielen, Dochy y Dierick (2003), la validez de constructo y la validez consecuencial. La primera tiene que ver con la necesidad de valorar la naturaleza misma de lo que se pretende medir. Si hablamos de competencias, entendidas como un conjunto coordinado de recursos (conocimientos, habilidades, estrategias) que nos permiten resolver problemas prácticos, ya sea de la vida cotidiana o profesional, su evaluación deberá, necesariamente, plantear problemas semejantes a los de referencia. Por consiguiente deberemos presentar al alumno tareas situadas, de carácter práctico, que requieran la puesta en marcha de estrategias de resolución.

La validez consecuencial se refiere precisamente a los efectos de la evaluación sobre la enseñanza y el aprendizaje. Cuando el estudiante percibe una relación coherente entre lo que se explica, la forma y el nivel con que se hace, y lo que se demanda, es decir la manera y la exi-

¹ MONEREO, C. "La autenticidad de la evaluación" en CASTELLÓ M. (Coord) (2009) La evaluación auténtica en enseñanza secundaria y universitaria, Barcelona, Edebé, Innova universitat.

gencia con que se evalúa, se produce un impacto positivo sobre la manera en que este estudiante se enfrenta al estudio de esa materia en cuestión (Seges, Dierick y Dochy, 2001).

Deberíamos añadir una última razón, no menor; al aumentar la autenticidad de las evaluaciones, incidimos positivamente en la motivación de los alumnos como muestran los estudios de Herrington y Herrington (1998) o Sambel, McDowell y Brown (1997). Pero, ¿cómo definimos y cuales son las características estructurales del concepto de “evaluación auténtica”?

0. El concepto de evaluación auténtica

La expresión “evaluación auténtica” suele emplearse de manera genérica para describir una amplia variedad de nuevos enfoques e instrumentos de evaluación que se contraponen a los utilizados reiteradamente en la evaluación tradicional, la denominada “*testing culture*”, la cultura del examen (Álvarez, 2005; Angelo y Cross, 1993; Monereo, 2003). Entre estos enfoques suelen citarse la evaluación no formal o alternativa (*alternative assessment*), la evaluación del proceso de realización (*performance assessment*), la evaluación basada en problemas (*problem-based assessment*), la evaluación dinámica, la evaluación formadora o la evaluación formativa (ver al respecto Campbell, 1994; Custer, 1998; Darling-Hammond, 1995; Darling-Hammond y Snyder, 2000; Nunziati, 1990; Sternberg y Grigorenko, 2003; Weber, 1999; Weinstein y Meyer, 1998). Las tres primeras podríamos situarlas dentro del paraguas de la evaluación auténtica, no así las tres últimas que tienen algunas características diferenciales que hemos tratado de sintetizar en la Tabla 1.

Centrándonos en el término “auténtico”, el diccionario² establece que algo es auténtico cuando es cierto u original, es decir no es falso, y se corresponde con lo esperado o buscado, por lo tanto existen pruebas que atestiguan o demuestran su veracidad. Como afirman Honebein, Duffy y Fishman (1993), la garantía de autenticidad solo puede establecerse en comparación con alguna otra “realidad” (Álvarez, 2005).

En todo caso existen distintas interpretaciones sobre este tópico, que de forma resumida expresamos en la siguiente lista:

a) Auténtico en calidad de evaluación **genuina o legítima** al tratarse de una evaluación propia a la forma de evaluar del profesor. En este sentido podríamos hablar de proximidad ecológica en cuanto a que parte de los recursos “naturales” del profesor.

b) Auténtico en calidad de **propedéutico**: como preparación para enfrentar evaluaciones similares en esta u otra materia de éste o de próximos cursos.

c) Auténtico en calidad de **funcional**: útil para resolver necesidades del alumno en sus diferentes escenarios de desarrollo (también en el académico).

d) Auténtico en calidad de centrado en la evaluación de la **actuación o realización** de una tarea (*performance-assessment*): orientado a valorar las operaciones necesarias para solucionar el problema o tarea.

e) Auténtico en calidad de **extra-académico**: La situación o problema que se plantea está vinculada a escenarios no escolares (vida cotidiana, noticias de impacto social y cultural, futuro profesional, ocio, etc.). Los problemas intra-académicos (intramatemáticos, intralingüísticos o intracientíficos), que aludirían a cuestiones propias del universo disciplinar (p.e. símbolos, estructuras y objetos matemáticos) serían, desde esta perspectiva, poco auténticos.

² P.e. Diccionario Sopena.Barcelona, 2006

f) Auténtico en calidad de **verosímil**: la situación que se plantea al alumno podría realmente ocurrir, aún cuando no necesariamente al propio alumno.

g) Auténtico en calidad de **real**: la actividad desarrollada no solo se produce en condiciones prácticamente idénticas a las que habitualmente conocemos, sino que además se trata de una actividad en la que el alumno o ya ha estado involucrado en la vida cotidiana o lo estará en un futuro próximo.

De entre estas diferentes perspectivas y acepciones las que sin duda han tenido mayor eco en la literatura especializada han sido las que hacen equivalente la evaluación auténtica con una evaluación del proceso de ejecución - *performance-assessment*- (p.e. Hart, 1994; Torrance, 1995) y las que han subrayado el realismo de la tarea que se propone (Herrington y Herrington, 1998). Con el propósito de dilucidar la cuestión Reeves y Okey (1996) afirman que la diferencia sustancial entre ambos enfoques está en el grado de fidelidad de las condiciones en que se ejecuta la tarea. Cuanto más fielmente se reproduzcan las condiciones en que esa tarea tiene lugar en su contexto habitual, más auténtica será la evaluación. Por lo tanto toda evaluación auténtica incluye una evaluación del proceso de realización, pero no toda evaluación de proceso es auténtica.

En este mismo sentido Savery y Duffy (1995) dan una vuelta más de tuerca y definen la autenticidad de una tarea como la similitud, no solo en las condiciones de ejecución, sino especialmente entre las demandas cognitivas requeridas en la evaluación y las solicitadas en el contexto en el que esa tarea o problema suele producirse (Calfée, 1994). Por lo tanto cuanto más se confunda una tarea de evaluación con su homónima en un contexto de práctica social, más auténtica podremos considerarla. Demostrar que algo es auténtico, autenticarlo o refrendarlo, supone aportar pruebas de que eso pertenece al mismo sitio, funciona del mismo modo, obliga al mismo esfuerzo o tiene la misma importancia que el original.

Tratando de resumir las principales diferencias entre una evaluación tradicional y una evaluación auténtica, la siguiente Tabla 2, adaptada de Wiggins (1990), esquematiza las principales discrepancias entre ambos enfoques.

Evaluación tradicional	Evaluación auténtica
Para evaluar, emplea ítems indirectos	Se examina directamente la ejecución del aprendiz en base a tareas relevantes
Revela solo si los estudiantes reconocen aquello que han aprendido en las aulas, fuera de ellas	Requiere que los estudiantes apliquen el saber adquirido a situaciones de la vida extraescolar
Las pruebas convencionales suelen ser de "lápiz y papel" y las preguntas a menudo suponen una única respuesta	Se evalúan a la vez un conjunto de competencias y contenidos necesarios para resolver aquella situación
Suelen tenerse en cuenta solo los resultados	Se requiere la justificación argumentada de las respuestas
La validez se determina mediante el apareamiento de ítems con el contenido del currículum	Se evalúa la capacidad de actuar en contextos sociales de manera idónea
Se evalúan elementos estáticos, y arbitrarios, que pueden resolverse casi siempre a través de la aplicación mecánica de principios y formulaciones.	Se proponen problemas mal estructurados, ambiguos, que reflejen la complejidad del mundo extraacadémico.

Tabla 2. Diferencias entre evaluación tradicional y auténtica (adaptado de Wiggins, 1990)

Como puede observarse una evaluación auténtica se caracteriza por valorar especialmente el proceso de decisión necesario para resolver un problema complejo, en el que deban activarse y aplicarse coordinadamente distintos conocimientos y competencias, y haya que demostrar la capacidad estratégica imprescindible para autorregular la propia conducta y ajustarse a los cambios inesperados, justificando posteriormente las acciones efectuadas. En el próximo apartado profundizaremos en las dimensiones que debe contemplar una tarea de evaluación auténtica.

1. Análisis de la autenticidad de una tarea

Analizar la autenticidad de una determinada actividad o tarea de evaluación significa situarla en un continuo de mayor o menor proximidad a la actividad que queremos en realidad evaluar. Ese continuo vendría definido por el realismo de la tarea, su relevancia académica, personal o profesional, su proximidad a las prácticas usuales del centro y el grado en que permite la socialización dentro de una determinada comunidad de práctica. Algunas propuestas como la de Díaz Barriga (2003), emplean un sistema de coordenadas, donde las ordenadas constituyen el grado de relevancia funcional y las abscisas el nivel de realismo que se otorga a la actividad o tarea en cuestión. A estas coordenadas podemos añadir dos dimensiones que nos parecen importantes, y que acabamos de citar, el grado de cercanía o proximidad ecológica de la tarea con el contexto educativo en el que pretendemos inscribirla y aplicarla, y el nivel de enculturización identitaria (*identity enculturation*) que favorece. Seguidamente trataremos de profundizar en el significado de estas cuatro dimensiones que resultan clave en toda práctica auténtica de evaluación, a partir de los aspectos en que converge la literatura especializada (Cronin, 1993; Gijse-laers, 1996; Hill, 1998; Gulikers, Bastiaens, y Kirschner, 2004; Hung & Der-Thanq, 2007):

1.1. Realismo

Antes de nada es preciso matizar el uso que se hace del término “real”. Todas las comunidades sociales en las que se desarrollan las personas, incluida por supuesto la comunidad educativa, son “reales” en el sentido de que las prácticas que allí se realizan son genuinas y permiten a cualquier participante socializarse y llegar a ser un miembro plenamente integrado y activo en esa comunidad. Más específicamente, siguiendo a Engeström (2005), cada comunidad tiene un conjunto de creencias, actitudes, normas y roles que suponen orientaciones epistemológicas hacia el conocimiento y constituyen identidad.

Cuando nos referimos a prácticas auténticas en los centros educativos debemos distinguir entre aquellas que se inscriben en el contexto propiamente escolar y se orientan a socializar al individuo en calidad de alumno, de aquellas otras que se orientan a hacer lo mismo pero en el ámbito profesional. En realidad deberíamos hablar de un continuo de prácticas que permitirían a un alumno transitar hacia una identidad profesional, y viceversa, caso por ejemplo de una persona del mundo laboral que, después de muchos años, decidiese incorporarse nuevamente a una comunidad educativa.

Como ya apuntamos con anterioridad, el realismo de una tarea de evaluación vendrá definido por el grado de fidelidad de las condiciones y exigencias cognitivas en que se realiza la actividad, con las condiciones y exigencias que demandan esas mismas actividades en los contextos sociales o profesionales de referencia.

Determinar el realismo de una tarea de evaluación supone, en primer lugar, haber identificado aquellos problemas profesionales que el alumno deberá ser capaz de resolver en el futuro y, en segundo lugar, analizar el conjunto de competencias, estrategias y conocimientos que el alumno debe haber adquirido para hacerlo.

En cuanto a la tipología de problemas a los que un profesional deberá hacer frente, hemos descrito en otro sitio (Monereo y Pozo, 2007) dos grandes grupos de problemas:

1) *Problemas prototípicos*: estaríamos hablando de tareas que por su frecuencia resultan habituales en el trabajo de un determinado profesional. Desde el currículo universitario debería garantizarse que el estudiante, cuando termine sus estudios de grado, será competente para resolver ese tipo de problemas.

2) *Problemas emergentes*: se referiría a dos tipos de demandas profesionales. En primer lugar a aquellas que, si bien resultan poco frecuentes en el momento actual, existen suficientes pruebas (estudios de prospectiva, sociológicos, de mercado, etc.) de que su incidencia se incrementará en un futuro próximo. En segundo lugar deberían incluirse aquellas situaciones, fenómenos o sucesos que, si bien la sociedad no considera aún conflictivas o problemáticas, inciden negativamente en el desarrollo de las personas, y es obligación de la universidad hacer visible su nociva influencia. En la Tabla 3 se muestra un ejemplo de cuales podrían ser, a grandes rasgos, algunos de los problemas prototípicos y emergentes en los distintos niveles educativos (en la universidad se toma el caso de la licenciatura de Psicopedagogía).

	Enseñanza obligatoria	Enseñanza Postobligatoria y grado (Psicopedagogía)	Postgrado Universidad (Psicopedagogía)	Doctorado
Finalidad	Formar ciudadanos	Formar profesionales técnicos	Formar profesionales especialistas	Formar investigadores
Problemas Prototípicos	Seleccionar y comprar un producto	Realizar una adecuación curricular	Desarrollar un proyecto de innovación	Enmarcar teóricamente un estudio
Problemas emergentes	Principios ecología doméstica	Búsqueda especializada en internet	Cooperación redes inter-profesionales	Analizar el discurso de una muestra
<u>Competencias</u>	Evaluación de necesidades y productos	Evaluación y coordinación de actores y recursos	Análisis sistémico del contexto	Análisis de datos cualitativos

Tabla 3. Un ejemplo de problemas y competencias según la finalidad de cada nivel educativo.

Una vez determinados los problemas prototípicos y emergentes que, al finalizar un período lectivo, el estudiante deberá ser capaz de afrontar, el segundo paso consistirá en analizar qué recursos ponen en juego los profesionales experimentados para manejar convenientemente esos problemas y situarlos, en la medida de lo posible, al alcance de los estudiantes. Cuanto más afinados y pertinentes sean los problemas que se planteen, y más fidedignas las condiciones para su solución, más realistas serán los procesos cognitivos y emocionales implicados y, por lo tanto, más auténtica la evaluación del alumnado. En este sentido Wiggins (1989; 1990) y posteriormente Gulikers, Bastiaens y Kirschner (2004), además de la autenticidad de los problemas a los que se enfrenten los alumnos, se refieren también a la autenticidad del contexto físico (fidelidad con el lugar en que trabajan los profesionales, con condiciones de tiempo, orden, limpieza, recursos, etc. similares a las que un día el alumno encontrará), del contexto social (acceso a los recursos sociales habituales en el contexto de referencia: interacción con otros y posibilidad de pedir ayuda cuando se necesite), del resultado (si la calidad del producto que se obtenga se podrá medir con los mismos parámetros con los que se miden esos mismos productos en la vida profesional) y de los criterios de evaluación (si se basan en los utilizados en las situaciones homónimas del ejercicio profesional).

1.2. Relevancia

Apela al nivel de utilidad del conjunto de competencias que tratan de enseñarse y evaluarse, nuevamente en relación a los contextos profesionales de referencia.

Una práctica de evaluación puede ser relevante para el estudiante por distintos motivos: porque resulta propedéutica para preparar aprendizajes posteriores, porque es directamente funcional para apoyar al alumno en sus estudios y trabajos educativos, porque plantea una situación verosímil, es decir que podría ocurrir realmente; porque se refiere a tareas no propiamente académicas o, finalmente, porque es una legítima muestra de las prácticas de la comunidad profesional de destino. Como veremos en el próximo apartado la transición hacia actividades plenamente auténticas probablemente deberá graduar estas distintas opciones.

1.3. Proximidad ecológica

Se refiere a la conexión de la modalidad de evaluación elegida con las prácticas de evaluación habituales del profesorado que las pone en juego, dentro de su zona de desarrollo docente. Se trata, por una parte, de que la innovación sea asumible, en tanto se percibe cercana, pero también de que suponga un enriquecimiento de la práctica cotidiana. No es lo mismo que esa tarea de evaluación parta o tenga su origen en actividades anteriores realizadas en el propio centro, que se trate de una tarea importada de otro contexto e instaurada allí. En el primer caso la actividad, incluyendo las modificaciones que sufra, está legitimada o acreditada por una práctica educativa anterior que la convierte en un producto de esa cultura; dicho de otro modo, alguien avala la autenticidad de una determinada tarea al considerarla como propia, perteneciente a una experiencia previa y singular.

Siguiendo esta línea argumental, distintas investigaciones demuestran que cuando el profesorado considera una práctica como propia, y por lo tanto no extraña o ajena, las posibilidades de que logre implicarse en su desarrollo y aplicación aumentan de manera considerable. Apropiándonos del concepto introducido en su momento por Vigotsky (1977), el cambio, la innova-

ción propuesta, debe estar en la zona de desarrollo del docente para que tenga posibilidades de prosperar.

1.4. Identidad

El concepto de identidad se utiliza aquí en relación al grado de socialización profesional que favorece la práctica en cuestión.

La literatura sobre comunidades de práctica ha insistido en la importancia del proceso de enculturización identitaria (*identity enculturation*) que debe seguir un miembro legítimo de toda comunidad (Lemke, 2002). La identidad debe considerarse como un constructo que evoluciona dentro de un proceso dialéctico de búsqueda de significados entre personas. Co-construimos nuestras identidades en relación a otros, en contextos sociales en los que nos comprometemos a participar (Bauman, 1996; Barresi, 1999; Guile & Young, 2003), y estos *selves* se apropian de concepciones, discursos, estrategias y emociones que forman paquetes integrados listos para activarse ante problemas –prototípicos y emergentes–, propios de la comunidad, en nuestro caso la profesional (Monereo, 2007).

Facilitar desde los centros educativos (secundaria y universidad) una progresiva integración en el mundo profesional no supone pues únicamente favorecer el proceso de adquisición de conocimientos, implica además ofrecer un contexto de enculturización donde apropiarse de una identidad (Brown & Duguid, 2000; Gee, 2000; Wenger, 1998). Una evaluación auténtica debería ser especialmente sensible a este componente y tratar, al menos en alguna ocasión, de evaluar-en-contexto al estudiante.

En este sentido, y a modo de resumen, una tarea sería más auténtica cuando:

a) Su nivel de realismo sea elevado, es decir cuando las condiciones para su aplicación y las exigencias cognitivas implicadas, sean idénticas o muy similares a las condiciones y exigencias de esa misma tarea ejecutada en el entorno cotidiano o profesional en el que habitualmente se produce.

b) Su relevancia sea elevada, es decir los aprendizajes que conlleva sean directamente útiles para que el alumno enfrente situaciones o solucione problemas que son o serán habituales o prototípicos en su vida personal y/o profesional.

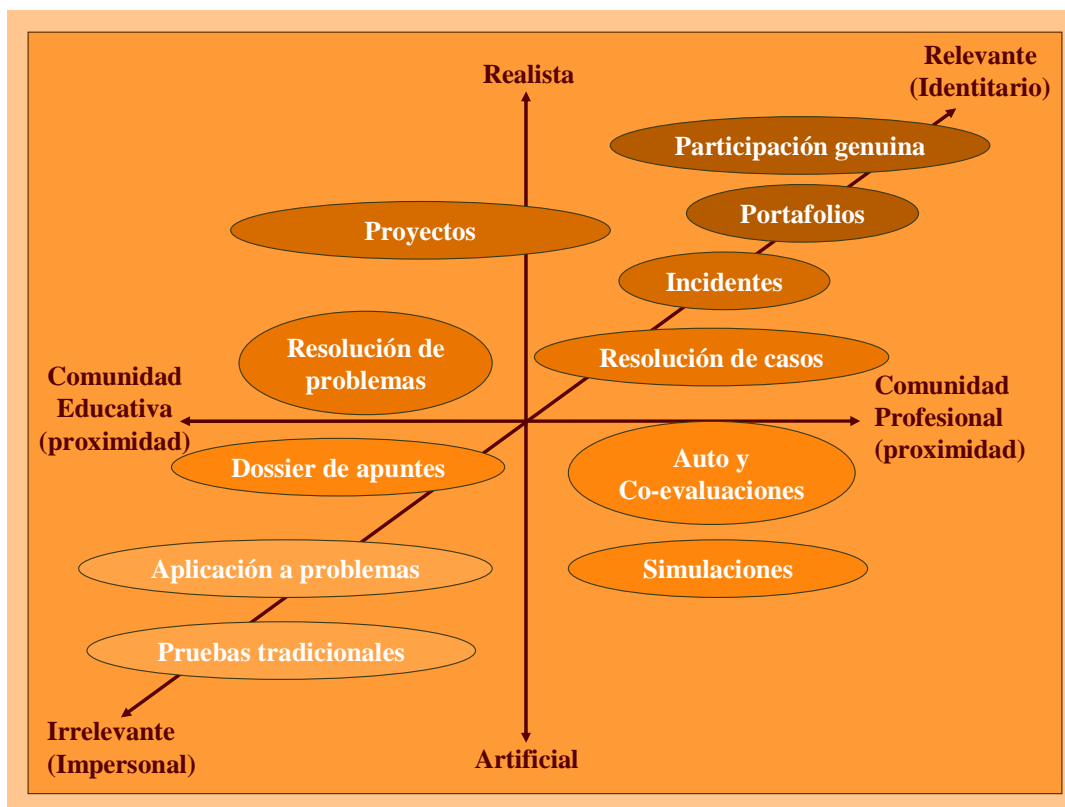
c) Su proximidad con las tareas y actividades propias del contexto educativo en el que se inscribirán y aplicarán, sea también elevada. Dicho de otro modo, cuando la tarea no se aleje de los planteamientos que habitualmente tienen los docentes de ese centro cuando enseñan o evalúan.

d) Su implicación identitaria sea asimismo alta. Se evalúe al alumno en situaciones que supongan distintos grados de inmersión profesional.

Lógicamente estas dimensiones pueden variar en su nivel de cumplimiento, situando cualquier tarea de evaluación en un continuo de mayor o menor autenticidad. Por otra parte las dimensiones que acabamos de presentar podrían emplearse también para estimar la autenticidad de una actividad de enseñanza.

2. Métodos de evaluación auténtica

Tomando como ejes organizadores las dimensiones que hemos identificado y analizado en el apartado anterior, podríamos clasificar algunos instrumentos de evaluación según su grado de autenticidad. En la figura 1 puede observarse esta clasificación que daría lugar a cuatro cuadrantes:



a) Métodos usuales en la comunidad educativa (proximidad ecológica), pero irrelevantes desde el punto de vista del alumno y artificiales, por cuanto su única referencia es intra-académica.

Formarían parte de este grupo las pruebas tradicionales de lápiz y papel, tipo “ejercicio”, es decir aquellas que pueden resolverse aplicando una rutina o algoritmo conocido, como por ejemplo la aplicación de una fórmula en matemáticas o de una regla ortográfica en lengua. También, aunque con un mayor grado de realismo, la aplicación de conocimientos a situaciones de la vida cotidiana, aunque poco verosímiles y bastante descontextualizadas (p.e. los típicos problemas de matemáticas en los que dos trenes se cruzan o un agricultor vende su mercancía). Un nivel más alto de autenticidad en este cuadrante lo tendrían propuestas como la elaboración de un dossier de apuntes de clase en la que el estudiante debe ampliar la información expuesta en el aula por su profesor a través de experiencias personales y ejemplos de la vida extra-académica.

b) Métodos próximos a los empleados en la comunidad profesional, relevantes para el alumnado puesto que evalúan aspectos funcionales de su futuro desempeño y suponen cierta enculturización (aunque tímida), pero que mantienen cierto nivel de artificiosidad al estar insuficientemente contextualizados y no enfrentarse a la complejidad e incertidumbre de las condiciones de situaciones profesionales reales.

En esta esfera están los métodos de evaluación basados en las emulaciones y simulaciones en las que el aprendiz puede manipular algunas variables controladas de un determinado sistema y demostrar sus conocimientos y pericia a través de los resultados que obtiene (p.e. en programas expertos por ordenador), o puede asumir determinados roles en una situación problemática y manifestar así su competencia en la gestión de dicha situación (p.e. un alumno que actúa como psicólogo en una entrevista clínica). Otros conjunto de métodos que se situarían en esta dirección son los basados en sistemas de auto-evaluación (el alumno evalúa sus propios avances) y co-evaluación (el alumno evalúa a un compañero y es evaluado a su vez por otro compañero). En numerosos medios profesionales la autorregulación de las propias acciones, y la evaluación por parte de los compañeros, están a la orden del día y su dominio es obligado. En este tipo de actividades resulta fundamental que el profesor logre compartir con los alumnos los criterios de evaluación de una determinada competencia con el fin de que éstos los interioricen y empleen autónomamente. En uno y otro caso el profesor puede introducir elementos que minimicen el sesgo en el que pueden caer los alumnos al evaluarse y evaluar a sus compañeros; por ejemplo, manteniendo premiando a aquellos alumnos que se desvían muy poco de las estimaciones que haría el profesor (o un experto profesional), y penalizando a los que presentan una desviación inaceptable.

c) Métodos con una buena dosis de realismo y relevancia, aunque tradicionalmente vinculados a la comunidad educativa.

Dentro de este cuadrante encontraríamos la resolución de problemas auténticos que pueden conllevar la presentación de situaciones extraídas de la vida profesional, muy bien documentadas, aunque finalmente la actividad se produce en el aula, por lo general en base a materiales escritos y que requieren una solución bastante convergente y cerrada, aquella que en su momento solucionó efectivamente el problema en cuestión.

Otro grupo de métodos, en este caso mucho más realista y relevante (dependiendo de la manera en que formule), lo constituyen los proyectos, que pueden tener un cariz educativo (p.e. buscar información sobre un tema y presentarlo en clase) o, en el otro extremo, plenamente profesional (p.e. crear un producto, buscar financiación, publicarlo, etc.). Evaluar a los alumnos mediante el desarrollo de un proyecto constituye uno de los planteamientos didácticos que mejor puede apoyar la transmisión desde una identidad como alumno de una carrera, a otra como incipiente profesional (ver en esta misma obra el capítulo 2).

d) En el último cuadrante se situarían los métodos de evaluación que atesoran un mayor grado de autenticidad. Se trata de prácticas de evaluación sumamente realistas, relevantes, próximas a las tareas profesionales y que favorecen la socialización en la comunidad profesional.

En el rango más alejado estaría la resolución de casos profesionales, comúnmente complejos, mal definidos y susceptibles de recibir soluciones dispares, como ocurre en la realidad; aquí los alumnos, teniendo todos los recursos a su alcance – de manera parecida a los que posee un profesional –, debe poner a prueba un gran número de competencias y estrategias. Más realista y

relevante resulta el método de incidente, conflicto preparado por el evaluador -y por lo tanto fingido-, en el que se inmiscuye al alumno, que cree que el conflicto se está efectivamente produciendo y debe actuar en consecuencia, como lo haría un profesional en circunstancias parecidas (p.e. se hace creer al estudiante de arquitectura que en un proyecto se ha producido un error en algún cálculo y éste debe plantear alguna solución de urgencia).

Tanto el método de evaluación a través de portafolios como la participación en prácticas profesionales genuinas, constituyen máximos exponentes de sistemas de evaluación auténticos. En el primer caso el estudiante debe recopilar muestras o evidencias de sus logros, en relación a unos objetivos competenciales previamente pactados con su profesor. Por ejemplo, para demostrar que es capaz de dirigir una reunión de trabajo, graba en vídeo una reunión real, realizada en su período de prácticas, y la expone a la evaluación de su profesor, quien puede analizarla con él y darle indicaciones para que repita la experiencia tomando en consideración algunos aspectos que deberá mejorar.

En cuanto a la participación genuina, estaríamos hablando de períodos de prácticas profesionales prolongadas (p.e. las que realizan los médicos residentes) en las que el estudiante participa directa y personalmente en la gestión de casos reales, bajo la supervisión de un tutor o mentor más experimentado. Hung & Der-Thanq (2007) sitúan también este tipo de actividades en la cúspide de las prácticas de carácter auténtico y subrayan como su característica más significativa la “co-evolución”. Con este término destacan la naturaleza recíproca de este tipo de actividades: el alumno recibe la clara influencia de sus tutores profesionales y compañeros de trabajo, pero a su vez influye sobre ellos al aportar nuevas concepciones, puntos de vista, ideas, etc. Un ejemplo frecuente lo encontramos en prácticas en las que el alumno aprende mucho sobre tratamientos convencionales de determinados problemas, y el profesional veterano también adquiere conocimientos sobre nuevas teorías, investigaciones o procedimientos (p.e. la utilización de Internet), proporcionados por el estudiante en prácticas.

En este caso la evaluación no solo se centraría en constatar la competencia del alumno para aplicar lo aprendido, y reconocer el conocimiento profesional que va construyendo, sino que también debería registrar el impacto de la participación del alumno en ese contexto y la influencia, positiva o negativa, que ha ejercido sobre el mismo.

La evaluación auténtica sin embargo, como cualquier opción educativa, tiene algunos requisitos y limitaciones que vamos a señalar a continuación.

3. Requisitos y limitaciones para instaurar una evaluación auténtica

Una primera condición *sine qua non* es el nivel de familiaridad que pueda tener el alumno con este tipo de tareas de evaluación. Para que una evaluación auténtica resulte efectiva es imprescindible que los alumnos hayan podido practicar con alguna de sus modalidades, antes de enfrentarse a una situación de evaluación, ya tenga una finalidad formativa (supervisar y mejorar el proceso) o sumativa (tomar decisiones de acreditación o promoción). Así lo confirman estudios como los realizados por Schnitzer (1993). Con el fin de garantizar esa familiaridad, las tareas de enseñanza-aprendizaje deberían parecerse a las tareas de evaluación, lógicamente con los ajustes y limitaciones temporales y materiales que sean pertinentes. Si estas tareas, además, cumplen con el requisito de proximidad a la cultura de prácticas del centro, el efecto novedad disminuirá notablemente.

Otro problema lo plantea el nivel de conocimientos que deben poseer los alumnos para enfrentar tareas complejas en situaciones pseudo-profesionales. Probablemente algunos alumnos

puedan sentirse superados por la situación planteada, lo que podría tener un impacto negativo sobre su aprendizaje (Sweller, Van Merriënboer y Paas, 1998). En este caso se recomienda simplificar las tareas realizando una cierta abstracción de la práctica profesional real; según Van Merriënboer (1997) este tipo de tareas aún pueden conservar el grado de autenticidad suficiente para que los alumnos puedan demostrar globalmente que están en posesión de las competencias necesarias. En una línea parecida estaría la propuesta de Hung & Der-Thang (2007) quienes proponen un conjunto de métodos de enseñanza y evaluación que pueden favorecer la transición entre la comunidad educativa y la profesional, tal como se exponen en la Figura 2.

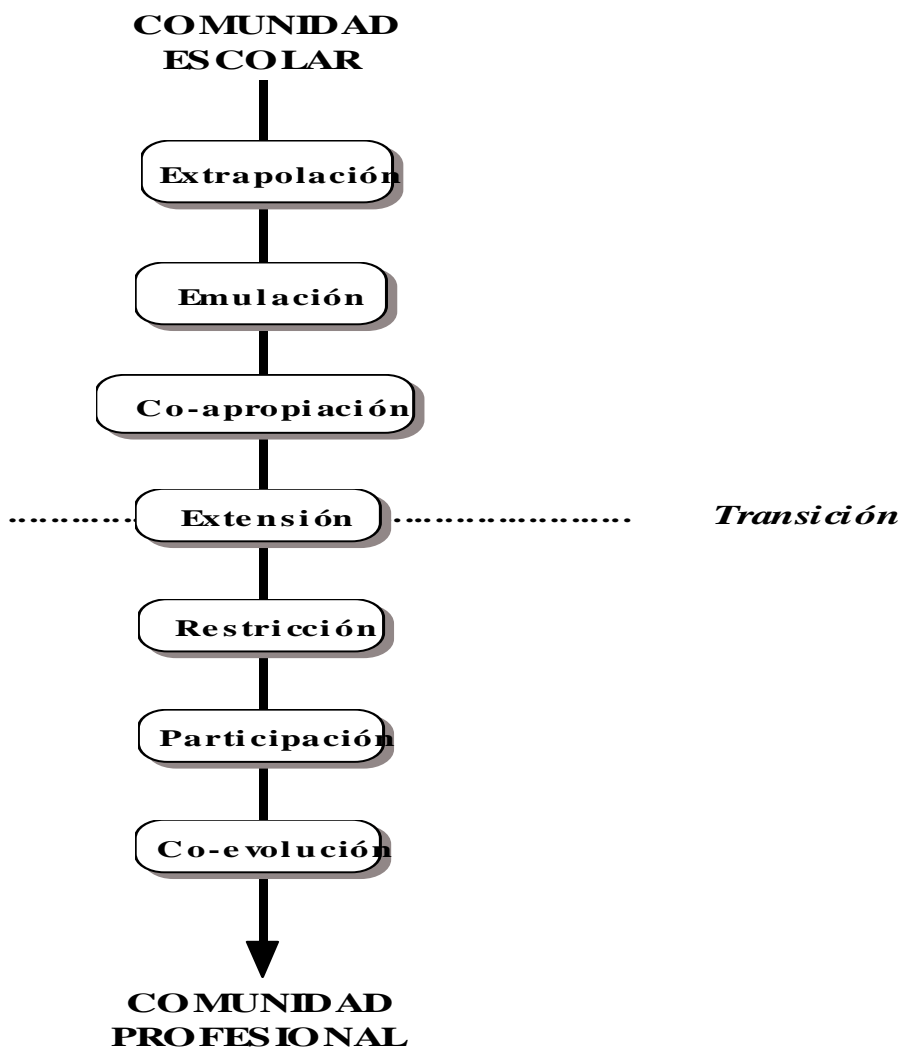


Figura 2. Métodos que promueven la transición entre comunidades

Brevemente estos métodos consistirían en:

a) Extrapolación. Potenciación de la aplicación de conceptos abstractos a problemas del “mundo real”.

b) Emulación o simulación: se reproducen escenarios y prácticas del mundo extra-escolar en base a las cuales el alumno puede experimentar o ser evaluado.

c) Co-apropiación: Se parte de actividades de enseñanza o evaluación que exigen la cooperación en parejas (*Peer Learning*) o en grupos pequeños de estudiantes.

d) Extensión: Se introducen en clase elementos propios de la cultura profesional para favorecer un acercamiento del estudiante a esa cultura. Estos elementos pueden ser: herramientas profesionales; experiencias profesionales a través de casos, historias de vida, iniciativas profesionales; la participación directa de profesionales en clase que resuelven problemas de su dominio, presentando a los alumnos modelos expertos de pensamiento; la introducción de nuevos conceptos y términos que suponen formas en la que los profesionales describen e interpretan sus asuntos.

e) Restricción: tareas o actividades profesionales reales, pero de menor complejidad, con el fin de minimizar los posibles errores.

f) Participación: Supone situarse en el contexto profesional de manera activa y comprometida, realizando tareas plenamente profesionales, generalmente bajo la supervisión de un tutor. Puede tomar modalidades diversas: Practicum clínico (con análisis de casos reales), cursos “sándwich” (con una preparación previa al inicio, después la práctica y un análisis teórico después de la misma), experiencias de campo, intercambios de puestos de trabajo (*clerkships*), grupos de trabajo colaborativos y cooperativos, etc.

g) Co-evolución: Implica la realización de un trabajo plenamente profesional, junto con otros profesionales, compartiendo los distintos procesos de intervención de igual a igual, y desarrollando un papel activo en el que no solo se recibe soporte y formación, sino que también se presta ayuda y se efectúan aportaciones al equipo, estableciendo una dinámica de influencia bidireccional y recíproca que permitiría la evolución del equipo de trabajo y la plena socialización de todos sus miembros.

Continuando con los problemas derivados de una evaluación auténtica, a menudo también se plantean dudas sobre si esta forma de evaluación cubre todos los objetivos y contenidos que deben evaluarse en una determinada disciplina. En realidad, puede que no siempre sea posible la evaluación mediante pruebas y tareas auténticas. En ocasiones, y en función precisamente de los objetivos, los contenidos contemplados y las características de los aprendices, otras formas tradicionales o innovadoras de evaluación pueden ser perfectamente adecuadas.

Un problema de fondo lo plantea lo que ha venido en denominarse la “subjetividad de la autenticidad”. Por ejemplo, hemos insistido en que una tarea auténtica debe ser relevante y realista, pero relevante y realista para quien, ¿para el profesor o para el alumno? Autores como Petraglia (1998), Nicaise et al. (2000), Huang (2002) o Stein, Isaacs, y Andrews (2004) han mostrado en sus estudios la distancia que puede existir entre lo que es auténtico para los docentes y para los alumnos, y han enfatizado la necesidad de compartir con éstos últimos el significado y sentido de lo que es educativamente auténtico. Se trataría de pasar de una situación de “pre-autenticación”, en la denominación que emplean estos autores, a una verdadera autenticación con el concurso de los usuarios finales, los alumnos. En este proceso de consenso con los alumnos, variables como el nivel educativo, la edad, los intereses personales o el conocimiento práctico que tengan éstos del mundo laboral y profesional, pueden tener una fuerte incidencia y

deberán anticiparse (Honebein et al., 1993). Es posible, por ejemplo, que en los últimos cursos universitarios, cuando los alumnos tienen más conocimientos y contacto con el mundo profesional, deba aumentarse la autenticidad del contexto físico de la tarea.

Un último problema radica en la posibilidad de realizar una correcta valoración de la eficacia de la evaluación auténtica. La respuesta nos remite a un concepto clave en educación, pero ciertamente complejo: el grado de transferencia de las competencias adquiridas y evaluadas. En una evaluación auténtica, por el mismo hecho de situar al alumno en una situación de resolución muy afín a la real, podríamos pensar que la transferencia es automática, sin embargo no basta con esa percepción, es necesario comprobar que efectivamente se produce. Por consiguiente, una evidencia fehaciente de que la evaluación ha sido eficaz consistiría en demostrar que el alumno, en una situación absolutamente real, sin aderezos ni aditivos, tiene el comportamiento y emplea conocimientos y competencias en la línea en que la evaluación había predicho.

Tabla 1. Relaciones conceptuales entre Evaluación dinámica, formativa, formadora y auténtica.

	Qué evalúa	Cómo evalúa	Cuando evalúa	Quién evalúa	Para qué evalúa	A qué se opone
Evaluación Dinámica	Potencial de aprendizaje (Andamiajes)	Tareas representativas (abstractas)	Durante el desarrollo de la tarea (puntual)	Especialista (test)	Identificar las ayudas individuales que el alumno necesita	Evaluación estática (qué puede hacer el alumno sin ayuda)
Evaluación Formativa	Proceso de enseñanza-aprendizaje (construcción)	Registro secuencial de los cambios	Durante el desarrollo de las actividades (extensiva)	Profesor o tutor	Mejorar los procesos de enseñanza y aprendizaje (nivel aula)	Evaluación tradicional (del rendimiento o producto)
Evaluación Formadora	Proceso de aprendizaje (construcción)	Registro de los materiales que produce el alumno	Durante el desarrollo de las actividades (extensiva)	Los propios alumnos (auto / co-evaluación)	Hacer a los alumnos conscientes de sus progresos	Evaluación centrada en el profesor
Evaluación Auténtica	Competencias funcionales (resolución de problemas)	Problemas representativos de situaciones reales socio-personales	Durante el desarrollo de la resolución del problema (puntual o extenso)	Profesor, tutor o especialista	Tomar decisiones educativas (institucional, nacional o internacional)	Evaluación académica, simplificadora y cerrada