

Universidad de Talca

Instituto de Investigación y Desarrollo Educacional – Proyecto Mecesup Tal 0101

Evaluación de Competencias en la Educación Superior

Prof. Gustavo Hawes B.

Prof. Oscar Corvalán V.

Talca, 2005

Tabla de contenidos

1. Presentación.....	7
1.1. Visión general.....	7
1.2. Plan de trabajo.....	9
2. Definiciones asociadas.....	9
2.1. Dominios de competencia.....	9
2.2. Competencias.....	10
2.3. Capacidades.....	10
2.4. Tareas.....	11
3. Evaluación.....	11
3.1. Concepto habitual de la evaluación.....	11
3.2. Una concepción renovada.....	14
4. Demandas de cambio.....	20
4.1. Demandas de cambio sobre los docentes.....	20
4.1.1. Demandas sobre la didáctica universitaria.....	20
4.1.2. Demandas sobre la evaluación.....	21
4.1.3. Demandas sobre la manera de organizarse.....	21
4.2. Demandas de cambio sobre los estudiantes.....	21
4.2.1. Requerimientos de renovación metodológica.....	21
4.2.2. Demandas sobre los hábitos y estilos de estudio.....	22
4.3. Demandas sobre la organización de los estudios.....	22
4.4. Demandas sobre la administración académica.....	23
5. Criterios y Estrategias evaluativos.....	25
5.1. Criterios a tener en consideración.....	25
5.2. Estrategias apropiadas.....	26
5.3. Estrategias inapropiadas.....	26
6. Medición de competencias.....	27
6.1. El espacio de la construcción de competencias.....	27

6.2. Competencias y medición.	29
6.3. Portafolios	32
6.4. Listas de chequeo.	34
6.5. Escalas de apreciación.	35
7. Calificaciones.....	36
7.1. Conceptos generales.....	36
7.2. Representación de los espacios evolutivos de competencias	37
7.3. El modelo de QAA para escalar las apreciaciones.	39
7.4. Una propuesta desde la perspectiva de la gestión de recursos humanos	40
7.5. Transformación a nota.....	43
8. Referencias.	46
9. Anexo. Lista de cotejo para establecer la calidad de la evaluación de los estudiantes (McInnis, 2004)	48

TABLAS

Tabla 1 - Recursos evaluativos asociados a ámbitos C, P, IP	30
Tabla 2 - Ejemplo de niveles de competencia según MA Alles	42

ILUSTRACIONES

Ilustración 1 - Ciclo de la gestión del conocimiento.....	8
Ilustración 3 - Estructura de logros en el model del QAA	40

1. Presentación.

1. Visión general.

Los temas relacionados con la evaluación están entre los que más problemas ponen a quienes ejercen la docencia, en cualquiera de sus niveles. Específicamente, la enseñanza de nivel universitario se hace más difícil y compleja en muchos casos debido a que los docentes no tienen formación especializada en el campo de la enseñanza y otros campos afines, y de manera particular en las cuestiones evaluativas. En efecto, un buen estudiante puede sobrevivir a una docencia de mala calidad, pero se le hará mucho más difícil sobrevivir a una evaluación de mala calidad. Por otra parte, para una gran mayoría de los estudiantes, la evaluación determina en buena parte lo que es el curriculum: su selección, su contenido, sus jerarquías y su enfoque de aprendizaje (McInnis, 2004).

Las reformas curriculares orientadas a la formación sobre la base de competencias acarrearán consigo la problemática de la evaluación; no podría esperarse menos de un cambio curricular tan profundo como el que se produce al pasar de un marco centrado en objetivos a uno centrado en competencias. No obstante, suele encontrarse resistencia, no sólo en los docentes, sino también en los estudiantes para quienes una evaluación sólo es tal cuando lleva asociada una nota.

Un dato fundamental de esta reforma tiene que ver con la enseñanza focalizada tanto sobre las propuestas locales del módulo o asignatura como, con igual importancia, focalizada sobre las competencias de salida del profesional que se forma. Por ello es que, según Le Boterf (2003:229), se trata de una gestión anticipada de las competencias. Existen miradas erróneas sobre el punto. Para Le Boterf que no se trata de predicciones, ni profecías, ni una bola de cristal. Más bien, la gestión anticipada de competencias tiene que ver con (a) hipótesis de evolución sobre las competencias necesarias; (b) un dispositivo de actualización para corregir regularmente estas hipótesis; (c) decisiones (formación, profesionalización, organización, movilidad interna...) para hacer probables las hipótesis consideradas positivas; (d) amarrar los cabos para que los individuos orienten el desarrollo de su profesionalismo. La complejidad que se deriva de este enfoque se encuentra reflejada en los puntos relativos a 'una concepción renovada' (Cfr. pág. 14 de este informe).

La pregunta básica que se plantea es ¿cuáles son las formas más efectivas para documentar si los estudiantes han logrado las competencias requeridas por el curso o programa de formación? Es decir, la evaluación responde fundamentalmente a la necesidad de cons-

truir, alimentar y utilizar un sistema de conocimiento institucional relativo a un estudiante, con el ulterior propósito de certificar y acreditar ante la sociedad la calidad profesional del graduado. Ciertamente un requisito para la existencia de un sistema de evaluación, cualquiera sea su naturaleza, requiere de un grado de credibilidad por parte de la comunidad frente a la cual se informan los resultados; esto no siempre es así como lo informan Wilcox y Ryder (2002).

De acuerdo a Bück (2003:164), para que el sistema de gestión de conocimientos sea pertinente y active las competencias, necesariamente debe concebirse bajo la forma de un proceso con etapas claramente definidas: (1) producción e identificación del conocimiento a compartir; (2) Determinación del valor de utilidad y su transferibilidad, (3) Captura y formalización del saber-hacer; (4) Fijación del nivel de seguridad; (5) Difusión y compartir del conocimiento formalizado; (6) Evaluación del valor de utilidad individual y colectivo; (7) Valorización, enriquecimiento y mantenimiento del conocimiento por parte de los usuarios. Gráficamente se trata de un ciclo como el que se reproduce a continuación.

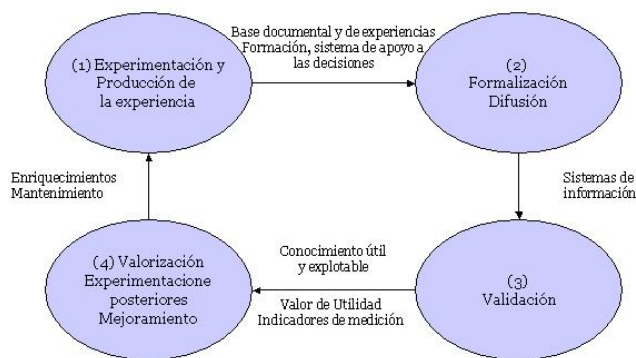


Figura 5.1. Ciclo de la gestión del conocimiento

Ilustración 1 - Ciclo de la gestión del conocimiento

Mirando las prácticas habituales en la educación superior, podemos concebir que se cumplen al menos las 4 primeras etapas señaladas por Bück; no obstante, las últimas, que tienen que ver más de cerca con la comunicación y compartir, se advierten débilmente estructuradas. En otras palabras, descuidamos y subutilizamos la información proveniente de los procesos evaluativos a los que sometemos a los estudiantes.

Como puede advertirse –y en esos términos se plantea este trabajo- el conocimiento producto de la evaluación tiene un trasfondo epistemológico que no puede dejarse de lado.

2. Plan de trabajo.

En las siguientes secciones se plantean algunos tópicos prácticos, sin pretender un desarrollo teórico importante, atendiendo al carácter eminentemente aplicado de este documento.

Primero, se retoman algunas definiciones cruciales en materia de competencia, a saber: dominios de competencias, competencias (por áreas), habilidades y tareas, así como la relación entre estos conceptos y procesos. A continuación se hace una primera digresión comparando el concepto tradicional con uno renovado de evaluación; se enfatizan las demandas que se plantean especialmente en referencia a la puesta en práctica de las nuevas ideas. La cuarta sección explora, desde la perspectiva de la educación basada en competencias (EBC), cuáles serán las estrategias evaluativas más apropiadas así como las menos apropiadas para enfrentar este modelo general curricular y docente.

Las dos últimas secciones exploran cuestiones más específicas, instrumentos de evaluación (en sus variadas expresiones) y la conversión de los puntajes a nota, enfatizando las perversiones asociadas a la nota como representativa de un haz de conductas.

2. Definiciones asociadas.

En esta sección se retoman algunos conceptos ya presentados y discutidos anteriormente, relacionados directamente con las competencias.

3. Dominios de competencia

Un dominio de competencia se describe como el complejo de capacidades de ejecución –acciones y decisiones- de diversa naturaleza que se conjugan y ponen en acción en el profesional de manera tal que le habilitan para desempeñar un rol específico: al profesor, enseñar; al médico, diagnosticar y tratar; el ingeniero, calcular y diseñar; el administrador, gerenciar.

El Dominio de Competencias se describe en términos de acciones y decisiones referidas a un ciclo completo de acción.

4. Competencias

Un artículo de Proffer (1994) declara, aludiendo a la educación en los Estados Unidos, que se establecen nuevos estándares académicos sobre la creencia de que los estudiantes tendrán mejores rendimientos si se los desafía a ello. Frente a esto, no hay mejor defensa que tener pocas expectativas: jamás se verá uno frustrado. No sucede lo mismo con los estudiantes europeos o asiáticos.

El informe “*Tuning Educational Opportunities in Europe*”, propone al respecto un concepto en el que “...las competencias representan una combinación de atributos (con respecto al conocimiento y sus aplicaciones, aptitudes, destrezas y responsabilidades) que describen el nivel o grado de suficiencia con que una persona es capaz de desempeñarlos” (González & Wagenaar, 2003).

Las competencias poseen tres dimensiones fundamentales: cognitiva, procedimental e inter-personal o afectiva). En los hechos ninguna competencia se ejerce de manera aislada (sólo cognitiva, sólo procedimental o sólo interpersonal): siempre una competencia da cuenta de las tres dimensiones.

Lo que distingue a una competencia profesional es que suele implicar una toma de decisiones fundada: “diagnosticar el estado de salud bucal de un paciente”, “producir un plan de trabajo”, “operar una máquina de control numérico”, “conducir una reunión de trabajo con profesionales de otras disciplinas”.

Una competencia, por su parte, requiere de ciertas capacidades o subcompetencias.

5. Capacidades

Una capacidad es un comportamiento específico y complejo que forma parte de una competencia, definido en términos de acciones sobre contenidos con un nivel de desempeño definido.

Se expresa como “verbo-sustantivo-adverbio”: el verbo expresa la acción; el sustantivo, el objeto sobre el que se ejecuta dicha acción; el adverbio, la calidad o amplitud del desempeño esperado. Por ejemplo, “Conocer – los fundamentos de la profesión – en términos generales”; “Calcular – la estructura de un puente – sin errores”.

6. Tareas

Cada una de las ejecuciones que compone una capacidad. Suelen ejecutarse en secuencias relativamente predefinidas; cuando no son procesos algorítmicos, el sujeto debe establecer la forma y secuencia en que realizará las diferentes tareas.

Tarea se denomina a cada una de las ejecuciones que compone una capacidad. Lo más propio de una tarea es que en sí misma carece de sentido, a menos que se la entienda en el contexto de una capacidad. De esta forma “Ajustar una sierra” o “Medir la talla”, aunque sean en sí mismas acciones complejas, carecen de sentido a menos que se especifique si “ajustar una sierra” es en el contexto del procesamiento de maderas o de una operación quirúrgica, por ejemplo.

3. Evaluación

Examinamos en esta sección tres dimensiones: las características del modelo habitual o estándar con que se evalúa en la universidad, una concepción renovada del mismo a la luz de los conceptos centrales de competencia (ver sección anterior) así como las demandas de cambio que se imponen desde esta visión renovada.

7. Concepto habitual de la evaluación.

El modelo habitual o estándar por el que se enseña a los estudiantes de la educación superior no se diferencia mucho de la manera en que se hacía hace 30 o 50 años atrás. Este inmovilismo se explica en parte por el efecto de la reproducción de los modelos que han sido exitosos con uno mismo, los que luego se convierten en el criterio para actuar de los profesionales que enseñan sin tener una formación apropiada y actualizada al respecto.

En el campo de la evaluación la situación no es muy diferente. Una consideración de las características del modelo vigente pone en evidencia no sólo sus problemas o defectos, sino que también permite tener conciencia de sus fortalezas.

Entre los aspectos más característicos del modelo estándar de evaluación podemos señalar los siguientes:

1. Normalmente se trata de una evaluación que privilegia el dominio de los contenidos del programa de estudios; eventualmente, focalizada sobre objetivos de logro. La principal referencia de la evaluación es “la materia”, en particular aquellos tópi-

cos tratados en clase por el docente. No se consideran las destrezas o competencias superiores desarrolladas en el proceso de aprendizaje. Un estudio de Algozzine, Bateman et al. (1999) de nivel nacional en los Estados Unidos sobre la incorporación de nuevos métodos basados en las TIC, pone en evidencia la importancia y positivo efecto de la incorporación de actividades de nivel de complejidad superior a la mera rememoración o reconocimiento de unidades de información.

2. Una evaluación centrada fundamentalmente sobre desempeños cognitivos, en particular recuperación de información y, en lo procedimental, la aplicación de algoritmos y reglas. Se enfatizan principalmente conductas cognitivas (conocimiento simple y destrezas intelectuales según la taxonomía de B. Bloom); los procedimientos reciben atención sola o casi solamente en las asignaturas que comprenden secciones de taller, laboratorio o “pasos prácticos”. Ciertamente, un curriculum cuyo propósito es que los estudiantes conozcan “reproductivamente” los contenidos de las materias de enseñanza, no puede sino tener un formato evaluativo que insista en la reproducción de los contenidos como criterio fundamental de evaluación. Abbott y Ryan (1999) sostienen al respecto que la educación centrada en metas u objetivos curriculares no siempre estimula aprendices creativos y emprendedores; señalan además que la nueva investigación del cerebro apoya la noción de que el aprendizaje constructivista entra en conflicto con muchas disposiciones institucionales de aprendizaje.
3. Una evaluación que deja completamente fuera los requerimientos afectivos e interpersonales de la formación profesional. Esto se justifica en el sentido de que “sólo se puede evaluar aquello que se ha enseñado explícitamente”; cabe preguntarse por lo sensato y correcto de esta expresión. Realmente, sostenemos, lo que debe evaluarse es la calidad de los logros del estudiante, en términos de las competencias esperadas, y de la ayuda y colaboración recibida de parte del docente.
4. Se evalúa en una dimensión exclusivamente individual –referida al sujeto, sin evaluación asociada de otros componentes de la formación. Se entiende entonces que el aprendizaje es un logro y una responsabilidad propio del estudiante, sin incluir los componentes que vienen de la enseñanza, el curriculum, la disponibilidad de recursos, los medios de aprendizaje, los ambientes y contextos. Evaluar, entonces, es una acción más amplia, abarcando muchas otras dimensiones.

5. Se privilegian las respuestas escritas, y en particular aquellas que implican marcar más que producir una respuesta. Con esta modalidad se enfatiza una concepción de la inteligencia como fundamentalmente convergente (sólo hay una respuesta correcta) y 'cristalizada', es decir, no dinámica ni evolutiva.
6. Las respuestas se convierten en cifras que son manipuladas aritméticamente. No habiendo definiciones explícitas y consensuadas respecto de qué significan los números utilizados al calificar los rendimientos de los estudiantes. En una situación de completa anomia respecto del contenido de las calificaciones numéricas, cualquier análisis aritmético es prácticamente un sinsentido. Sin embargo, se opera sobre estas débiles bases y se toman importantes decisiones sobre las vidas de los estudiantes.
7. Se asume que los promedios representan algo:
 - en parte, porque se considera que los diversos tipos de logro sobre diversos tipos de material son completamente comparables;
 - en parte, porque se considera que los logros obtenidos al comienzo del proceso son completamente comparables con los obtenidos al término del proceso;
 - igualmente, se considera que el desempeño futuro del sujeto será el promedio del período más que el logro obtenido al final del mismo.

Este modelo evaluativo es posibilitado por una serie de convicciones y creencias, a la vez que está validado por años de práctica con resultados que, al final de cuentas, no han sido tan negativos: en otras palabras, un modelo evaluativo que ha sido funcional a un modelo de formación que se basa en un conjunto de convicciones acerca de la cultura y el curriculum, la naturaleza de las profesiones y de las instituciones formadoras, de la enseñanza y del aprendizaje.

Por cierto, si modificamos los conceptos centrales de la formación, parece claro que tendremos que preguntarnos por los efectos que ello traerá sobre el modelo evaluativo a utilizar en este nuevo marco.

8. Una concepción renovada

Si el modelo formativo es uno asociado a competencias, entonces se requiere pensar en evaluación basada en competencias. Esto no es únicamente cuestión de técnicas sino que se entiende en un cambio cultural de la concepción misma de la evaluación y su sentido.

A continuación se proponen algunos principios y se examinan sus efectos sobre el modelo evaluativo en la universidad.

1. Concebir los logros de los estudiantes.

Por una parte, esto implica afirmar que los estudiantes trabajan más efectivamente cuando saben para qué están trabajando. En otras palabras, el aprendizaje es más probable como función del grado de significación de la propuesta curricular. Igualmente, los estudiantes van a valorar las tareas evaluativas cuando perciban que son “reales”, es decir, propuestas evaluativas que representen desafíos serios y auténticos a su inteligencia y capacidad. Los estudiantes pueden reaccionar positivamente no sólo ante una calificación en juego, sino también ante una tarea de naturaleza tal que involucre conocimientos y destrezas (McInnis, 2004).

Bajo el modelo de competencias, se entiende que los logros de los estudiantes son estructuras complejas antes que simples. “La evaluación de competencia debería incluir un insumo multivariado/multidimensional y multimétodo, incluyendo perspectivas múltiples en todos los niveles de desarrollo profesional” (Christiansen et al., 2004).

Entre los efectos principales, se encuentra desarrollar metodología que de cuenta de una variedad de insumos y de una diversidad de resultados multidimensionales. Esto supone una previa descripción de las competencias de salida que se esperan del estudiante.

Si bien lo que interesa en último término es un resultado cualitativamente sostenible, es preciso tener en cuenta que no siempre el resultado es completamente uniforme u homogéneo en su descripción, razón por la cual diferentes estudiantes podrían obtener la misma competencias utilizando diferentes recursos y desempeñándola de acuerdo a una apropiación personal.

2. Toda evaluación de competencias supone una perspectiva evolutiva (Christiansen et al., 2004) y su principal naturaleza y propósito es proveer retroalimentación oportuna (Chickering & Gamson, 1987; McInnis, 2004).

La retroinformación es crítica para el desarrollo del proceso formativo de los estudiantes, especialmente en las primeras etapas de un curso o de una materia. En otras palabras, se requiere considerar el proceso de desarrollo y evolución del estudiante en términos de los logros buscados. La perspectiva evolutiva requiere del docente un apropiado manejo del tiempo: ésta ya no es más la medida física estándar, igual para todos, sino que también se conjuga con los procesos personales de cada uno.

Entendiendo entonces la perspectiva temporal y que lo que realmente importa es el logro de la competencia propuesta, entonces el promedio como indicador global de logro pierde sentido en este referente. (Ver análisis más detallado en 27).

3. Toda competencia requiere la definición de un estándar.

“Se requiere definir el nivel mínimo de competencia para que el “umbral” de la competencia básica funcione profesionalmente, así como también los niveles aspiracionales de competencias avanzadas” (Christiansen et al., 2004). Hamilton, McCaffrey et al (2003) estudiaron la relación entre los rendimientos estudiantiles y las prácticas de enseñanza en programas de la National Science Foundation’s Systemic Initiatives, encontrando relaciones consistentes entre el uso de instrucción basada en estándares y el logro estudiantil. En el caso de los propios docentes, las propuestas y un proyecto desarrollado por Hirsh (Hirsch & Weber, 1999; E. D. Hirsh, 1999; S. Hirsh, 2004) logra documentar que los planes más efectivos de desarrollo profesional de los profesores son aquellos que se orientan a resultados, se basan en estándares y se focalizan sobre la experiencia laboral diaria del profesor.

En el caso de la formación profesional inicial, el criterio para determinar si hubo logro o no, está dado externamente más que por la evaluación del mejoramiento de la eficiencia o calidad previamente evaluada, y es funcional al perfil de salida, incluyendo componentes profesionales de diversos contextos de desempeño. En este sentido, los procesos evaluativos han de tomar en cuenta de manera relevante la utilización de los diferentes saberes desarrollados y aprendidos en el curso (programa, módulo) en situaciones de desempeño lo más próximamente reales que sea posible en el marco de la enseñanza. (Ver desarrollo más amplio en 21, página 39).

Ciertamente, la consistencia de los calificadores/evaluadores, con todo el impacto que tiene sobre la apropiada retroinformación al estudiante, puede apoyarse de

manera importante sobre estándares de logro definidos de antemano, consensuados y operacionalizados por un equipo docente (McInnis, 2004).

4. El promedio no siempre basta o es suficiente.

“La evaluación de competencias debería involucrar una combinación de modelos compensatorios y no-compensatorios”(Christiansen et al., 2004).

Entendiendo que una competencia es el complejo de acciones o capacidades puestas en acción, abarcando aspectos cognitivos, procedimentales, afectivos e interpersonales, para certificar la competencia se requiere que el sujeto haya dominado todos y cada uno de sus componentes. En otras palabras, el carácter no-compensatorio indica que no es posible proceder por simple promedio para determinar aritméticamente que un sujeto ha logrado la competencia. En el mundo universitario es el clásico tópico de las asignaturas que tienen componentes denominados “teóricos” y “prácticos” o “cátedra” y “laboratorio”. Para aprobar una asignatura de este tipo, el estudiante necesariamente debe aprobar cada uno de sus componentes: no basta que el promedio de las calificaciones arroje un valor que se ubica en o sobre la nota mínima. Esto, precisamente porque los componentes de un área no son compensatorios (no reemplazan o sustituyen con equivalencia) a los componentes de la otra.

5. La evaluación tiene diferentes funciones asociadas a diferentes decisiones.

“En la psicología profesional se requiere tanto la *evaluación formativa* (por ej., mientras todavía se está en entrenamiento y progresando en un programa o para mejoramiento continuo de la carrera) como *evaluación sumativa* (por ejemplo, para el licenciamiento)” (Christiansen et al., 2004).

Las decisiones evaluativas no pueden ser simplemente de aprobación-reprobación, sino que deben ser seguidas por acciones que tiendan a compensar o remediar, ampliar o reorientar, los resultados obtenidos por el estudiante.

La *evaluación formativa* se entiende en el sentido expresado en la sección anterior. En el caso de la *evaluación sumativa*, se trata de acreditar públicamente, ante la sociedad (caso de títulos grados) o ante la universidad (caso de asignaturas), que el graduado o estudiante pueden llevar a cabo las acciones y decisiones que describen las competencias evaluadas.

6. “La evaluación de las prácticas de competencias durante el entrenamiento debería *modelar* cómo los profesionales deberían continuar evaluando competencias a lo largo de la propia carrera” (Christiansen, Roberts et al, 2004).

Además, se asegura ante la sociedad y la profesión que el egresado está en condiciones y capacitado para decidir si es capaz o no de enfrentar un problema determinado, tomando los cursos de acción respectivos: iniciando las acciones o derivando otro profesional. La auto-evaluación pasa a ser un rasgo propio de la conducta profesional. Un estudio de Reiger (2001) informa de un grupo sustancial de graduados que evalúan a su centro educativo, sosteniendo al respecto que esto es representativo de la democracia en acción: cuando los propios graduados son consultados para evaluar el programa educativo que se les ha provisto.

7. “La evaluación de competencias requerirá un ‘viraje cultural’ en los formadores, los formados, y los profesionales.” (Christiansen, Roberts et al, 2004)

Un cambio en la forma de concebir la evaluación como parte de una enseñanza renovada, y el uso de técnicas más amplias y complejas para trabajar. El cambio no sólo afecta a los docentes sino que también a los estudiantes, a los egresados, a los miembros de la profesión (el gremio), al conjunto de la sociedad.

8. “La evaluación basada en competencia debería incluir indicadores mensurables de conductas, destrezas y actitudes/valores” (Christiansen, Roberts et al, 2004).

La demanda de la mensurabilidad tiene que ver fundamentalmente con la «observabilidad» de los comportamientos (que corresponden a los indicadores) relativamente complejos que dan cuenta de una competencia en particular.

Una observación para ser “medida” en términos genéricos cuando es posible construir categorías en las cuales clasificar sus diversas expresiones. Cuando estas categorías están ordenadas sobre alguna dimensión, entonces puede expresarse como variables numéricas, aunque su propia naturaleza no sea tal.

9. “La destreza se evalúa mejor mediante un muestreo de conductas (por ej., a través de evaluación directa/observación) a través de perspectivas múltiples, incluyendo ‘expertos’, pares, sí mismo, clientes, miembros de la comunidad. (Christiansen, Roberts et al, 2004)”

Un principio básico de realismo indica que no es posible observar todas las posibles conductas de un sujeto en el marco de las posibles demandas que se le hagan desde la profesión o desde otros componentes del currículum. Se hace necesario entonces proceder muestralmente. Algunas cuestiones que surgen son:

- no se puede considerar un muestreo aleatorio, a menos que el material (contexto de desempeño) sobre el que se desarrolla la competencia sea homogéneo y, por consiguiente, intercambiable; al menos, que sean estructuralmente comparables;
- el muestreo intencionado, aparte de su debilidad estadística, permite centrarse sobre aquellas conductas que son completamente relevantes para la competencia (asumiendo que sean observable por algún medio o dispositivo)
- los observadores tampoco pueden ser sorteados al azar sino que se los asume como expertos en el área de competencia que se evalúa, pudiendo por tanto fijar los estándares de desempeño más altos
- la variedad de observadores colabora a la triangulación de la competencia evaluada y, por consiguiente, mejora de manera sustancial la apreciación (=evaluación) que se haga del desempeño del estudiante

10. “La evaluación de destrezas debe hacerse en un enfoque integrativo y no como un conjunto de competencias evaluadas individualmente” (Christiansen, Roberts et al, 2004).

La evaluación tradicionalmente se plantea en términos analíticos, focalizando conductas y contenidos muy específicos, muchas veces sin relación con contextos de desempeño (aspectos relacionados con la práctica de la profesión en condiciones reales). Finalmente, se suman y promedian escalas que no significan lo mismo en términos numéricos y que, menos aún, significan la misma cosa en términos de las dimensiones que se están promediando.

La demanda de un enfoque integrativo no es compatible con el concepto estadístico de “promedio”. Para ello se requiere una modalidad en que una comunidad o equipo docente, bajo un modelo específico de gestión del conocimiento, proporcionen un juicio de expertos fundado que describa en su globalidad e integridad el nivel de

desarrollo alcanzado por un sujeto en términos del perfil de competencias que le corresponde, tal como se ve a continuación.

11. La evaluación debe colaborar a la gestión, alimentación y mantenimiento de un sistema de gestión de conocimiento, a la vez que hacer uso del mismo para la toma de decisiones (Bück, 2003).

El evaluador debe considerarse a sí mismo integrado en un proyecto mayor, en el que la gestión del conocimiento¹ es la parte sustantiva. Al final de cuentas, en términos evaluativos, la decisión sobre la promoción y acreditación profesional de un estudiante tiene que ver con un juicio institucional, respaldado por un juicio compartido de los miembros de un equipo formador (o comunidad formadora), fundado y consensuado. No se sólo cuestión de un promedio aritmético, sino que de un juicio de mayor calado, en el que se integran las diferentes dimensiones del estudiante.

La trascendencia de un enfoque como éste implica un giro radical en la modalidad analítico-cartesiana con que estamos evaluando en la actualidad.

12. Las prescripciones producidas bajo un modelo de competencias son de carácter abierto, orientadas a la profesión completa, con reducción de niveles jerárquicos y responsabilidad ampliada (Le Boterf, 2003). Esto implica entonces que se requieren “consideraciones multiculturales en todos los niveles de la evaluación de competencias. Hay una necesidad de asistir tanto a los aspectos culturales de la evaluación de competencias como a la evaluación de la competencia cultural” (Christiansen et al, 2004).

Por contraposición al modelo analítico-tayloriano, de prescripción estricta, con numerosos niveles jerárquicos y una lógica de puesto de trabajo, esta concepción exige disponer de estructuras decisionales y de descripciones más amplias e integradoras en que, más que la exigencia unidimensional (sólo técnica, sólo cognitiva, sólo procedimental) se consideren exigencias pluridimensionales (técnica, económica, de calidad, relacionales, medioambientales, etc).

¹ Se habla aquí de “gestión del conocimiento” en sentido lato: no sólo del conocimiento producto de los actos evaluativos, sino que del conocimiento curricular, el organizacional, el conocimiento legal, los saberes procedimentales de la administración, etc.

4. Demandas de cambio

El viraje hacia un modelo evaluativo basado en competencias pone importantes demandas sobre los diferentes actores y funciones asociados: docentes, estudiantes, la organización de los estudios, la administración académica.

Estas demandas de cambio no se pueden referir únicamente a modificaciones en ciertas prácticas más o menos irrelevantes o en normativas a las que se adhiere más bien formalmente. Por el contrario, la renovación tiene que ver con culturas y subculturas, con maneras de entender y practicar la enseñanza universitaria, de comprender y practicar la función docente en un marco diferente.

9. Demandas de cambio sobre los docentes

Sobre los docentes en general se plantean demandas de cambio en relación a la didáctica, la evaluación, la manera de organizarse.

10. Demandas sobre la didáctica universitaria

La renovación didáctica requiere pasar de la enseñanza basada en los contenidos a la enseñanza basada en competencias: de lo que es capaz de reproducir o aplicar el estudiante, a las acciones que revelan que domina una competencia determinada.

Se requiere aquí un proceso de cambio que mueva el foco de la insistencia desde lo que es capaz de reproducir o aplicar el estudiante, al complejo de acciones que revelan que domina una competencia determinada.

Además de una variación de la mirada, ciertamente se quiere una variación en las prácticas didácticas dentro y fuera del aula. Una docencia que se apoya fuertemente en la conferencia, en la capacidad de un expositor de mantener un discurso (monólogo) relativamente animado con dispositivos audiovisuales, tenderá a una evaluación reproductiva de los conceptos enseñados o aplicación de algoritmos, sin una clara insistencia en otras dimensiones cruciales como la comprensión del material, la capacidad de sintetizar y evaluar las ideas, principios y prácticas, la interacción del sujeto con otras personas y con los recursos, etc. Por su parte, la enseñanza que se oriente al dominio de competencias por parte de los estudiantes, tendrá que proveer las condiciones, recursos y contextos necesarios para que ellos desarrollen y se desarrollen en el marco de las competencias que deben aprender. Existen variadas experiencias en este campo (Durkin & Main, 2002; Lammers & Murphy, 2002; Lenning & Ebbers, 1999; Taylor, Moynihan, McWilliam, & Gresty, 2004).

11. Demandas sobre la evaluación

A manera de enunciados breves, que conforman en sí un programa de trabajo, las principales demandas sobre la evaluación en los estudios universitarios se resumen como:

- superar la restricción del examen de papel y lápiz u oral, de la memorización y la aplicación de algoritmos, para centrarse en la toma razonada de decisiones y acciones profesionales en condiciones de desempeño;
- desarrollar un modelo operacional de evaluación por competencias en un marco de dominio, con referentes métricos fuertes en el caso de las competencias mayores o de salida. Existen diversas aproximaciones; por ejemplo, McInnis (2004) propone una lista de cotejo para establecer la calidad de la evaluación de los estudiantes (Ver Anexo en página 48)
- Otra demanda, conceptual y cultural se basa en la necesidad de alcanzar un nivel de holismo evaluativo, entendiendo que la evaluación es más que simplemente calificar u otorgar una nota a un estudiante;

12. Demandas sobre la manera de organizarse

Bajo este modelo de competencias, el trabajo docente no puede comprenderse sino en un referente de comunidades de aprendizaje o, al menos, comunidades docentes. Ello implica que los docentes se relacionen por escuela constituyendo comunidades en que la tarea principal sea aprender acerca de cómo enseñar más y mejor a los estudiantes concretos que dependen de ellos.

13. Demandas de cambio sobre los estudiantes

Sobre los estudiantes el nuevo modelo pone requerimientos de renovación metodológica y de hábitos y estilos de estudio.

14. Requerimientos de renovación metodológica

La renovación metodológica también es demanda de los estudiantes. En el mundo del aula, el comportamiento pasivo y silente en el asiento individual debe dar paso a la actividad y la interacción de los estudiantes entorno a una “mesa de trabajo”, en que la propia interacción sea fuente de aprendizaje significativo. Similar cosa se plantea respecto del trabajo autónomo del estudiante, en los períodos extra-aula. Esto tiene un impacto fuerte sobre estudiantes que durante años han sido entrenados para el trabajo y el rendimiento indivi-

duales, sin considerar –al momento de la evaluación- ni dar a entender que la colaboración y la actividad son dimensiones importantes del crecimiento y el aprendizaje que persigue la universidad.

15. Demandas sobre los hábitos y estilos de estudio

El modo genérico de la enseñanza básica y media ha tendido tradicionalmente a favorecer hábitos y estilos de estudio heterónomos y reactivos. Los estudiantes han pasado la mayor parte de sus vidas como tales sometidos a demandas exógenas (generadas casi exclusivamente en los profesores). Están modelados entonces por una escuela en que se premia la conducta reactiva frente a demandas más que proactiva, que responde más que pregunta, que acepta resignada antes que requerir explicaciones, que obedece más que tiene iniciativa.

La heteronomía y dependencia de los modelos estándares de aprendizaje requieren de un esfuerzo especial para lograr superar los esquemas memorísticos y algorítmicos impulsados habitualmente por la educación básica y media². Un currículo basado en competencias ciertamente exige estrategias de aprendizaje muy diferentes a las utilizadas habitualmente para aprendizajes memorísticos y mecánicos.

Un tercer elemento tiene que ver con capacidades reales de trabajo en equipo, donde se pongan en juego los distintos plexos de recursos de cada uno de los integrantes y del contexto para el logro de metas compartidas y comunes. Más que un ensamblaje de partes construidas de manera independiente (cada uno escribe una sección por su cuenta y luego las juntan), se trata de una construcción colectiva.

16. Demandas sobre la organización de los estudios

Tradicionalmente, los currícula se han ensamblado sobre la base de la división de las disciplinas, usando el esquema de ciencias básicas vs. ciencias aplicadas. En cada uno de estos conjuntos, se seleccionan subconjuntos de las diversas ciencias (en ciencias básicas, se enseñan partes de las matemáticas, o partes de la física, o partes de la sociología, o partes de

² En una mirada un tanto pesimista podríamos decir que la única solución sería que los niños ingresaran a la universidad directamente desde el jardín infantil, porque todavía no han perdido la curiosidad, la capacidad de maravillarse, la eterna pregunta del 'por qué' de todo, la relación personal y confiada con el adulto, el trabajo desinteresado con sus pares, la creatividad inagotable. Doce años de escolaridad básica y media hacen perder estas fantásticas cualidades. Eso, en la hipótesis – también aventurada- de que en la universidad las cosas son diferentes...

la filosofía), asumiendo que el aprendizaje de esas materias es funcional al aprendizaje de las “ciencias aplicadas”: la matemática y la física son funcionales al aprendizaje de la ingeniería; la biología lo es a las ciencias médicas; la sociología es funcional al administrador público, como la filosofía es funcional al aprendizaje de las diversas ramas del derecho.

Lo que está implicado en esta mirada es que cada disciplina tiene su propia naturaleza autónoma, relacionándose entre sí a modo de hipótesis pero no de manera expresa, velando con extremo cuidado los solapamientos o “invasión” de territorios que no son los propios, en una mirada de especialista. Corresponderá al estudiante producir la “síntesis” de una cantidad de aprendizajes inconexos entre sí, síntesis que finalmente es considerada como referente de la evaluación final de la profesión.

Un currículo basado en competencias pide concebir las unidades de enseñanza y aprendizaje de manera diferente. Primero, el foco de la enseñanza es el estudiante que aprende, entendiendo que el aprendizaje se produce en un marco de extrema complejidad, con lo cual los modelos reduccionistas (como los del conductualismo de raíz skinneriano) no podrán funcionar adecuadamente.

En segundo lugar, es preciso concebir que el eje de la organización curricular no son las partes de la disciplina en cuestión, sino las competencias que deben lograr el estudiante. Esto significa que la enseñanza de una competencia puede requerir del apoyo de diferentes docentes, cada uno aportando su saber especializado en un marco de colaboración con sus pares. Lograr este nivel de colaboración aparece como una de las tareas de mayor rango, complejidad y dificultad de un proceso de cambio hacia el curriculum basado en competencias.

Por consiguiente, en relación al concepto de asignatura se necesita su ruptura, flexibilizar el concepto y superar definitivamente la práctica de considerar la asignatura como coto o ghetto disciplinario.

17. Demandas sobre la administración académica

En este plano se proponen cuatro demandas críticas: flexibilización, mediación, desarrollo docente y señales institucionales.

En cuanto a flexibilización, se requiere la disolución progresiva de las rigideces impuestas por el manejo del tiempo y los recursos, posibilitada por la mayor incorporación de las TIC en los procesos formativos. El concepto tradicional del horario o del período académico (semestre o año) debe dar cabida a duraciones mayores o menores, que atiendan a las ne-

cesidades de la formación de los estudiantes, más que a la necesidad de encuadrar los tiempos de formación en marcos rígidos para facilitar los procesos administrativos y burocráticos.

El concepto de mediación aparece cuando se concibe que la evaluación es mediadora entre diferentes momentos y dimensiones de la formación del profesional. Una consecuencia mayor es la emisión de un juicio responsable que da fe pública del logro satisfactorio en el aprendizaje de una competencia o conjunto de competencias, o bien del conjunto del perfil profesional: la evaluación está mediando para posibilitar dichos juicios. Por ello es que la idea del promedio aritmético debe ser superada en la medida en que el promedio no da cuenta del logro efectivo de una competencia al ocultar los detalles que constituyen la varianza intrínseca del proceso formativo. Los supuestos para evaluar y calificar con el promedio se debilitan ante requerimientos más exigentes provenientes de las naturaleza de las tareas y los contextos de desempeño específicos de la profesión en un entorno sociocultural dado.

En tercer lugar, existe la necesidad imperiosa de que la institución haga un compromiso de apoyo al desarrollo docente. Normalmente quienes enseñan en las universidades no han sido formados profesionalmente en el campo de la pedagogía sino que vienen de sus propios campos profesionales y disciplinarios, en cuyo seno y contexto han obtenido altas calificaciones y especializaciones (grados académicos de doctor y magister). Lo notable es que las altas calificaciones en las disciplinas no necesariamente habilitan a quienes las detentan a ejercer la docencia en mejor forma. Lo corriente es que ello no suceda. Por eso es que, si la institución pretende introducir una reforma curricular del calibre de lo que implica el curriculum basado en competencias, el compromiso con sus docentes debe ser claro y explícito en cuanto a proporcionar las oportunidades y apoyos que requieran para mejorar sus desempeños, optimizar sus prácticas, y lograr mejores y más significativos aprendizajes en sus estudiantes.

Finalmente, se alude a las señales institucionales que privilegien la enseñanza, yendo un poco a contrapelo con las prácticas habituales que suelen premiar la investigación o la asistencia técnica, dejando a la docencia en una suerte de penumbra académica donde cada uno hace lo que buenamente puede.

5. Criterios y Estrategias evaluativos

18. Criterios a tener en consideración.

Las estrategias evaluativas, cualesquiera sean, deben responder a cuatro criterios fundamentales (Westat, 2002:45), a saber: credibilidad de los hallazgos, capacidades del personal, costos, y restricciones de tiempo.

La credibilidad de los hallazgos tiene que ver fundamentalmente con el tipo de información que es considerada por quienes son la audiencia o público al cual va dirigida, o bien por la cobertura del objeto. Ocasionalmente, la audiencia va a otorgar mayor credibilidad a cierto tipo de información (cuantitativa, cualitativa) así como a la instrumentación utilizada (pruebas de papel y lápiz, paneles de expertos, historias de vida). La cobertura condiciona la credibilidad en relación al tipo de información que puede requerirse: un examen nacional será más creíble si utiliza instrumentos cuyos resultados sean procesables objetivamente antes que por muchos sujetos que corrigen basados en criterios personales. En el caso específico de la enseñanza orientada a la formación profesional ...

Las capacidades del personal docente condicionan la posibilidad de llevar a cabo procesos evaluativos. Los métodos cuantitativos (producción de tests, procesamiento de datos) así como los cualitativos (entrevistas, observación en terreno, listas de chequeo, grupos de discusión) requieren una buena capacidad del personal docente. Algunos métodos pueden ser dominados con ayuda de manuales o cursos breves de entrenamiento; otros requerirán procesos más complejos o personal más entrenado para desarrollar instrumentos y gestionar la recolección y análisis de datos.

Los costos pueden interponerse en la aplicación o utilización de un método que se considere más apropiado. Un formato evaluativo fácil y rápido puede ser barato en tiempo, dinero o esfuerzo, pero no necesariamente proporcionará evidencia útil para la toma de decisiones. Se requiere desarrollar y utilizar estrategias que combinen alta eficacia en la producción de información, minimizando los costos y riesgos de su producción.

La necesidad de comparabilidad y complejidad de los datos, las capacidades instaladas y los recursos disponibles se ven también condicionados con el tiempo requerido para realizar los procesos de preparación de instrumentos, observación y análisis de los datos. Hay métodos que pueden ser más demandantes de tiempo que otros. Si el tiempo es escaso, como suele suceder en el marco de la enseñanza universitaria, se hace necesario diseñar

procesos que sean más eficientes en términos de su capacidad como dispositivos de enseñanza.

19. Estrategias apropiadas

Son estrategias evaluativas apropiadas aquellas que

- se ajustan a las características de desempeño que constituyen los indicadores de la competencia que se está evaluando
- proporcionan estimaciones apropiadas que permiten un juicio acerca de las capacidades desarrolladas por el estudiante
- integran apropiadamente –o permiten integrar– las diversas áreas de competencia: cognitiva, procedimental, interpersonal.

20. Estrategias inapropiadas

Son estrategias evaluativas inapropiadas las que

- No se asocian operacionalmente a los indicadores que revelan a la competencia evaluada
- Son reductivas en cuanto sólo consideran un área de competencia, o se refieren a parcelas del contenido, sin referencia a la integridad del desempeño de la competencia
- Se toman indiscriminadamente de un repertorio de estrategias sin considerar previamente si son pertinentes a las competencias y contextos de desempeño correspondientes a la profesión y el estilo formativo.

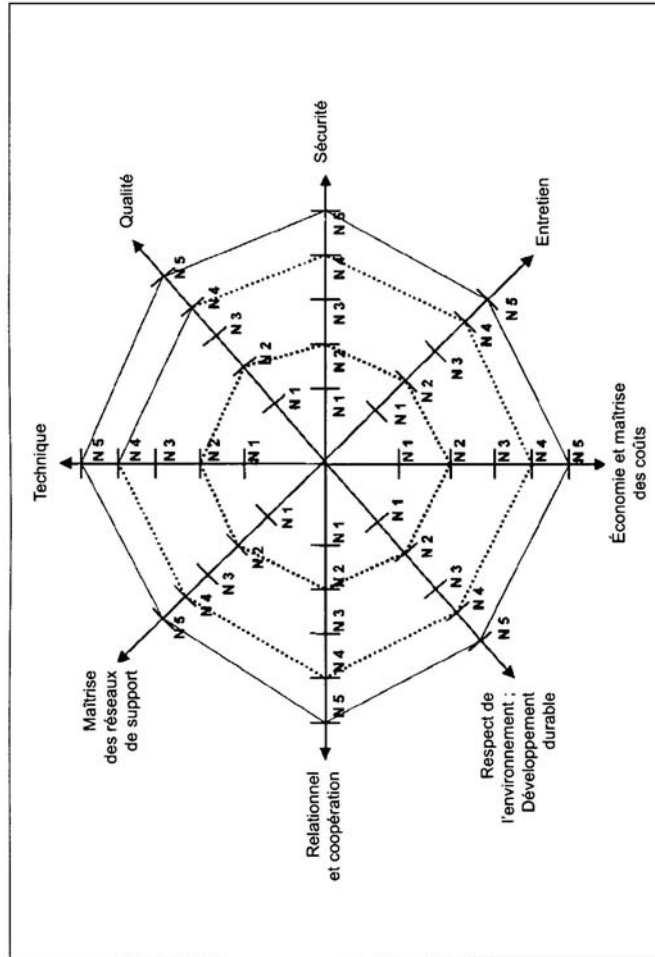
6. Medición de competencias

21. El espacio de la construcción de competencias.

Atendemos a la idea inicial de que las competencias no son unitarias sino que múltiples y que no se las puede enseñar independientemente sino que en un conjunto indisoluble en la práctica aunque analíticamente desmontable. Esto tiene dos implicancias:

primero, que las competencias constituyen un espacio pedagógico, epistémico y prático, que es a la vez unitario y complejo;

segundo, que el aprendizaje de las competencias se produce en el tiempo de manera que el logro final de las mismas debe medirse necesariamente al final del período de instrucción.



© Éditions d'Organisation

285

En relación a la primera consideración, se constituye entonces un espacio en el cual se articulan las diversas dimensiones que constituyen la competencia o, en el caso de una profesión, el perfil profesional. Este recurso, presentado por Le Boterf (2003:285) parte de la base que el desarrollo de un profesional es un proceso unitario y complejo, que debe ser logrado con niveles de suficiencia en todos sus aspectos para poder lícita y éticamente acreditarlo ante la sociedad. Si un sujeto no logra el nivel de suficiencia o “umbral”³ en alguna de las dimensiones consideradas necesarias para la constitución de la competencia o

³ En inglés se lo traduce como “threshold”, literalmente “umbral”.

dominio, entonces no puede certificarse su logro. Esta noción se asocia al carácter no-compensatorio (cfr. pág. 16) de los logros en las competencias.

El nivel “suficiencia” representa, como se indica en 12, una calidad tal que el sujeto se desempeña sin representar un peligro ni una amenaza para los demás, para sí mismo, para la organización y para la sociedad. Esto tiene que ver con la eficacia (capacidad de producir los resultados) pero también con la eficiencia (costos asociados a la producción de los mismos).

En relación a la primera consideración se constituye, entonces, un espacio en el cual se articulan las diversas dimensiones que constituyen la competencia o, en el caso de una profesión, el perfil profesional. Este recurso, presentado por Le Boterf (2003:285) parte de la base que el desarrollo de un profesional es un proceso unitario y complejo, que debe ser logrado con niveles de suficiencia en todos sus aspectos para poder, lícita y éticamente, acreditarlo ante la sociedad. Si un sujeto no logra el nivel de suficiencia o “umbral” en alguna de las dimensiones consideradas necesarias para la constitución de la competencia o dominio, entonces no puede certificarse su logro. Esta noción se asocia al carácter no-compensatorio de los logros en las competencias.

En cuanto al carácter evolutivo del aprendizaje de las competencias, ello se asocia a las cuestiones de enseñanza o didáctica de competencias, que no son abordadas aquí.

22. Competencias y medición.

No es posible evaluar y emitir un juicio de logro o no logro si no se puede observar sistemáticamente y clasificar las conductas relacionadas con las competencias. Toda competencia ha de expresarse en indicadores o comportamientos que puedan ser observados de alguna forma. La observación, por otra parte, requiere de instrumentos que den cierta garantía de estabilidad de las observaciones y comparabilidad de los registros a lo largo del tiempo y entre diversos grupos de estudiantes o sujetos que están siendo evaluados.

En la tabla siguiente se presentan los principales dispositivos relacionados con la medición de indicadores relacionados con las diversas áreas de competencia.

Tabla 1 - Recursos evaluativos asociados a ámbitos C, P, IP

Recurso evaluativo	Dimensiones asociadas en cada ámbito		
	Cognitivo	Procedimental	Intra-Interpersonal
Informes escritos: informes de lectura, monografías, desarrollo de guías o protocolos	Pensamiento crítico; síntesis-análisis. Pensamiento crítico. Pensamiento complejo.	Cumplimiento de normas de protocolos y guías; capacidad heurística	Responsabilidad
Informes orales sobre lecturas, guías o protocolos, monografías, experiencias en terreno	Pensamiento crítico; síntesis-análisis. Pensamiento crítico. Pensamiento complejo.	Cumplimiento de normas de protocolos y guías; capacidad heurística. Capacidades discursivas	Responsabilidad. Autoconfianza; relación con los demás; enfrentamiento de una audiencia; manejo del estrés
Presentaciones en formato multimedial	Pensamiento analítico-sintético. Pensamiento crítico. Pensamiento complejo.	Uso de recursos multimediales en la diagramación estética del contenido del mensaje.	Sensibilidad estética. Responsabilidad. Autoconfianza; relación con los demás; enfrentamiento de una audiencia; manejo del estrés
Exámenes orales de la materia de una asignatura	Pensamiento analítico-sintético. Pensamiento crítico. Pensamiento complejo. Capacidad de argumentación. Dominio de la dialéctica.	Manejo escénico, del habla, del espacio. Organización y concatenación de ideas. Ejemplificaciones apropiadas.	Autoconfianza; relación con los demás; enfrentamiento de una audiencia; manejo del estrés
Exámenes escritos de respuesta abierta de la materia de una asignatura	Memoria y retención de información. Comprensión y aplicación de conceptos, principios y teorías. Capacidades analítica, sintética y evaluativa de materiales conceptuales.	Organización de ideas. Producción de textos con sentido. Estructuración de un texto.	Prolijidad del trabajo. Honestidad en la producción de las respuestas.

	Capacidad de argumentación.		
Exámenes escritos de respuesta cerrada de la materia de una asignatura	Memoria y retención de información. Discriminación de contenido veritativo de enunciados.		Honestidad.
Listas de chequeo		Ejecuciones; Actividades de laboratorio Ejercicios de simulación, Actividades en terreno, Desempeño práctico con otras personas	Cantidad y frecuencia de las relaciones. Contenido de las interacciones.
Escalas de apreciación		Ejecuciones; Actividades de laboratorio Ejercicios de simulación, Actividades en terreno, Desempeño práctico con otras personas	Calidad de las relaciones. Contenido de las interacciones.

Existen muchos otros recursos evaluativos, entre los que pueden citarse: Informes de supervisores, Construcción de proyectos, Realización de investigaciones, Auto-informes – individuales y grupales, Portafolio de evidencias, Recursos biográficos: memorias, diarios, autobiografías; Desarrollo de proyectos individuales; Desarrollo de proyectos grupales; Actividades de pre-práctica profesional; Actividades de práctica profesional; Entrevistas.

Cada uno de éstos tiene un potencial como recurso evaluativo que puede ser explorado teniendo a la vista las condiciones específicas de la enseñanza.

Los diferentes dispositivos tienen propiedades que los caracterizan, a la vez que limitan y delimitan su aplicabilidad a diferentes tipos de comportamiento así como diferentes áreas disciplinarias.

Su utilización se justifica según se insertan en determinados enfoques o estrategias evaluativas. Entre las más conocidas se encuentra la de portafolios, que últimamente ha recibido mucha atención y que se describe a continuación. Otro enfoque estratégico, de historia más antigua, es el llamado “testing” que se basa en una batería de exámenes estandarizados y de puntuación objetiva; entre las utilidades más conocidas en nuestro país se encuentran las pruebas SIMCE y PSU; las primeras se orientan a medir competencias cognitivas en el sistema educacional con propósitos de evaluación del sistema; las segundas pretenden medir competencias cognitivas en sujetos que pretenden proseguir estudios universitarios –siendo una prueba de selección, la PSU no tiene entonces carácter de evaluación del sistema medio, sino que de ordenamiento de postulantes para el proceso de selección y admisión. Finalmente, no deja de ser importante la estrategia más común entre nosotros y que consiste en un enfrentamiento de la evaluación por la vía del reduccionismo del promedio, calculado sobre observaciones de diferente naturaleza, de distinto rango, pero asumiendo que por ser numeral una escala es necesariamente numérica.

23. Portafolios

Douwe K. Wielenga (2002) describe la forma en que conciben y utilizan la metodología de portafolio como dispositivo evaluativo en la Amsterdam Faculty of Education⁴ - se trata de un portafolios basado en la web.

⁴ Cfr. el sitio <http://www.efa.nl/publicaties/english.html>

Lo más notable, de acuerdo a la autora es que “el sistema de portafolios ayuda a la gente, bajo su propia responsabilidad, a reunir materiales que pueden ser usados y reutilizados en diferentes momentos con diferentes propósitos:

- aprender (poner a prueba el propio proceso de aprendizaje, a menudo con ayuda de retroinformación por parte de otras personas)
- evaluar (probar la competencia frente a una autoridad externa; por ej., ser admitido a la siguiente etapa del curso, o conseguir un trabajo, o ser admitido a un nivel superior en el trabajo)
- compartir (productos y experticia deberán ser parte de un conocimiento base; la gente “desconocida” debería poder encontrarlos).

Una ventaja del portafolios basado en web es que uno puede usar los mismos materiales para esos diferentes propósitos. A lo largo del tiempo, los materiales pueden ser ordenados según diferentes propósitos en diferentes oportunidades. Para cada propósito se toma una selección de los materiales y se los agrupa (con una línea de historia –storyline) en una estructura que sirva a tal propósito (Wielenga, 2002:12)

Un portafolios puede contener:

Artefactos: materiales producidos por los alumnos como resultado de su paso por el curriculum. Se incluye aquello que puede ser útil para los propósitos que piensa el dueño del portafolios. Siempre es mejor tener de más que de menos. Ejemplos: grabaciones video, un examen, artículos o papers, sitio web, resultado de una serie de lecciones dictadas, la respuesta que dio a una pregunta abierta en una prueba de comprensión

Reproducciones: artefactos producidos fuera del contexto del curriculum. No fueron asignados, pero pueden ser muy importante. Puede usar cualquier aspecto de su vida: dirigente de boy scouts; comentario a una presentación musical amateur. Cosas así.

Testimonios: retroinformación de otros. Su opinión sobre su trabajo, su conocimiento, sus destrezas, su forma de cooperar. Podría ser únicamente una nota en un test. Podría ser una evaluación oral o escrita sobre el rol que Ud. jugó en un proyecto. Podría ser una respuesta a una pregunta evaluativa que Ud. respondió.

Afirmaciones: en ocasiones uno debe hacer afirmaciones acerca de sí mismo. Breves y poderosas. ¿En qué dirección voy? O específicamente, ¿en qué dirección NO voy? ¿Dónde estoy hoy con relación a ciertas metas de competencia y cómo voy a proceder? Por cierto, estas afirmaciones se estructuran con hipervínculos a los artefactos, reproducciones o testimonio en el portafolios.

Comentarios: es necesario dar un significado a todo material del portafolio (artefactos, reproducciones, testimonios, afirmaciones), usando explicaciones breves para mantenerlos conectados a su fuente e historia. Otros podrían aportar significado señalando las razones por qué el material en cuestión se incorpora al portafolios. También pertenecen aquí las categorías para estructurar los materiales.

Creaciones: la información no sólo se relaciona con archivos de texto y otros medios como película, sonido, imagen. Una cantidad de información se encuentra en la *forma* en que se presenta el material: una metáfora usada consistentemente puede ser muy convincente o entusiasmante; o una forma inesperada de ordenar el material que muestra una visión inesperada. Así, las creaciones son descubrimientos e invenciones en comunicación en citas educativas usando portafolios. (Wielenga, 2002)

24. Listas de chequeo.

Particularmente apropiadas para observar y registrar situaciones de desempeño. Se pueden desarrollar y aplicar para ayudar a los docentes a formular juicios acerca del logro de competencias en áreas afectivas y de relaciones interpersonales, en las cuales no es funcional el modelo tradicional de la evaluación fundada en conocimientos. Las listas de chequeo se responden marcando en una escala dicotómica, de dos puntos: Sí o No. No se consideran juicios intermedios.

Casos notables son los del área médica. Por ejemplo, se encuentra la **Interpersonal and Communication Skills Checklist** (Cohen, Colliver, & Marcuy, 1996) Algunos ítemes de ejemplo de esta última son:

	Muy de acuerdo	De acuerdo	Quizás acuerdo	Indeciso	Quizás desacuerdo	Desacuerdo	Muy desacuerdo
<i>El doctor quería comprender cómo veía yo las cosas</i>	7	6	5	4	3	2	1
<i>El doctor me dio oportunidades de hacer preguntas</i>	7	6	5	4	3	2	1
<i>El doctor me explicó las cosas de manera que ahora sé qué es lo que me está sucediendo</i>	7	6	5	4	3	2	1
<i>Me siento satisfecho(a) con el cuidado médico recibido</i>	7	6	5	4	3	2	1

Existe otra escala que se usa para evaluar la competencia de interacción médico-paciente en un estudiante de medicina. Algunos de sus ítemes son:

El estudiante inició la entrevista presentándose a sí mismo(a)

Sí	<input type="checkbox"/>	No	<input type="checkbox"/>
Sí	<input type="checkbox"/>	No	<input type="checkbox"/>
Sí	<input type="checkbox"/>	No	<input type="checkbox"/>
Sí	<input type="checkbox"/>	No	<input type="checkbox"/>
Sí	<input type="checkbox"/>	No	<input type="checkbox"/>
Sí	<input type="checkbox"/>	No	<input type="checkbox"/>

El estudiante se veía confiado(a)

El estudiante explicó cómo iba a proceder en mi evaluación o tratamiento

El estudiante me permitió hacer preguntas

El estudiante mantiene contacto visual

El estudiante mantuvo una distancia cómoda y apropiada durante la entrevista

25. Escalas de apreciación.

Las escalas de apreciación, a diferencia de las listas de chequeo, permiten al observador hacer un juicio más ponderado porque admiten una gama de respuestas más amplia que Sí o No. Normalmente hacen uso de escalas numéricas en que los números representan cierto “monto” de la variable como es el caso de “acuerdo-desacuerdo” o “siempre-nunca”.

Entre éstas, pueden citarse las siguiente.

- **Humanism Scale** (Linn, Dimatteo, Cope, & Robbins, 1987)
- **Patient Assessment of Interpersonal and Communication Skills**, de la American Board of Internal Medicine

- **Interpersonal Skills Rating Form.** (Schnabl, Hassard, & Kopelov, 1991)

	Completo acuerdo	Acuerdo	Posible acuerdo	Indiferente	Posible desacuerdo	Desacuerdo	Completo desacuerdo
1.- El médico buscaba comprender las cosas como yo las veía	7	6	5	4	3	2	1
4.- La respuesta del médico fue generalmente tan fija y automática que no pude comunicarme con él	7	6	5	4	3	2	1
5.- El médico me trató con respeto y cortesía	7	6	5	4	3	2	1
7.- El médico me explicó las cosas de manera que yo supiera qué pasaba conmigo	7	6	5	4	3	2	1
11.- El médico utilizó un lenguaje comprensible y no técnico	7	6	5	4	3	2	1
13.- Me siento satisfecho(a) con el cuidado médico que he recibido	7	6	5	4	3	2	1

7. Calificaciones

26. Conceptos generales

La calificación es la expresión del juicio evaluativo utilizando una escala numérica, suponiendo que los diferentes valores de la escala representan diversos juicios ordenados de mayor a menor logro. Una condición necesaria para poder expresar el juicio evaluativo en numerales es describir con la mayor propiedad y claridad posible, qué representa cada grado de la escala. Los valores tienen sentido porque hay un contenido previamente acordado, consensuado y cuya descripción es pública para los diferentes actores que participan en el proceso evaluativo.

Utilizar escalas numéricas sin previamente acordar o tener definido el nivel de desempeño que representa cada uno de los puntos de la escala puede llevar a manipulaciones indebidas de los datos. Entre éstas, una de las más comunes es la del cálculo estadístico (promedios, varianzas, correlaciones, análisis predictivo, etc). Las escalas de calificación, por la propia naturaleza de las variables educativas, no pueden ser más allá del nivel ordinal: los puntos están ordenados monotónicamente pero no conocemos ni puede establecerse la

equivalencia entre los diferentes puntos de la escala. Ello hace viables ciertas operaciones estadísticas pero no otras. Sin embargo, la consideración meramente numérica de la escala lleva a auténticas perversiones.

El criterio fundamental para establecer calificaciones requiere aclarar dos componentes: (a) los significados de la escala y (b) las dimensiones de la escala.

Los significados de la escala implican

- la descripción de los estándares de desempeño requeridos para certificar el dominio de una competencia
- la determinación de los niveles que describen los diversos niveles de logro en torno (sobre y bajo) el estándar de desempeño definido
- la comparabilidad y conversión entre escalas

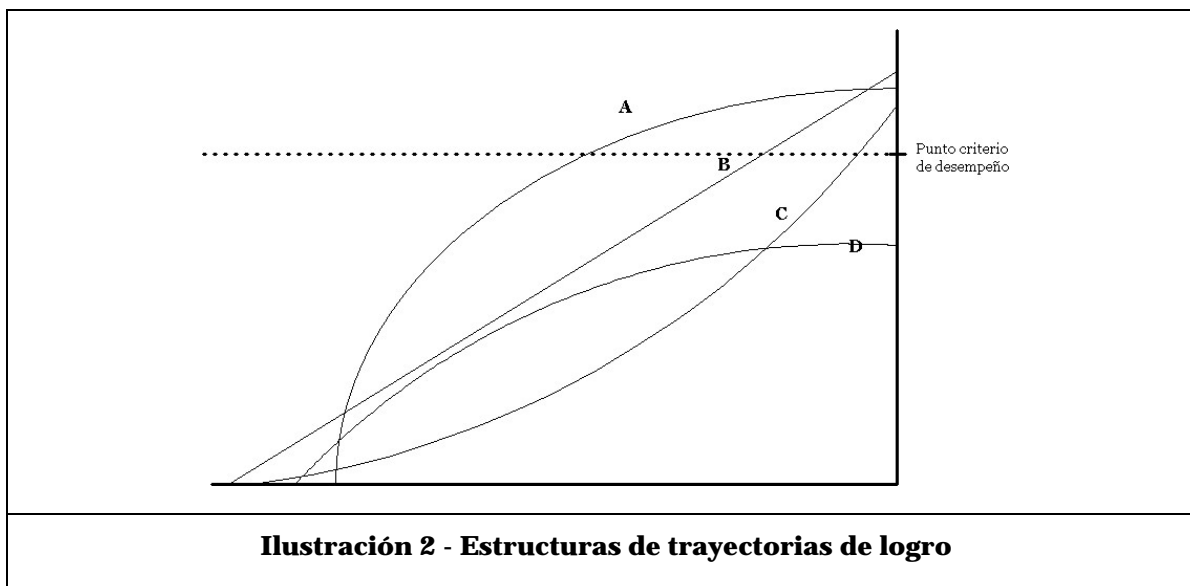
Las dimensiones de la escala, por su parte, tienen que ver con decidir acerca de

- la naturaleza de la escala: letras, puntos, conceptos
- la amplitud de la escala en puntos (5, 7, 10, 20, 100 puntos), en letras (A, B, C, D, E), en conceptos (Insuficiente, Suficiente, Bueno, Muy Bueno)
- la existencia de puntos de corte predefinidos (nota de aprobación) o criterios normativos para su definición

27. Representación de los espacios evolutivos de competencias

Una situación altamente compleja se presenta al docente cuando se ve enfrentado a la necesidad de evaluar a un sujeto en términos de su evolución en el dominio de una competencia. Podemos pensar razonablemente que un estudiante, cuando inicia un módulo de aprendizaje, no es capaz de desempeñarse en la competencia que es el objeto de la enseñanza; si el proceso ha sido el apropiado, también es posible pensar que al final del módulo el que era ignorante ahora sabe, el que era incompetente ahora es competente. Es decir, ha habido un cambio entre el estado inicial y el estado final.

Desde el punto de vista de su representación gráfica, podemos pensar en una curva o pendiente monotónica creciente que se mueve entre el punto de inicio (el cero) y el punto de logro, tal como aparecería en la ilustración siguiente. Como puede advertirse, no importa tanto la forma de la curva, siempre que su punto final corresponda al área delimitada por el punto de corte del nivel de logro esperado.



En relación al punto de inicio, puede esperarse razonablemente que no todos los estudiantes parten del mismo punto: ello depende de sus propias historias académicas, orientaciones disciplinarias⁵. Ciertamente un grupo de características más bien homogéneas siempre será una ventaja para el docente y la docencia, toda vez que no se ve obligado a proporcionar información ni ejercitación diferenciadas según los niveles de entrada de los estudiantes.

En la ilustración se representan cuatro casos hipotéticos. Los tres primeros, A a C, son sujetos que partiendo de condiciones diferentes a la entrada, todos ellos logran superar el punto de quiebre que define el desempeño satisfactorio del insatisfactorio. No obstante, el caso D no logra el nivel de desempeño requerido, aunque su nivel de inicio es comparable con el de sus condiscípulos. ¿Cómo leer esta información? ¿Qué decisiones corresponde tomar al respecto?

Los datos de la ilustración son completamente insuficientes para cualquier aproximación a las razones de que D no haya logrado el desempeño esperado. Éstas deberán ser proporcionadas por el docente luego de una apropiada investigación y análisis de las condiciones de enseñanza, de aprendizaje, y las circunstancias en que se encuentra el estudiante en relación al programa del curso.

⁵ En particular cuando se trata de asignaturas o módulos que son cursados por estudiantes de diferentes carreras, lo que introduce mayor heterogeneidad o varianza al grupo.

Las decisiones, por su parte, son variadas. Primeramente, una decisión profesional siempre debe ser fundada, no bastando la constatación de que un sujeto no respondió apropiadamente una exigencia de desempeño (un examen, una situación práctica, unos pasos de laboratorio). Un docente profesional debe estar en condiciones de decir y justificar por qué el estudiante D de la ilustración no logró el desempeño esperado, pero sí lo hicieron sus compañeros. Ahora bien, entre las posibles decisiones asociadas se encuentran la de simplemente reprobado al sujeto obligándolo a cursar nuevamente la asignatura (un semestre o un año después); también, puede dársele la oportunidad de responder a la misma exigencia con un tiempo adicional, o una exigencia alternativa pero equivalente, o proporcionando ayuda adicional. En fin, existe un sinnúmero de posibilidades para enfrentar estas situaciones. Se trata ahora de que los docentes también asuman la docencia como una actividad netamente profesional y académica, con igual rango que la investigación o la administración académica.

28. El modelo de QAA para escalar las apreciaciones.

La Quality Assurance Agency (Agencia para Aseguramiento de la Calidad) del Reino Unido propone un modelo conceptual para concebir la evaluación como acreditación del logro de competencias, muy consistente con la propuesta anteriormente (ver la ilustración).

La QAA distingue primeramente lo que denomina, en inglés, “threshold”, literalmente umbral. Éste delimita el punto en que se entiende que un sujeto tiene un logro de la competencia tal que hace posible certificarlo ante una comunidad determinada.

Bajo el punto de umbral se encuentra un segmento –de tamaño variable- en que se ubican aquellos que no han logrado o alcanzado el umbral de rendimiento y, por consiguiente, no serán certificados. Sobre el punto de acreditación se encuentran dos grupos claramente identificables –aunque sus límites no sean completamente claros: uno, el grupo modal, correspondiente a la “mayoría” de los acreditados, que se desempeñan habitualmente bien en sus profesiones; el otro, el grupo destacado o “top” que sobresalen por la excelencia de su trabajo, su creatividad u otras dotes propias del desempeño de la profesión.

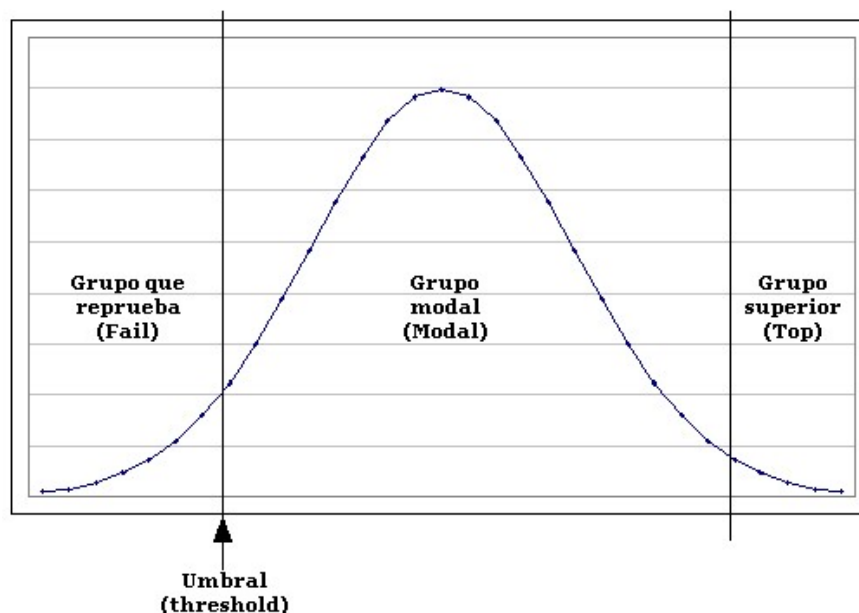


Ilustración 3 - Estructura de logros en el model del QAA

Con todo, la propuesta del QAA en su formato gráfico puede prestarse a tergiversaciones, entendiéndose erradamente que los resultados evaluativos deben distribuirse siguiendo la curva normal de probabilidades, con un cierto promedio y una cierta desviación estándar, tal que ahora pueden fijarse “cuotas” para los extremos. Así, por ejemplo, podría decidirse que el 5% inferior corresponde al grupo de reprueba. ¿Y si hay un 12% que no ha logrado la competencia? ¿Significa esto que el 7% restante quedará aprobado por obra y gracia de la magia de la curva normal? Más grave aún es el caso de que todos aprueban y por adhesión mecánica y acrítica a un modelo se obliga a que el 5% más bajo (pero que han dominado satisfactoriamente la competencia) se vea reprobado, injustamente por cierto.

Por el contrario, si entendemos que se trata de conceptos más que de funciones matemáticas podemos concebir “porcentajes” diferentes de “casos” en cada uno de los sectores en que se divide el grupo.

29. Una propuesta desde la perspectiva de la gestión de recursos humanos

Una propuesta interesante es la propuesta por M.A. Alles (2004) con orientación al comportamiento en la empresa en la perspectiva de la gestión de recursos humanos. Si bien no es utilizable de manera directa en la formación de profesionales, puede aprenderse de la manera en que define una escala de cuatro grados que designa con letras:

Grado A: logro del nivel más alto de comportamiento; identificado con el 100% de una escala;

Grado B: segundo nivel, corresponde al 75% de la escala;

Grado C: tercer nivel, corresponde a logro de 50% de la escala;

Grado D: cuarto nivel, ubicado en el 25% de la escala, descrito como “competencia en su grado mínimo”;

(Sin denominación): quinto nivel, correspondiente al 0%; descrito como “competencia NO desarrollada”.

Ocasionalmente utiliza sólo tres grados, eliminando el correspondiente a 25%.

Dada la orientación a la gestión de recursos humanos, es comprensible cuando dice que *“hemos relacionado los comportamientos de los niveles más altos, grados A o B, con posiciones jerárquicas o de mayor nivel; y los grados C o D –este último indica un nivel mínimo de la competencia- fueron relacionados con posiciones de menor nivel en la estructura organizacional”* (2004:60). Es claro a partir de esta expresión que el propósito de la escala tiene que ver con selección y ubicación del personal más que con cuestiones relacionadas con formación y promoción profesional.

Un ejemplo es la definición y escala propuestas para la competencia de “prudencia” que se transcribe más abajo (2004:70-71).

¿Qué puede aprenderse a partir de esta propuesta? En primer lugar, que es posible hacer descripciones de las competencias, aunque éstas no siempre sean todo lo claras que podría desearse (de hecho, las definiciones de M.A. Alles suelen incorporar términos sumamente ambiguos y polisémicos); también parece que es posible hacer un esfuerzo por utilizar términos menos ambiguos o bien fijar significados previa convención.

Tabla 2 - Ejemplo de niveles de competencia según MA Alles

<p>PRUDENCIA: Sensatez y moderación en todos los actos, en la aplicación de normas y políticas de la organización, sabiendo discernir lo bueno y lo malo para la empresa, para el personal y para sí mismo. Implica también que piensa y actúa con sentido común. Ambas características, sensatez y moderación, y sentido común, aplicadas en todos los actores, en todo momento, en todos los aspectos de la vida.</p>
<p><i>Comportamientos habituales con relación al discernimiento, la sensatez y la moderación en los ámbitos personal y profesional.</i></p>
<p>Grado A.</p> <ul style="list-style-type: none"> • Establece los objetivos organizacionales guiándose por la sensatez y la moderación • Tiene presente lo bueno y lo malo para la empresa, para sí y para su personal, al definir las políticas de la organización • Es reconocido por su capacidad para analizar y resolver situaciones basándose en el sentido común • Es moderado y sensato en todos sus actos, tanto en su vida laboral como personal • Da <i>coaching</i> a su gente, para desarrollar su capacidad de actuar con prudencia y premia sus logros en este sentido. <p>Grado B.</p> <ul style="list-style-type: none"> • Adopta e implementa las políticas de la empresa, traduciéndolas a objetivos y acciones claros, basados en la sensatez y la moderación • Guía a sus equipos de trabajo con discernimiento de lo que puede ser bueno o malo para todos y para cada uno de sus miembros • Planifica, diseña e implementa acciones con sentido común • Motiva y estimula a su gente a actuar con moderación, sensatez y sentido común • Es moderado y sensato en todos sus actos laborales, tanto en relación con sus tareas como con la gente <p>Grado C.</p> <ul style="list-style-type: none"> • Establece procesos y ámbitos de trabajo orientados al logro de los objetivos organizacionales, basándose para su análisis, planificación, diseño y adjudicación de responsabilidades, en su sensatez y moderación • Comprende lo que puede ser bueno o malo para su equipo de trabajo o sus pares, y se esfuerza por orientar su accionar sobre la base de este conocimiento • Actúa ante los problemas que se le plantean con sentido común, moderación y sensatez • Brinda asesoramiento cuando se lo consulta sobre cuestiones laborales o personales, basándose en su sentido común <p>Grado D.</p> <ul style="list-style-type: none"> • Lleva a cabo sus tareas y objetivos con moderación y sensatez • Tiene discernimiento respecto de lo que es bueno o malo para sí o para sus compañeros, y actúa de acuerdo con este conocimiento, cuando las situaciones lo permiten • Consulta cuando tiene dudas para actuar de acuerdo con el grado de moderación que se requiere para una situación imprevista • Actúa con sensatez y moderación en situaciones conocidas <p>Competencia NO desarrollada.</p> <ul style="list-style-type: none"> • Sus actos son imprudentes o faltos de moderación, lo que genera riesgos para la operación del negocio y para su gente • Analiza y resuelve situaciones con falta de sentido común • Demuestra escasa preocupación por lo que pueda ser bueno o malo para su equipo de trabajo, mientras se alcanzan los resultados o beneficios que él espera • Sólo se preocupa por su propio bienestar o beneficios, sin que le importe lo que suceda con los demás en la organización • Ante situaciones imprevistas reacciona emocionalmente o intempestivamente, sin darse tiempo para reflexionar sobre cuál puede ser la mejor alternativa de acción

Lo segundo, es que los diferentes niveles que se establezcan tienen una relación de continuidad creciente (o decreciente según se lo mire en sentido inverso) de tipo monótonico. Además, cada nivel supone la adquisición o dominio del nivel inferior en la escala: el grado A supone el logro del grado B, éste el logro del grado C, y así sucesivamente. La relación de implicancia podría dar pie para pensar en una suerte de invariancia (las etapas se logran progresivamente) e irreversibilidad (no se vuelve atrás), en el sentido de la epistemología genética piagetana; sin embargo, no se tiene evidencia alguna que confirme o rechace esta presunción.

30. Transformación a nota

La transformación de la evaluación a nota o calificación reviste problemas no menores, aunque a veces llegue a confundirse una con la otra. Conviene aclarar desde el inicio, a partir de los conceptos planteados en secciones anteriores, que las escalas son convencionales; que los puntos de la escala (incluyendo los extremos) por lo general no tienen definición; que las competencias no siempre son promediables.

Entre los recursos más utilizados, aunque no siempre con fundamento, se encuentran los siguientes:

- a. Conversión lineal entre la escala de puntos de un examen y la escala de calificaciones utilizada.

El procedimiento consiste en simplemente contar el número de puntos de un examen y luego dividir el recorrido de la escala por el puntaje máximo resultante. Se logra así un factor que multiplicado por el número de puntos obtenido y sumado al valor inicial de la escala, se convertirá en la calificación. En este caso, la calificación de aprobación se considera el punto medio de la escala (si la escala tiene 30 puntos, la nota de aprobación es la que corresponde a 15 puntos; si tiene 40 puntos, entonces es la que corresponde a 20 puntos).

Por ejemplo,

- un examen tiene un máximo de $P_{je} = 30$ puntos;
- si se usa la tradicional escala de 1 a 7 se tiene un recorrido de $R = 7 - 1 = 6$ puntos.
- al dividir el recorrido de la escala por el puntaje se obtiene el factor $f = 6/30 = 0.2$.

- El valor inicial o mínimo de la escala de calificaciones es $\text{Mín} = 1$ (no existe el cero en la escala de notas).
- Un estudiante “i” que obtiene $P_{je}(i) = 18$ puntos en el examen tendrá una calificación de $1 + 0.2 \cdot 18 = 1 + 3.6 = 4.6$.
- La expresión general del procedimiento es:

$\text{Nota} = \text{Mín} + (R/P_{je}) * P_{je}(i)$

- b. Conversión semi-lineal, a partir de un criterio de suficiencia previamente establecido.

Esta modalidad es utilizada cuando se asigna previamente un criterio para la asignación de la nota de aprobación. Por ejemplo, se puede indicar que para obtener la nota de aprobación mínima, un estudiante debe conseguir el 60% (u otro valor) del puntaje total. Lo que resulta de cortar la escala en un punto diferente del punto medio es que se tienen dos tramos de desigual longitud, de manera que se hace preciso hacer un doble cálculo.

Por ejemplo,

- un examen tiene un máximo de $P_{je} = 30$ puntos;
- se decide que para aprobar se requiere acumular 20 puntos; en la escala tradicional de 1 a 7, la nota 4 corresponde a 20 puntos en el examen.
- Para determinar la conversión entre 4 y 7, se divide el tramo $7-4=3$ entre la diferencia que hay entre el puntaje máximo y el puntaje de corte, es decir, $30-20 = 10$.
- al dividir el recorrido del segmento superior de la escala por el puntaje se obtiene el factor $f = 3/10 = 0.3$.
- Para determinar la conversión bajo 4, se divide el tramo $4-1=3$ entre el puntaje de corte, es decir, $3/20 = .15$
- Las reglas decisionales son las siguientes:
 - Si un estudiante tiene un puntaje igual a 20, su calificación de 4.0
 - Si tiene un puntaje mayor que 20, su calificación se calcula como

$$\text{Nota} = 4 + (\text{Pje}-20)*.3$$

Un estudiante obtuvo 25 puntos en el examen. Por consiguiente su calificación será de $\text{Nota} = 4 + (25-20)*.3 = 4 + 5*.3 = 4 + 1.5 = 5.5$

- Si tiene un puntaje menor que 20, su calificación se calcula como

$$\text{Nota} = 1 + \text{Pje}*.15$$

Un estudiante obtuvo 17 puntos en el examen. Por consiguiente su calificación será de $\text{Nota} = 1 + 17*.15 = 1 + 2.55 = 3.55$

c. Asignación por cuotas

El modelo de asignación por cuotas parte de una distribución previa de la frecuencia de casos en cada tramo de la escala. Esta modalidad es especialmente apropiada cuando se trata de procesos de selección en que ciertamente el propósito no es pedagógico (formativo) sino que de otra naturaleza.

Existen diversas aproximaciones al cálculo de la calificación bajo este modelo. Las más frecuentes son:

- determinar el número o porcentaje de casos que serán considerados para la aprobación sobre la base de algún criterio preestablecido;
 - se procede luego a la evaluación de los productos de examen y se los rankea en orden decreciente;
 - si el criterio es un número (por ejemplo, 15), se cuentan los 15 primeros y se consideran aprobados; desde el caso 16 en adelante se consideran reprobados;
 - si el criterio es un porcentaje (por ejemplo, el 25%), se cuenta el total de casos, se calcula el porcentaje, y con el número resultante se procede de la misma manera que en el caso anterior
- fijar las cuotas sobre la base de alguna distribución teórica, generando así porcentajes de sujetos que serán asignados a los diferentes tramos; en este campo es frecuente encontrar referencias a la distribución normal de probabilidades o “campana de Gauss” o simplemente “la gaussiana”.

- Por ejemplo, puede decidirse que la nota de aprobación corresponderá a quienes estén entre -0.5 y $+0.5$ unidades de desviación estándar; el tramo superior, entre 0.5 y 1.5 ; el superior, entre 1.5 y 2.5 , etc.
- Se calculan los puntajes, su promedio y desviación estándar;
- Se procede luego a calcular los puntajes de corte entre los tramos;
- Se asignan calificaciones según los puntajes de cada sujeto se encuentran en uno de los tramos definidos.

8. Referencias.

- Abbott, J., & Ryan, T. (1999). Constructing knowledge, reconstructing schooling. *Educational Leadership*, 57(3), 66.
- Algozzine, B., Bateman, L. R., Flowers, C. P., Gretes, J. A., Hughes, C. D., & Lambert, R. (1999). Developing technology competencies in a college of education. *Current Issues in Education*, 2(3).
- Alles, M. A. (2004). *Diccionario de comportamientos. Gestión por competencias. Cómo descubrir las competencias a través de los comportamientos*. Buenos Aires: Ediciones Granica S.A.
- Buck, J.-Y. (2003). *Le Management des Connaissances et des Compétences en pratique*. Paris: Éditions d'Organisation.
- Chickering, A. W., & Gamson, Z. F. (1987). Seven Principles of Good Practice in Undergraduate Education. *American Association of Higher Education Bulletin*, 39(7), 3-7.
- Christiansen, M. D., Roberts, M., Borden, K., Belar, C., Breault, L., Edwards, H., et al. (2004). *Assessment of Competence. The Ultimate Frontier for Improving Professional Competence*.
- Cohen, D. S., Colliver, J. A., & Marcuy, M. S. (1996). Psychometric properties of a standardized-patient checklist and rating form used to assess interpersonal and communication skills. *Academia Medica*(71), S87-S89.
- Durkin, K., & Main, A. (2002). Discipline-based study skills support for first-year undergraduate students. *Active learning in higher education*, 3(1), 24-39.
- González, J., & Wagenaar, R. (Eds.). (2003). *Tuning Educational Structures in Europe. Informe Final. Fase Uno*. Bilbao (España): Universidad de Deusto y Universidad de Groningen.
- Hamilton, L. S., McCaffrey, D. F., & Stecher, B. M. (2003). Studying Large-Scale Reforms of Instructional Practice: An Example from Mathematics and Science. *Educational Evaluation & Policy Analysis [H.W. Wilson - EDUC]*, 25(1), 1.
- Hirsch, W. Z., & Weber, L. E. (1999). *Challenges Facing Higher Education at the Millenium*. Oxford: Pergamon/Elsevier, IAU.
- Hirsh, E. D. (1999). Equidad y excelencia: Metas alcanzables en educación. *Estudios Públicos*(73), 165-178.
- Hirsh, S. (2004). Putting Comprehensive Staff Development on Target. *Journal of Staff Development*, 25(1), 12.

- Lammers, W. J., & Murphy, J. J. (2002). A profile of teaching techniques used in the university classroom. A descriptive profile of a US public university. *Active learning in higher education*, 3(1), 54-67.
- Le Boterf, G. (2003). *Ingénierie et Évaluation des Compétences*. Paris: Éditions d'Organisation.
- Lenning, O. T., & Ebbers, L. H. (1999). *The powerful potential of learning communities : improving education for the future*. Washington, DC: Graduate School of Education and Human Development George Washington University.
- Linn, L. S., Dimatteo, M. R., Cope, D. W., & Robbins, A. (1987). Measuring physicians' humanistic attitudes, values, and behaviors. *Medical Care*, 25(6), 504-515.
- McInnis, C. (2004, Feb 5, 2004). *Repositioning Assessment to Enhance Learning Outcomes*. Paper presented at the Quality Enhancement Workshopt. Using Assessment to Motivate Learning, Glasgow. <http://www.cshe.unimelb.edu.au/assessinglearning>; Accesado en Octubre 28 de 2004.
- Proffer, L. (1994). A vision of competence. *State Legislatures*, 20(2), 12+.
- Reiger, R. C. (2001). Graduating seniors make democracy become alive by evaluating their high school program. *Education*, 121(4), 839.
- Schnabl, G., Hassard, T. H., & Kopelov, M. L. (1991). The assessment of interpersonal skills using standardized patients. *Academia Medica*(66(1 suppl)), S34-S36.
- Taylor, M. J., Moynihan, E., McWilliam, J., & Gresty, D. (2004). Teaching business IT ethics. *Active learning in higher education*, 5(1), 43-55.
- Wielenga, D. K. (Ed.). (2002). *Probing and Proving Competence*: Amsterdam Faculty of Education.
- Wilcox, K. C., & Ryder, R. (2002). Standardized testing and improving educational opportunity in Brazil. *The Educational Forum [H.W. Wilson - EDUC]*, 66(3), 214.

9. Anexo. Lista de cotejo para establecer la calidad de la evaluación de los estudiantes (McInnis, 2004) ⁶

1. La evaluación es tratada por el personal y los estudiantes como componente integral del proceso global de enseñanza y aprendizaje.
2. Se reconocen los múltiples roles de la evaluación. Se comprende el poderoso efecto motivador de los requerimientos evaluativos sobre los estudiantes, y se diseñan tareas de aprendizaje para estimular hábitos de estudio considerados valiosos.
3. Existe una política de la facultad/departamento que guía las prácticas evaluativas. La evaluación de las materias está integrada a un plan general de evaluación de curso.
4. Existe una clara alineación entre los resultados esperados del aprendizaje, lo que se enseña y se aprende, y el conocimiento y destrezas evaluados.
5. Las tareas evaluativas evalúan la capacidad de analizar y sintetizar nueva información y conceptos más que simplemente recuperar información que ha sido presentada anteriormente.
6. Se utiliza una variedad de métodos de evaluación de manera que se minimizan las limitaciones de métodos particulares.
7. Se diseñan las tareas evaluativas para evaluar destrezas genéricas relevantes así como conocimientos y destrezas propios de la materia.
8. Existe una progresión constante en la complejidad y demandas de los requerimientos evaluativos en los años finales de las carreras.
9. En ciertas oportunidades existe una provisión para la elección de los estudiantes de tareas evaluativas y ponderaciones.
10. En la programación y diseño de las tareas evaluativas se consideran las cargas de trabajo de estudiantes y profesores.
11. Se evita una evaluación excesiva. Las tareas evaluativas se diseñan para muestrear el aprendizaje estudiantil.
12. Se ponderan las tareas evaluativas para equilibrar los roles evolutivo (“formativo”) y de juicio (“sumativo”) de la evaluación. Se utilizan niveles menores y menores ponderaciones al comienzo para proporcionar retroinformación a los estudiantes.
13. Las notas se calculan e informan sobre la base de resultados de aprendizaje claramente articulados y de criterios para los niveles de logro.
14. Los estudiantes reciben retroinformación explicativa y diagnóstica junto con las calificaciones.
15. Se verifican las tareas evaluativas para asegurarse que no existen sesgos inherentes que puedan desfavorecer a grupos particulares de estudiantes.
16. Se minimiza la deshonestidad a través de un cuidadoso diseño de la tarea, educación explícita y monitoreo apropiado de la honestidad académica.

⁶ La traducción de nuestra (Nota de los AA).