
Herramientas para la calidad del *e-learning* académico

PID_00274332

Patrizia Maria Margherita Ghislandi

Tiempo mínimo de dedicación recomendado: 5 horas





Patrizia Maria Margherita Ghislandi

Patrizia Ghislandi es profesora titular de «Tecnologías educativas y aprendizaje a distancia» en el Departamento de Psicología y Ciencias Cognitivas de la Universidad de Trento, Italia. Imparte clases de Metodología de la Investigación Cualitativa y Diseño e-Learning en el mismo departamento. Ha realizado investigaciones en el campo de las tecnologías y metodologías didácticas desde principios de la década de 1980, y ha publicado monografías y libros sobre el tema. Fundó y dirigió dos centros universitarios de tecnologías didácticas en la Universidad de Milán – Center for Learning Technologies (CTU), en 1976– y la Universidad de Trento –OnLine Didactics (DOL), en 2000–.

El encargo y la creación de este recurso de aprendizaje UOC han sido coordinados por la profesora: Juliana E. Raffaghelli

Primera edición: octubre 2020
© de esta edición, Fundació Universitat Oberta de Catalunya (FUOC)
Av. Tibidabo, 39-43, 08035 Barcelona
Autoría: Patrizia Maria Margherita Ghislandi
Producción: FUOC



Los textos e imágenes publicados en esta obra están sujetos –excepto que se indique lo contrario– a una licencia Creative Commons de tipo Reconocimiento-NoComercial-SinObraDerivada (BY-NC-ND) v.3.0. Se puede copiar, distribuir y transmitir la obra públicamente siempre que se cite el autor y la fuente (Fundació per a la Universitat Oberta de Catalunya), no se haga un uso comercial y ni obra derivada de la misma. La licencia completa se puede consultar en: <http://creativecommons.org/licenses/by-nc-nd/3.0/es/legalcode.es>

Índice

Introducción	5
1. Qué es la evaluación	9
2. Calidad	13
3. <i>Quality assurance</i> y <i>quality enhancement</i>	15
4. Calidad de la enseñanza (académica)	17
5. Los modelos de evaluación	22
6. <i>Ranking</i> de universidades	24
7. 7. Calidad del <i>e-learning</i>	28
8. Modelos para la calidad del <i>e-learning</i>	31
9. Herramientas para la calidad del <i>e-learning</i>	34
9.1. Herramientas utilizadas durante el <i>e-learning</i>	34
9.1.1. Entrevistas con los <i>stakeholders</i>	35
9.1.2. Prototipos	35
9.1.3. Revisión con el grupo de trabajo	36
9.1.4. Test de usabilidad con los usuarios	36
9.1.5. <i>Checklist</i>	37
9.1.6. Test de aceptación	37
9.2. Herramientas utilizadas después de la realización del <i>e-learning</i> ..	37
9.2.1. Entrevista	39
9.2.2. Grupo focal	39
9.2.3. Análisis de documentos	40
9.2.4. Observación	40
9.2.5. Cuestionario	42
9.3. Herramientas utilizadas durante o después de la realización del <i>e-learning</i>	43
9.3.1. Estándar	43
9.3.2. <i>Benchmark</i>	44
9.3.3. Buenas prácticas	46
9.3.4. Líneas guía	47
9.3.5. Rúbrica	47
10. Ejemplo de herramienta: la rúbrica adAstra	51

Abreviaturas	59
Bibliografía	60

Introducción

adAstra (*per aspera...*)

¿Por qué adAstra (*per aspera...*)?

Ad astra per aspera es un aforismo latino que dice que es posible llegar, en lo que se emprenda, a las estrellas, pero que alcanzarlas requiere mucho esfuerzo y comporta también, a veces, algún fracaso.

Es un aforismo que se encuentra citado, con varias declinaciones, en Cicerón, Virgilio, Séneca, Salustio y Horacio. Muchos y en muchos ámbitos lo han utilizado, desde la Royal Air Force inglesa hasta la hinchada *rossonera*, o Franco Battiato.

Es un aforismo tan utilizado porque la paronomasia le confiere musicalidad, eficacia e incluso un poco de misterio.

Para la autora y para su grupo de investigación, hace referencia a un proyecto de investigación con más de diez años a sus espaldas: el de las rúbricas adAstra, que son una herramienta para la evaluación de la calidad del *e-learning*. Pero de esto se hablará, tendidamente, en el texto que introducen estas líneas.

¿Qué exergo, pues, es mejor que *ad astra per aspera*, si el texto habla de calidad?

Feliz lectura.

En pocas palabras

El texto «Herramientas para la calidad del *e-learning* académico» quiere presentar, como el título permite intuir, una introducción a las herramientas de recopilación de datos para la evaluación de la calidad del *e-learning* académico.

Pero evaluar es emprender una actividad de investigación.

Y no se puede hablar de recopilación de datos para una evaluación, que también es investigación, si antes no se ha aclarado el enfoque teórico que está en el trasfondo de este estudio. Cualquiera que lleve a cabo una investigación sabe bien que la misma debe estar fundada sobre una sólida base epistemológica, sobre un enfoque teórico claro, sobre una metodología definida, sobre herramientas adecuadas para la metodología elegida (Crotty, 1998). Si no se aclara el contexto teórico de referencia, las herramientas pueden conducir a resultados engañosos, parciales, opacos. En definitiva, no se puede decir que

Sobre la autora

Patrizia Maria Margherita Ghislandi es profesora del Departamento de Psicología y Ciencia Cognitiva de la Universidad de Trento, Italia.

aporten algo fiable sobre la calidad del producto que se pretende analizar. Por un motivo simple: la investigación efectuada bajo estas premisas no se puede afirmar que sea de calidad.

Por todo lo dicho, el texto que sigue, antes de presentar las herramientas para la calidad afronta una serie de cuestiones que lo sitúan en la perspectiva teórica elegida. Así pues, habla de evaluación, de calidad, de aseguramiento de la calidad (*quality assurance*) y mejora de la calidad (*quality enhancement*), de modelos de evaluación, de calidad de la didáctica y del *e-learning*. Solo tras haber aclarado la posición de la autora respecto a estos elementos, se afronta el tema de las herramientas de recopilación de datos para una evaluación de la calidad correctamente contextualizada.

El módulo concluye con la presentación de un proyecto, con más de una década de trayectoria, de la autora y su grupo de investigación para la evaluación de la calidad del *e-learning*: la rúbrica adAstra.

Para este módulo, principalmente por motivos de espacio, se han tomado algunas decisiones:

- Se tratará, en particular, la evaluación del *e-learning* académico y no, por ejemplo, la evaluación del *e-learning* para la formación dual. Ambas evaluaciones tienen, sin embargo, varios puntos en común y lo que aquí se dirá también podrá ser utilizado, en muchos casos, para este tipo de formación.
- El *stakeholder* que se considerará de un modo preferente es el docente que evalúa el propio curso de *e-learning* que, normalmente, él mismo ha diseñado y erogado.
- Se tratarán, preferentemente, productos de *e-learning* indicados como «hechos a medida», es decir, productos realizados por un docente para su propia clase o curso, y para utilizarse en años sucesivos si procede. No obstante, en algunos casos este contenido también es válido para productos *off-the-shelf*, es decir, productos disponibles en el mercado que, eventualmente, el docente puede adaptar para su propio curso.

Introducción a la materia

El tema de este texto para la asignatura Construir Culturas de Calidad del *e-Learning*: Instrumentos y Prácticas de Evaluación son las herramientas de recopilación de datos para la evaluación de la calidad del *e-learning*. Pero antes de entrar en la materia de las herramientas, es útil hacer una reflexión más general sobre lo que se entiende por *evaluación*, por *calidad*, por *calidad de la didáctica* y por *calidad del e-learning*.

Estos son conceptos tan amplios y complejos, tan imbuidos de epistemología, es decir, de cuestiones fundamentales para la perspectiva desde la cual se observa un tema de estudio, que no pueden sino constituir la base para profundizar en este asunto. De hecho, no se pueden tratar cuestiones sobre la metodología del trabajo, como las herramientas, sin antes haber dirimido en qué contexto teórico se insieren.

Y es que cada método y cada metodología, como dice Crotty, deben derivar de una epistemología y una perspectiva teórica que se puedan declarar explícitamente (Crotty, 1998, págs. 2-9).

1. Qué es la evaluación

Entre los autores de referencia para entender lo que significa *evaluar* están, sin duda, Michael Scriven, filósofo australiano nacido en Gran Bretaña, y Michael Quinn Patton, expresidente de la American Evaluation Association.

El primero publica, ya en 1977, el volumen *Evaluation Thesaurus*, que en 1991 iba por la cuarta edición (Scriven, 1991), y el segundo publica, por primera vez en 1980, *Qualitative research & evaluation methods* (Patton, 1990, 2002). Más recientemente, Jane Davidson escribe el utilísimo *Evaluation Methodology Basics* (Davidson, 2005). Tomando como referencia a estos autores y autoras, se pueden extraer los conceptos fundamentales del constructo *evaluación*.

El término *evaluación*, para Scriven (1991, pág. vii), se refiere a:

«[El] proceso de determinación del mérito, el valor y la importancia de las cosas –o del resultado de ese proceso–».

La evaluación puede referirse a:

- La evaluación de un producto, por ejemplo, un artículo para una revista científica.
- La evaluación del personal, por ejemplo, docentes que realizan enseñanza académica.
- La evaluación de un programa educativo, por ejemplo, un curso de estudios universitarios.
- La evaluación de una propuesta, como la de un proyecto de investigación que participa en una convocatoria para su financiación.
- La evaluación del desempeño, ya sea deportivo o de estudiantes ante una prueba.

En este módulo, la atención se centrará en la **evaluación de un producto de *e-learning***, que es uno de los campos de aplicación de la disciplina más general de la evaluación.

La evaluación puede ser:

- **Heteroevaluación:** por heteroevaluación se entiende la evaluación realizada por un sujeto externo al proceso. En el campo del aprendizaje, por ejemplo, se trata de la evaluación por parte del docente. En general, por

heteroevaluación se entiende la evaluación dirigida por un evaluador externo (auditor) a los procesos, los productos o los sistemas que son objeto de evaluación. La heteroevaluación (auditoría) en este caso es realizada con fines sumativos, de certificación y notificación de las propias acciones –es decir, para dar cuenta de las acciones propias: *accountability*, en inglés–. Se trata de recopilar datos sobre la base de los cuales tomar decisiones.

- **Autoevaluación:** es la evaluación de un desempeño realizada por la misma persona que lo efectúa. La autoevaluación es una habilidad, compleja y a veces sufrida, que se adquiere gradualmente y que ayuda a la persona a mejorar el propio desempeño y los productos de la propia actividad, así como a construir progresivamente una imagen positiva de la propia identidad (Mariani, 2013):

«Autoevaluarse [...] comporta un juicio subjetivo, y precisamente por eso uno de los retos más difíciles consiste en el esfuerzo de asegurar también en la modalidad de autoevaluación aquellos niveles de calidad, cosa difícil de asegurar, que deberían ser comunes a todas las actividades de evaluación, es decir, validez, fiabilidad, transparencia e intercambio».

- **Peer evaluation o coevaluación:** introduce una dimensión de socialización de la experiencia de evaluación, en cuanto está basada sobre el hecho de que la evaluación de un desempeño está hecha por personas con el mismo nivel de competencia (por ejemplo, compañeros o compañeras de clase) y no por una persona con una competencia mayor (por ejemplo, el o la docente). Además, con el fin de que la evaluación sea más transparente y compartida, los parámetros de evaluación en la coevaluación son siempre explícitos.

La evaluación también se puede diferenciar entre:

- **Formativa:** la evaluación formativa se realiza, normalmente, durante un proceso (un aprendizaje, un programa, un desempeño deportivo, etc.) por personas que forman parte del equipo interno. Su objetivo es mejorar lo que se está evaluando. Los resultados de la evaluación siguen siendo internos y sobre todo útiles para los actores que forman parte del proceso. Los resultados no son, por lo tanto, utilizados para tomar decisiones por parte de personas externas.
- **Sumativa:** la evaluación sumativa se lleva a cabo al final de un proceso, de un programa o de un desempeño y normalmente a beneficio de personas externas al proceso, que deben tomar decisiones sobre el propio proceso (por ejemplo, erogar financiación, hacer «pasar» el examen de un curso, activar un programa de estudio para los y las estudiantes).

Robert Stake (2004) expone muy bien estos dos conceptos en una frase muchas veces citada:

«Cuando el cocinero prueba la sopa, es formativa; cuando el comensal prueba la sopa, es sumativa».

Todas las modalidades de evaluación mencionadas en principio no son alternativas ni compiten entre sí, sino que podrían –y de hecho deberían– integrarse.

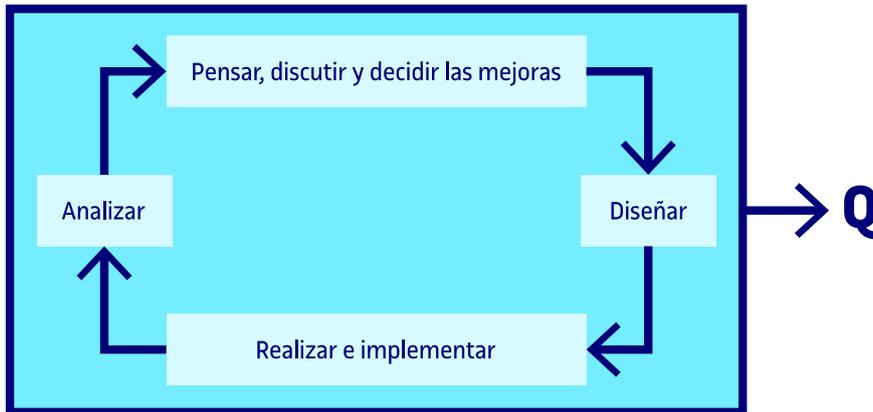
Lo dicho hasta aquí permite entender que la evaluación es un proceso complejo y con varias facetas, que debe realizarse teniendo en cuenta múltiples elementos. Muchos autores y autoras hablan de la evaluación como un proceso de investigación a todos los efectos (Patton, 1990, 2002; Guba y Lincoln, 1989; Shaw, 1999). En este documento se adopta la misma posición que estos estudiosos y estudiosas, es decir, se entiende que la evaluación debe seguir todos los pasos y el rigor de un proceso de investigación.

Por lo tanto, aquí se propondrán, entre las herramientas para la evaluación del *e-learning*, algunos instrumentos típicos de la investigación, preferentemente cualitativa. Hay que señalar que, en todo lo que guarda relación con la investigación evaluativa, un texto de óptimo para la consulta de los aspectos teóricos es el de Philips, McNaught y Kennedy (2012).

Como está indicado en el título de este módulo, aquí no se tratará, sin embargo, el análisis de los datos que se hayan recopilado con las herramientas que se proponen, ni las conclusiones que de este análisis se puedan extraer. No obstante, estas son operaciones indispensables en el ámbito del tipo de investigación en el que se opera con la evaluación que, en lo que respecta a la calidad de los cursos de *e-learning*, se propone que sea del tipo *Design Based Research*, o DBR (Anderson y Shattuck, 2012), el cual sigue un ciclo de cuatro fases (figura 1):

- 1) Proyectar.
- 2) Realizar e implementar.
- 3) Analizar.
- 4) Pensar, discutir y decidir las mejoras.

Efectivamente, la DBR se presta muy bien como metodología de investigación en el ámbito de la didáctica, con o sin el uso de tecnología.

Figura 1. Ciclo de evaluación de la calidad en DBR (*Design Based Research*)

Fuente: Anderson y Shattuck (2012).

2. Calidad

El concepto de calidad no nació en el mundo académico, sino más bien en el de la industria. Fue importado por la academia, a finales de los años ochenta, en el mundo anglosajón y más recientemente en los países latinos. Y en el mundo universitario, el constructo *calidad* se ha ampliado y profundizado. A continuación, los estudiosos de referencia ayudarán a entender por qué.

Entre los autores que se han ocupado de la calidad están, ante todo, Lee Harvey, exdirector del Center for Research and Evaluation de la Universidad Sheffield Hallam, y Diana Green, presidenta de la Society for Research into Higher Education, ambos británicos. Los dos estudiosos publicaron en 1993 un ensayo titulado *Defining Quality*, citado 2.759 veces según datos de Google Scholar (Harvey y Green, 1993). El trabajo, que analiza la calidad en relación con la educación universitaria, pero que en realidad examina las raíces políticas y filosóficas del concepto sometido a examen, dice que la calidad tiene distintos significados para diferentes personas. Al año siguiente, Diana Green (Green, 1994) publicó un libro en el que presentó cinco enfoques de la definición de calidad:

- Calidad como cumplimiento de un estándar (Green, 1994, pág. 13). La calidad de un producto se mide de acuerdo con el cumplimiento de un estándar, normalmente definido por una entidad acreditada para ello. Para decir si existe cumplimiento del estándar, se utilizan procesos de *benchmarking*, término utilizado para describir el proceso de evaluación de la validez de productos (Cocuzza, 2015-2020):

«[el] proceso de evaluación de la validez de productos, servicios y procesos mediante su comparación con las empresas con un mejor desempeño en el sector de referencia. Las comparaciones con los competidores directos o los operadores primarios de otros sectores se realizan precisamente sobre la base del *benchmark*, es decir, prestaciones de referencia respecto a las que realizar la comparación y hacia las que se debe apuntar al poner en práctica las acciones de cambio».

Green (1994, pág. 14) dice que este enfoque sobre la calidad tiene la ventaja de permitir a las distintas instituciones establecer los estándares que estimen más oportunos. Esto, sin embargo, hace difícil la medición, tan útil para la educación universitaria.

- Calidad como *fit for purpose* –ajustada al objetivo– cuando el producto o la práctica se ajustan al objetivo establecido. Sin embargo, como dice Green (1994, pág. 15), no es fácil definir exactamente cuál es el objetivo de la enseñanza académica. Por lo tanto, este enfoque es quizás más válido para el mundo de la industria y sus productos que para el mundo universitario.

- Calidad como eficacia en la consecución de los objetivos institucionales. Se trata de una subcategoría de la calidad *fit for purpose*, que es más adecuada para la evaluación de la calidad en el mundo académico, considerado de alta calidad y con una misión clara (Green, 1994, pág. 15). Gibbs (2011) también ha sugerido una visión alternativa al concepto de «ajustarse a un objetivo» con la idea de «práctica lo suficientemente buena».
- Calidad como satisfacción de las necesidades del usuario. Este modo de entender la calidad se centra en el conocimiento de los usuarios, de sus necesidades y en cómo satisfacerlas. Green (1994, pág. 16) se pregunta: «pero en el caso de la universidad, ¿el estudiante es el usuario, el producto o ambos? Y, además, ¿los estudiantes están capacitados para determinar completamente cuáles son sus propias necesidades?»
- Calidad como concepto tradicional. El concepto tradicional de calidad la define como excelencia. En términos prácticos, es una visión de calidad que clasifica las universidades de Oxford y Cambridge como universidades de excelencia, por los servicios que ofrecen a los estudiantes y por la competencia de sus graduados. Pero la calidad como excelencia no responde a la pregunta «¿Esto es bueno?» (como debería hacer), sino a la pregunta «¿Esto es mejor que lo de otros?». Lo que rápidamente conduce a un concepto relativo y de confrontación, de *ranking*, que no es la esencia de la calidad.

La calidad es, sin embargo, un concepto, o más bien una cultura con muchas perspectivas (Ghislandi y Raffaghelli, 2015; Ghislandi, Raffaghelli y Yang, 2013), en el sentido de que hay varias dimensiones que deben ser analizadas para entender totalmente todas sus facetas.

3. *Quality assurance* y *quality enhancement*

En un pasado no muy lejano, varios artículos analizaban la diferencia entre *quality assurance* –aseguramiento de la calidad– y *quality enhancement* –mejora de la calidad– (Raban, 2007).

Más tarde, en 2015, Noha Elassy (2015) escribe un importante ensayo en el que define los conceptos de *quality assurance* (QA) y *quality enhancement* (QE) en la enseñanza académica, y afirma que las dos prácticas deben formar parte de un continuo, siendo ambas necesarias para mejorar los procesos de enseñanza/aprendizaje universitarios. Otro importante estudio (Filippakou y Tapper, 2008) afirma que la QA juzga un producto o un proceso mediante ciertos criterios, mientras que la QE está menos definida y permite un discurso más complejo que deja espacio a una interpretación de la calidad más polifacética e integradora respecto a la innovación.

En la tabla siguiente (tabla 1), se enumeran a modo de ejemplo todas las diferencias que caracterizan la QA y la QE.

Tabla 1. Diferencia entre *quality assurance* y *quality enhancement*

QA	QE
Otorga un peso insuficiente a los procesos de enseñanza/aprendizaje.	Proporciona un peso considerable a los procesos de enseñanza/aprendizaje.
Tiende a asociarse más con la evaluación y la rendición de cuentas.	Tiende a asociarse más con la mejora y el desarrollo.
Se ajusta a estándares externos.	Se ajusta a estándares internos.
Se mueve del nivel más alto al más bajo.	Se mueve del nivel más bajo al más alto.
Un proceso sumativo.	Un proceso formativo.
Se centra en el pasado.	Se centra en el presente y el futuro.
Menos libertad (sigue reglas absolutas).	Más libertad (utiliza vías flexibles y negociadas).
Da un mayor espacio a administradores.	Da un mayor espacio a académicos.

Fuente: Elassy (2015, pág. 257).

Más recientemente, Ghislandi (2016a) afirma, en un texto en inglés traducido a continuación:

«Desde hace algún tiempo hemos asistido a un creciente interés de la universidad por la calidad. Interés que puede conducir a una politización y una distorsión de los resultados, dado que no basa las propias convicciones en estudios rigurosos y evidencias seguras. La universidad en estos días a menudo está dominada por una ideología gerencial y por una difundida retórica de la auditoría y la rendición de cuentas. Y a menudo, en la evaluación

de productos principalmente destinados al aprendizaje, también nosotros, pedagogos, aceptamos con demasiada pasividad el lenguaje propuesto por la industria.

Un lenguaje que está empapado por una visión positivista que trae consigo una hiperburocratización de todo el proceso, que termina preocupándose bastante más por el aparato organizativo que por el proceso de enseñanza/aprendizaje.

El influyente Alex Buckley (2014), dice: “Desde principios de los ochenta, el árbitro *de facto* de la calidad era una revista de noticias que clasificaba a escuelas y universidades en función de cosas como la reputación, las calificaciones en los exámenes de acceso de sus estudiantes y los salarios de los profesores. Estos *rankings* fueron diseñados para alimentar la creencia, entre el público general, que las instituciones más ricas y selectivas tenían que ser las mejores. Sin embargo, los que estaban verdaderamente preocupados por la calidad en la educación lamentaban la ausencia de la enseñanza y el aprendizaje en los discursos sobre calidad”. Estas también son las razones por las que debemos centrar nuestra atención, y diría que principalmente, en la mejora de la calidad por delante del aseguramiento de la calidad, con el objetivo de hacer más eficaz y válido el objeto o el proceso que estamos estudiando, a través de ciclos repetidos de evaluaciones y modificaciones [...]. En síntesis, la verdadera calidad es un proceso largo y complejo, que procede paso a paso, y no siempre en la dirección correcta [...]. Un proceso de certificación de la calidad –basado en una elección consciente e inteligente de condiciones previas epistemológicas, teóricas y metodológicas, así como en la identificación de los métodos correctos– es una condición *sine qua non* para una universidad que quiere garantizar una certificación ECTS (European Credit Transfer System) a los estudiantes que han completado un curso *online*».

4. Calidad de la enseñanza (académica)

Es útil, para empezar este párrafo, echar un primer vistazo a la definición de calidad de la enseñanza académica adoptado en este módulo, una definición que luego se enriquecerá de sentido en las páginas siguientes.

«La calidad de la enseñanza académica es una cultura compartida, mediada y transformadora que define (en un momento o lugar dados y para determinados *stakeholders*) la característica que la enseñanza debe tener, cómo evaluarla (criterios y métodos), y el proceso para tender a garantizarla y mejorarla».¹

⁽¹⁾Ghislandi, 2016b.

Y en esta definición la perspectiva transformadora, como se ha visto en el párrafo anterior, es el corazón de la *cultura de calidad para el aprendizaje*.

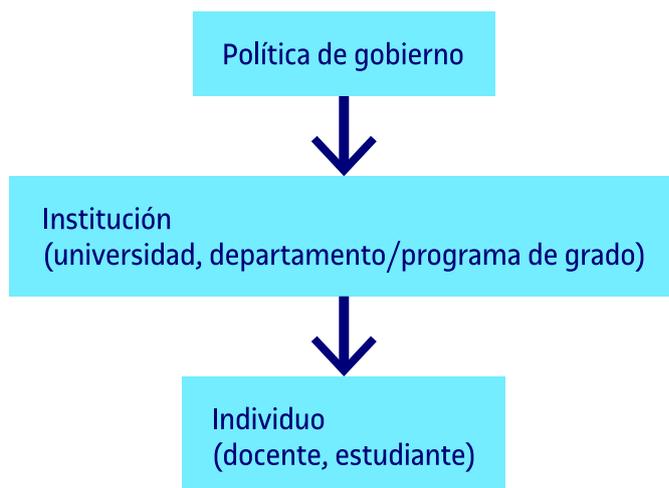
Partiendo del hecho de que el análisis de la calidad de la enseñanza académica se puede realizar a distintos niveles, la figura 2 distingue tres niveles de análisis: macro, meso, micro. El nivel macro o gubernamental es relativo a todo lo que se decide a nivel nacional o internacional, político y normativo, y a la erogación de los recursos pertinentes. El nivel meso se refiere a las instituciones (universidad o departamentos individuales) y es relativo (o debe referirse) a una visión y una cultura compartidas de la calidad, a ser posible teniendo en cuenta las posiciones de sus respectivas comunidades profesionales. Se ocupa, además, de la formación de los docentes. El nivel micro se refiere a la evaluación de la docencia y afecta a las motivaciones, los valores y la implicación personales, el desarrollo profesional, la implicación en la comunidad de práctica del docente individual.

Combinando los tiempos y los niveles de análisis macro y micro, se obtiene la matriz de la tabla 2, donde se identifican las combinaciones de las distintas posibilidades.

Tabla 2. Matriz de tiempo y nivel de análisis de la calidad de la enseñanza

Combinaciones de niveles y momentos de análisis de la calidad de la enseñanza		
Institución – <i>ex ante</i>	Institución – <i>in itinere</i>	Institución – <i>ex post</i>
Programa de grado – <i>ex ante</i>	Programa de grado – <i>in itinere</i>	Programa de grado – <i>ex post</i>
Docencia – <i>ex ante</i>	Docencia – <i>in itinere</i>	Docencia – <i>ex post</i>

Figura 2. Nivel de calidad de la didáctica de la enseñanza



Además de los niveles de análisis, hay otros constructos que componen el concepto de calidad de la enseñanza académica. En una convención en 2015, EM&M Italia, Multiconferenza Italiana su e-Learning, Media Education e Moodlemoot, celebrada en Génova del 9 al 11 de septiembre de 2015, la autora de este módulo presentó las diferentes dimensiones que conforman el constructo de calidad de la enseñanza en un octágono, mostrado en la figura 3 y revisada con profundidad para la presenta ocasión.

Figura 3. Octágono de la calidad del *e-learning* académico

Fuente: Ghislandi (2015), EM&M Italia.

Estas dimensiones involucradas son:

1. Epistemología/perspectiva teórica: qué se entiende por calidad.
2. Metodología de evaluación: cómo evaluar.
3. Los modelos de evaluación: con qué parámetros evaluar.
4. Tipo de producto a evaluar: listo para usar (*off-the-shelf*) / a medida.
5. Quién decide la evaluación: *stakeholders*.
6. Quién realiza la evaluación: evaluación externa / evaluación interna.
7. Tiempos/lugares de la evaluación: cuándo/dónde evaluar.
8. Herramientas de evaluación: con qué instrumentos evaluar.

Epistemología/perspectiva teórica: qué se entiende por calidad

Este elemento hace referencia a lo que se entiende por calidad y evaluación. Es un tema que ya se ha abordado muy brevemente en este módulo.

Para más información

Para los que deseen profundizar en la posición teórica en la que se basa este trabajo, la mejor referencia seguramente sea Guba y Lincoln (1989).

Metodología de evaluación: cómo evaluar

Con respecto a metodologías y métodos de análisis: se puede distinguir entre metodologías cuantitativas y cualitativas (Patton, 1990, 2002; Creswell, 2003). Acerca de los métodos (o herramientas): este es el tema principal del presente módulo. Sobre ellos se entrará en mayor detalle más adelante.

Los modelos de evaluación: con qué parámetros evaluar

Por las dimensiones del análisis: este elemento se basa en muchas investigaciones realizadas hasta el momento, y que identifican las dimensiones y los parámetros que deben ser considerados en la evaluación, en función de las evidencias probadas en numerosos estudios. A los modelos de evaluación se dedicará un párrafo a continuación, ya que son particularmente importantes.

Tipo de producto a evaluar: *off-the-shelf* (listo para usar) / a medida

Es posible que se requiera evaluar un producto existente, que se tenga, por ejemplo, que comprar o adquirir para realizar la formación. A este tipo de productos se le llamará productos *off-the-shelf* (listos para usar), es decir, aquellos que tomamos del estante, como en el supermercado, o del estante virtual de Internet. En otros casos, es posible que se tenga que evaluar un producto realizado por uno mismo o que se haya encargado a alguien, es decir, un producto a medida. Las herramientas necesarias para llevar a cabo los diversos tipos de evaluación pueden ser parcial o completamente distintos.

Quién decide la evaluación: *stakeholders*

¿Quiénes son los *stakeholders* de la enseñanza? Las prioridades, las expectativas y las necesidades cambian inevitablemente de una persona a otra en función del rol que desempeñe. Por lo tanto, es necesario considerar cuáles son los roles relevantes en la enseñanza académica, es decir, los *stakeholders* del proceso. A continuación se muestra alguna definición.

«*Stakeholder*: una persona con un interés especial en la compleción exitosa de un proyecto. Los *stakeholders* en el *e-learning* a menudo incluyen desarrolladores, facilitadores, alumnado, administración del alumnado, clientes, etc.»²

⁽²⁾ Learningcircuits.org (s. f.).

«Sujetos interesados en la actividad de una organización y en disposición, directa o indirectamente, de influir en sus resultados».³

⁽³⁾ INDIRE (s. f.).

Entre los *stakeholders*, pueden considerarse:

- La persona o los órganos de decisión (de una facultad o de una empresa) que deciden invertir recursos en formación *online*.
- Los usuarios del curso, aquellos que tienen que aprender.
- Las personas expertas en la materia o docentes que preparan los materiales.
- Las personas que diseñan el curso o docentes, encargadas de la posibilidad de interacción que ofrece el curso y de la estructura de los materiales.
- El tutor o la tutora, que facilita la interacción y la comprensión del contenido.
- El personal informático, que se ocupa de la plataforma en la que se realiza el curso.
- El servicio de asistencia técnica, que garantiza soporte técnico a todos los participantes.
- La institución que crea el contexto en el que se desarrolla el curso (expectativas generadas, leyes, etc.).

«Como se ha mencionado, es una lista ampliamente variable. De hecho, de un caso a otro, la responsabilidad propia de uno u otro de estos roles puede repartirse entre todos los demás en función del tamaño de la organización y de los recursos disponibles. Por lo tanto, la institución que quiera realizar un análisis de la calidad de sus propios cursos debe evaluar con qué figuras cuenta y cuáles son sus competencias. Y luego decidir si la asignación de esta responsabilidad es óptima o si debe ser modificada (tal vez añadiendo algún rol). Es importante señalar, analizando a los diferentes actores, que sus expectativas, así como sus responsabilidades, son diversas. Dado que en el objetivo general – la calidad del curso de aprendizaje– coexisten los distintos objetivos de cada *stakeholder* particular, deviene fundamental valorizar el rol de cada uno, definir sus límites, subrayar la importancia de la colaboración con otros *stakeholders* y de la negociación de los recursos disponibles. Por lo tanto, es útil que todos sean conscientes al detalle de cómo se dividen las competencias y la responsabilidad, con el fin de facilitar y hacer más eficaces la comunicación y la interacción».⁴

⁽⁴⁾Ghislandi, Raffaghelli y Yang (2013).

Quién realiza la evaluación: evaluación externa / evaluación interna

Si los *stakeholders* son las personas que tienen interés en evaluar la calidad de un producto, este puede ser analizado por la misma organización que lo ha realizado o utilizado, o por una organización externa.

Tiempos/lugares de la evaluación: cuándo/dónde evaluar

Acerca de los tiempos de evaluación: si el producto es a medida la evaluación puede ser *ex ante*, *in itinere* o *ex post*; es decir, durante el diseño, durante la realización o tras finalizar la realización del producto.

Si el producto es *off-the-shelf*, no puede ser *ex post*.

Herramientas de evaluación: con qué instrumentos evaluar

Los instrumentos de evaluación, en particular del *e-learning*, serán analizados a continuación, ya que conforman el tema principal de este módulo.

5. Los modelos de evaluación

Los parámetros que se tienen en cuenta en la evaluación pueden ser varios. Todos juntos constituyen un modelo de análisis, es decir, una lente a través de la cual la realidad es examinada, capaz de poner de relieve algunos elementos considerados más importantes para describir adecuadamente el objeto/proceso examinado. Es decir, un modelo puede ser definido como un esquema teórico que describe un objeto, definiendo las características consideradas más relevantes (Ghislandi, 2005, págs. 107-110). Mejor aún: un modelo es un esquema teórico que describe un objeto, definiendo sus categorías –bien definidas y con toda probabilidad mutuamente excluyentes– consideradas más relevantes. El modelo que se elija seguir en una investigación evaluativa depende de la intención investigadora que se declare al principio de la misma. Aquí se proponen algunos, muy conocidos en el mundo de la evaluación.

El primer modelo analizado es un estudio realizado para la enseñanza *online* pero que tiene valor, en las dimensiones especificadas, para la enseñanza tradicional (Barchechath, 1996):

«Hoy se entiende que el análisis de una estructura flexible para la formación a distancia debe tener en cuenta cuatro dimensiones: la pedagógica, la organizativa, la económica y la tecnológica. La posición jerárquica de estas cuatro dimensiones puede variar en función del punto de vista. En cualquier caso, constituyen “un sistema y no pueden ser consideradas separadamente [...]”».

Por ejemplo, para los y las docentes, seguramente la dimensión más significativa sería la pedagógica, pero no podrían siquiera empezar a diseñar su curso sin tener una idea clara de los recursos tecnológicos disponibles. Si se considerase oportuno invertir en nuevos recursos, seguramente se encontraría a alguien de la administración, más interesada en el aspecto económico, que con gusto desearía conocer cuáles son las motivaciones para ello. La dimensión organizativa, sin embargo, estaría siempre presente durante todo el proceso.

Otro modelo –siempre creado para la enseñanza *online* pero válido también para la didáctica tradicional– es el del OLC (Online Learning Consortium), hasta hace unos años llamado Sloan Consortium (Lorenzo y Moore, 2002).

El marco de referencia del OLC es el conocido como los «cinco pilares». Los cinco pilares identificados son:

- *Learning effectiveness* (eficacia del aprendizaje): la clave es la interacción con los y las compañeras, el cuerpo docente y los contenidos.

- *Student satisfaction* (satisfacción del alumnado): basado en la prestación de servicios oportunos y personalizados, y en una elevada calidad del aprendizaje obtenido.
- *Faculty satisfaction* (satisfacción del profesorado): basado en el apoyo moral y administrativo, y el respeto mutuo entre docentes *online* y profesorado tradicional. En este sentido, el artículo toma el ejemplo del Monroe Community College en Rochester, Nueva York, donde se pone en práctica lo que se ha llamado «modelo Monroe»: el profesorado recibe apoyo en todo lo necesario –formación, enseñanza, tecnologías, servicios a los estudiantes– para que destine el menor tiempo posible a cuestiones administrativas y, en su lugar, se centre en el diseño y la implementación de los cursos, y a la interacción con sus estudiantes.
- *Cost effectiveness* (control de los costes): se ocupa de cómo aumentar la eficacia didáctica, disminuir la tasa de abandono, superar el problema de los edificios saturados y disminuir los costes.
- *Access* (accesibilidad): el alumnado (con o sin discapacidad) debe poder descubrir fácilmente y por sí mismo cuán eficaz, satisfactorio y económicamente asumible puede ser el *e-learning* (o la enseñanza en general).

Por último, otro modelo muy conocido para la evaluación de la calidad de la formación, principalmente en el mundo del trabajo, es el de Kirkpatrick, mostrado en la tabla 3 (Kirkpatrick, 1998; Kirkpatrick y Kirkpatrick, 2016).

Tabla 3. Los cuatro niveles de Kirkpatrick: un modelo para la evaluación de la calidad de la formación

Nivel 1. Reacción	El grado en el que los participantes encuentran que la formación es favorable, motivadora y relevante para sus trabajos.
Nivel 2. Aprendizaje	El grado en el que los participantes adquieren el conocimiento, las habilidades, la actitud, la confianza y el compromiso previstos, en base a su implicación en la formación.
Nivel 3. Conducta	El grado en el que los participantes aplican lo aprendido durante la formación cuando regresan a sus puestos de trabajo.
Nivel 4. Resultados	El grado en el que los resultados previstos ocurren como resultado de la formación y el apoyo, y el conjunto de responsabilidades.

Fuente: Kirkpatrick (1998).

6. *Ranking* de universidades

Tras haber hablado de la evaluación de la enseñanza académica, se abre ahora un paréntesis de no poca importancia –visto el éxito de público que tiene actualmente– sobre los *rankings* de universidades. Cada año en agosto, cuando se acerca el período de elección de universidad por parte de los estudiantes, aparecen en prensa numerosos *rankings* de universidades a escala internacional. Entre los más conocidos están los listados en la tabla 4.

Tabla 4. Los *rankings* internacionales de universidades más conocidos

Rankings internacionales de universidades			
1	QS World Universities Ranking	Quacquarelli Symonds	https://www.topuniversities.com/university-rankings
2	ARWU (Academic Ranking of World Universities)	Shanghai Ranking Consultancy	http://www.shanghairanking.com/resources.html
3	THE World University Rankings	Times Higher Education	https://www.timeshighereducation.com/world-university-rankings
4	U multirank. Universities compared. Your way	European Commission	https://www.umultirank.org/

Para un análisis sobre los parámetros en los que se basan, se puede tomar como ejemplo el THE (Redazione ROARS, 2019), que construye su clasificación con estos porcentajes:

- 30 % enseñanza (el ambiente de aprendizaje).
- 30 % investigación (volumen, ingresos, reputación).
- 30 % citas (influencia de la investigación).
- 7,5 % influencia internacional (docentes, estudiantes, investigación).
- 2,5 % ingresos por la transferencia tecnológica.

Por lo tanto, la enseñanza influye en un 30 % sobre la clasificación general. Pero, ¿de dónde se obtiene este 30%?

- 15 % investigación sobre la reputación.
- 4,5 % relación docentes-estudiantes.
- 2,25 % porcentual de doctorados respecto a graduados trienales.
- 6 % docentes con doctorado.
- 2,25 % presupuesto de la institución (es decir, los recursos disponibles).

La inclusión de la «relación docentes-estudiantes» en los porcentajes de cálculo se basa en el hecho de que, normalmente, cuanto mayor sea el número de docentes mayor será el tiempo que dedican a los estudiantes. Y este elemento puede tener sentido para evaluar la calidad de la enseñanza, en tanto que la investigación ha demostrado que la implicación de los estudiantes es un indicador esencial del aprendizaje (Buckley, 2014). Los otros parámetros tienen poco que ver con los que, para la literatura científica, resultan fundamentales para una enseñanza de calidad.

Para orientar sobre los parámetros que son realmente importantes en la evaluación de la eficacia de la enseñanza académica, que es el elemento central de la calidad de la enseñanza, se puede recurrir a las reflexiones de Graham Gibbs. Gibbs, en 2010, publica un trabajo dedicada por entero a la calidad de la enseñanza académica (Gibbs, 2010). El ensayo se refiere a la calidad de la enseñanza en el Reino Unido en cursos equivalentes al grado de tres años italiano, y es una respuesta a los muchos *rankings* de universidades que se vienen proponiendo. Muchas de las tesis sostenidas por el autor han tenido validez también en contextos más amplios. Gibbs se propone como objetivo identificar una serie de dimensiones de la calidad de la enseñanza, verificando su validez en función de la investigación previamente realizada. Las variables consideradas son elegidas siguiendo el modelo de las «3 P» de Biggs (1993, citado en Gibbs, 2010): pronóstico, progreso, producto.

El **pronóstico** se refiere al contexto del estudiante fuera de la enseñanza, e implica que los mejores estudiantes se sienten atraídos por las universidades más prestigiosas, mejor financiadas y con mejores resultados en la investigación. Y las evaluaciones del alumnado al finalizar el curso están influenciadas por estos elementos. Por lo tanto, las evaluaciones finales del alumnado no pueden ser consideradas para evaluar la calidad de la enseñanza académica. Por la misma razón, la medición de la calidad del **producto** educativo, como el hecho de que cuanto se enseña sea recordado, o que el estudiante sea fácil y rápidamente empleable en el mundo laboral también están influenciadas por las variables de pronóstico. Todo esto hace más difícil, si no imposible, la comparación de las instituciones universitarias mediante las variables de pronóstico y de producto.

Son, en cambio, muy importantes para una comparación de la calidad educativa de las instituciones universitarias –es decir, para la evaluación de lo que el estudiante, independientemente de qué estudiante sea, aprende gracias al hecho de que la universidad a la que asiste aprovecha al máximo los propios recursos para proporcionar una enseñanza de calidad– las variables de **proceso**, es decir, lo que sucede en el proceso educativo, mientras el estudiante aprende y el docente enseña. Y las variables de proceso más importantes, puesto que la evidencia de la investigación sobre la influencia de estas variables está bien documentada, son las cuatro siguientes:

- La dimensión de la clase.

- Quién es el docente.
- El nivel de implicación del estudiante y su esfuerzo en el aprendizaje.
- La calidad y la cantidad del *feedback* recibido por el estudiante a lo largo de su aprendizaje.

Puesto que a menudo es difícil evaluar los parámetros del proceso, se puede, en un primer análisis, utilizar un indicador propuesto por Gibbs (2010):

«Uno de los indicadores más reveladores de la calidad de los resultados educativos es el trabajo que los estudiantes someten a evaluación, como su proyecto de último año o su disertación».

También hay que señalar que muchas veces es difícil comparar la enseñanza de toda una universidad, dado que la calidad de la enseñanza puede variar mucho entre los distintos departamentos. Por lo tanto, los *rankings* deben darse a un cierto nivel de granularidad, es decir, a nivel de departamento y no a nivel de institución para que sean útiles.

Como se ve, los parámetros que la investigación identifica como importantes para la calidad de la enseñanza son muy diferentes de los utilizados en los *rankings* más conocidos. Por lo tanto, difícilmente pueden ser considerados como fiables para una evaluación de la universidad, al menos en lo que respecta a la enseñanza.

A la luz de lo anterior, solo se puede estar de acuerdo con la afirmación de Pascarella (2001):

«Un [...] serio problema con el *ranking* de las revistas nacionales es que, desde el punto de vista de la investigación, son en gran parte inválidos. Es decir, se basan en los recursos institucionales y las dimensiones reputacionales, que, en el mejor de los casos, solo tienen una relevancia mínima para lo que sabemos sobre el impacto de la universidad en los estudiantes. Durante los últimos 15 años, mi colega Pat Terenzini y yo hemos pasado una cantidad considerable de tiempo leyendo y sintetizando literalmente miles de estudios sobre los factores que son importantes en la conformación del impacto de una universidad sobre sus estudiantes. Si me viera obligado a resumir esa investigación en una sola proposición, diría que las experiencias dentro de la universidad tienden a contar sustancialmente más que las características entre universidades».

También está de acuerdo con los numerosos artículos de crítica a los *rankings* la revista *online* italiana *Roars. Return on Academic Research*, que se define como una red «informal, con el objetivo de intervenir de un modo creíble y competente en un debate que tenga como interlocutores a los que deben gestionar el proceso de transformación de la universidad italiana y especialmente a las fuerzas políticas que son candidatas para gobernar el país en un futuro» (Redazione *ROARS*, 2015).

Se muestra, como ejemplo, una cita (THE, 2019):

«Lanzamos un reto a los académicos y periodistas italianos. Al día siguiente de la publicación del enésimo *ranking*, nos gustaría que algún profesor –aún mejor algún rector– de las universidades a la cabeza de la clasificación nacional tuvieran la honestidad intelectual de recordar los fallos, los límites y los peligros de estas clasificaciones, verdaderas armas de distracción masiva. Y también sería útil que algún periodista [...] recordase a los lectores las barbaridades y los resultados absurdos coleccionados a lo largo de los años de esta especie de horóscopos académicos».

Y no se podría estar más de acuerdo con De Nicolao (2015, 2018) y la redacción de *ROARS* (Redazione *ROARS*, 2015).

7. 7. Calidad del *e-learning*

Hasta aquí se ha hablado de la evaluación, de la calidad y de la calidad de la enseñanza académica. Pero en los años noventa en Europa se empieza a hablar de *e-learning* o enseñanza *online*.

Antes, sin embargo, hay que precisar qué se entiende aquí por *e-learning*.

Por *e-learning* se entienden «artefactos.... tecnología que pueden ser usados como herramientas para facilitar el aprendizaje» (Phillips, McNaught y Kennedy, 2012). Siempre diferenciando entre los artefactos las herramientas genéricas de *e-learning* y los objetos de aprendizaje. Para una descripción más detallada, se puede consultar Phillips, McNaught y Kennedy (2012).

La cuestión de la calidad en el *e-learning* se afronta en estudios de autores singulares o de autoría múltiple –como Creelman, Ehlers y Ossiannilsson (2014), o Wright (2003)–, y en muchos proyectos europeos y americanos. Se pueden citar algunos.

- EFQUEL (European Foundation for Quality in e-Learning),⁵ con el proyecto UNIQUE (EFQUEL, 2011).
- QM – Quality Matters Standards⁶ (Quality Matters, 2014).
- OLC – Online Learning Consortium⁷ (OLC, 2000).
- ICDE – International Council for Open and Distance Education⁸ (Mathes, 2016).

⁽⁵⁾<https://www.eurashe.eu/about/partners/efquel/>

⁽⁶⁾<https://www.qualitymatters.org/>

⁽⁷⁾<https://onlinelearningconsortium.org/>

⁽⁸⁾<https://www.icde.org/>

A continuación, se analizan brevemente las características de estos proyectos.

- La European Foundation for Quality in e-Learning (EFQUEL) es la promotora de la certificación UNIQUE. Esta certificación se centra en la autoevaluación, las revisiones externas y los planes de mejora. Todo está basado en 71 criterios de calidad. Los criterios utilizados se refieren más a la institución que al curso en particular.
- El programa QM – Quality Matters, ya desde sus inicios fue creado para los cursos *online*. QM realiza una evaluación de los cursos mediante una serie de rúbricas (Quality Matters, 2014) basadas en ocho estándares generales. Todo esto ha sido desarrollado en el marco del proyecto Quality Matters (2020b), patrocinado por MarylandOnline, el consorcio para la promoción de la excelencia en el *e-learning* que une escuela y universidad en el estado

⁽⁹⁾<http://marylandonline.org/>

de Maryland.⁹ Se verán con más detalle la rúbrica y el estándar de Quality Matters en el capítulo dedicado a las herramientas.

- El Online Learning Consortium, ya mencionado anteriormente, ofrece un sistema de setenta criterios que pueden ser utilizados tanto para cursos *online* como para los cursos MOOC-Massive Open Online Courses, un término acuñado en 2008 por Dave Cormier, gerente de Comunicación Web e Innovaciones en la Universidad de Prince Edward Island.
- El ICDE – International Council for Open and Distance Education, es una organización mundial sin ánimo de lucro que opera en el campo de la educación flexible y a distancia, en relación con la UNESCO, fundada en 1938 mediante una subvención del Gobierno de Noruega. El estudio sobre el estándar de la calidad en el *e-learning* ha sido llevado a cabo por un equipo de investigadores coordinado por la European Association of Distance Teaching Universities (EADTU), una organización europea líder en el campo de la enseñanza *online*. El objetivo del estudio era elaborar un análisis global sobre el estándar y las líneas guía para la enseñanza flexible, a distancia y *online* (Ossiannilsson, Williams, Camilleri y Brown, 2015; Mathes, 2016). *Quality on the Line. Benchmarks for success in Internet-based distance education*, de Jamie P. Merisotis y Ronald A. Phipps (2000) está disponible para su libre descarga en el sitio web del Institute for Higher Education Policy, donde aún forma parte de las «publicaciones más populares» (Merisotis y Phipps, 2000).

Así, se puede señalar (Ghislandi *et al.*, 2008b) que:

«Leyendo la muy citada Quality On The Line, la excelente rúbrica de Quality Matter y varios otros enfoques para la evaluación de la calidad, los factores importantes para alcanzar la excelencia son, con pocas diferencias, siempre los mismos (¡afortunadamente!). Las diferencias son fundamentalmente de dos tipos:

- Ámbitos que se excluyen o que son tratados con menor detalle. A algunos parámetros se les da una mayor o menor importancia en distintos enfoques (o son totalmente excluidos) según el nivel de análisis elegido por los autores: institución, curso de estudio o curso individual. Por ejemplo, si se considera el material informativo que la universidad pone a disposición de los futuros estudiantes, su calidad será considerada en la evaluación a nivel de institución, mientras que no será un factor relevante desde el punto de vista del tutor que se ocupa de un curso en particular.
- Agrupaciones ligeramente distintas de los mismos parámetros. Esto se debe a que las diferentes instituciones dividen las competencias y los procesos de manera diferente y, naturalmente, estas diferencias también tienen influencia en la organización de los métodos y las herramientas para la evaluación».

Hay que mencionar también algunos estudios que se han llevado a cabo específicamente para los MOOC y OER: Kawachi, 2015; Lowenthal y Hodges, 2015; EdX, 2014; Ehlers, Ossiannilsson y Creelman, 2013; Kocdar y Aydin, 2015; Lantero, 1996; Margaryan, Bianco y Littlejohn, 2015; Rosewell y Jansen, 2014; Williams *et al.*, 2011; Witthaus *et al.*, 2015; Yousef *et al.*, 2014.

Y también dos proyectos realizados para evaluar la calidad de los MOOC, aunque muchos de los parámetros que se proponen también se pueden utilizar para evaluar la calidad del *e-learning* en general:

- La iniciativa Openuped,¹⁰ financiada por la European Association of Distance Teaching Universities (EADTU) en colaboración con un importante número de universidades europeas, publicada en enero de 2014 (Creelman, Ehlers y Ossiannilsson, 2014).
- El estudio OpenCred, realizado por el Institute of Learning Innovation de la Universidad de Leicester, en colaboración el Institute for Prospective Technological Studies (IPTS) de la Comisión Europea (Witthaus *et al.*, 2015, págs. 1-12).

⁽¹⁰⁾<https://www.openuped.eu>

Todas estas iniciativas, y en particular las últimas citadas, también nacieron para contrarrestar el fenómeno de los *diploma mills* («fábricas de diplomas»), un término que designa organizaciones –particularmente activas en los Estados Unidos, pero también en otros países– que proporcionan diplomas en serie. En la mayor parte de los casos, estos diplomas no ofrecen una verificación precisa del aprendizaje.

8. Modelos para la calidad del *e-learning*

Antes de hablar de las herramientas de recopilación de datos, hace falta examinar los macroobjetivos de la evaluación.

Para los macroobjetivos se cuenta con la ayuda de las dos modalidades ya citadas anteriormente, la de Kirkpatrick (1998), y Kirkpatrick y Kirkpatrick (2016), y la del OLC – Online Learning Consortium (Moore, 2005).

Reinterpretando los dos modelos, unos más dedicado a la formación y el segundo más dedicado a la enseñanza, se obtiene una *checklist* de evaluación con diversos parámetros:

- **Satisfacción de los estudiantes y satisfacción de los docentes.** Los participantes (estudiantes y docentes) encuentran que la enseñanza y la formación son ventajosas, atractivas y relevantes para los objetivos de aprendizaje definidos.
- **Eficacia del aprendizaje.** Los participantes adquieren el conocimiento, la habilidad, las aptitudes y la dedicación deseados mediante la participación en la formación.
- **Aplicabilidad de lo aprendido.** Los participantes están en condición de aplicar cuanto han aprendido durante la formación.
- **Retorno de la inversión.** Los objetivos deseados se alcanzan como resultado conjunto de la formación, el apoyo y la responsabilización, pero teniendo presente también la inversión necesaria.
- **Accesibilidad.** El objeto para evaluar es accesible, es decir, puede ser consultado prescindiendo de cuál sea la cultura, la lengua, el tipo de medios tecnológicos a disposición y la habilidad mental o física del usuario (W3C Accessibility, 2018). La accesibilidad es la «capacidad de un sistema de brindar un acceso universal, esto es, de garantizar a todos sus usuarios la comprensión de los contenidos y la interacción con los otros elementos del sistema, prescindiendo de su cultura, de su idioma, del tipo de medios tecnológicos a su disposición y de su habilidad mental o física» (Schiavone, 2017).

A menudo, en lugar de recopilar datos para todos los cinco niveles, como sería el caso si realmente se desea certificar la calidad de un producto de *e-learning*, se recopilan datos para uno de los niveles.

La **verificación de la accesibilidad** es bastante lineal, en tanto que es un parámetro para el cual han sido establecidos unos estándares. No por ello se trata de un parámetro que siempre venga certificado, tal vez porque la cultura de la accesibilidad aún no ha penetrado en todos los ámbitos del diseño instruccivo, o porque en las primeras versiones del producto hay muchos otros elementos que deben funcionar y que son considerados prioritarios respecto a la accesibilidad, a menudo considerada, erróneamente, un «lujo» en lugar de ser la norma. Los estándares de accesibilidad han sido desarrollados por el World Wide Web Consortium (W3C), que es una comunidad internacional que desarrolla estándares y recomendaciones para la web. Las guías sobre accesibilidad constan de varias partes:

- Web Content Accessibility Guidelines (WCAG).
- Authoring Tool Accessibility Guidelines (ATAG).
- User Agent Accessibility Guidelines (UAAG).

Un parámetro, de los cinco, que habitualmente centra la atención es la **satisfacción de los estudiantes**, preguntándoles cuánto le ha gustado el producto, o si ha sido eficaz, y haciendo luego una recopilación y un análisis de los datos. Tal vez porque se cree que, de todos los parámetros, es aquel del que resulta más fácil recopilar información. Pero los estudiantes no están en condiciones de evaluar plenamente la calidad del producto del que han disfrutado, aunque puedan ofrecer indicaciones inestimables. Este tema se ha llamado, en la academia, SET – Students Evaluation of Teaching, y hay disponible sobre la cuestión una gran cantidad de literatura, desde los años ochenta hasta hoy. Quien escribe tiene en curso de publicación, como coautora, un artículo que ilustra el problema y avanza algunas sugerencias al respecto (Ghislandi *et al.*, 2020).

Con respecto a la **satisfacción de los docentes** en la universidad, a menudo no se recopilan los datos de este parámetro porque los docentes también son, con frecuencia, los autores del producto de *e-learning*, y difícilmente se toman en consideración los datos de una autoevaluación.

En lo que concierne al **retorno de la inversión**: no es objeto de este estudio realizar un análisis económico de los productos, pero es posible hallar un tratamiento profundo de este tema en Horton (2001, págs. 67-85).

Sobre la **aplicabilidad de lo aprendido**: es un parámetro muy tenido en cuenta especialmente en el mundo de la formación permanente; un poco menos, al menos para algunas materias, en el mundo académico. Se trata de verificar cuánto de lo aprendido es realmente aplicable a un problema auténtico. Un problema auténtico es una tarea que los estudiantes afrontan en una situación compleja, que está conectada con el plan de estudios y, en una relación significativa, con la realidad. Quien estudia, realiza y resuelve la tarea, demuestra la posesión y el ejercicio de conocimientos, competencias, recursos cognitivos y emocionales. Explora problemas interesantes y los analiza a la luz de lo que sabe. Busca respuestas y conocimiento adicionales. Hipotetiza un cambio.

Comparte sus resultados a través de un producto. Para ello, las estrategias didácticas a utilizar se basan en casos, problemas o proyectos concretos, según el campo de estudio. Se puede observar que en la academia este tipo de estrategias son poco practicadas, en tanto que aún la lección *ex cathedra* sigue siendo mayoritaria.

En cambio, en las empresas, probar la capacidad para resolver problemas concretos está a la orden del día, por lo que se pone de inmediato a prueba en qué medida los conocimientos, las habilidades y las competencias aprendidas han sido memorizados, y si son aprovechables y adecuados.

Por último, el parámetro quizás más difícil de investigar (aunque es también el más importante en el mundo académico) es cuánto ha influido el producto de *e-learning* en la calidad del aprendizaje, es decir, en la *eficacia del aprendizaje*. Esta es una cuestión difícil de estudiar no solo cuando se trata de productos de *e-learning*, sino en general cuando se quiere evaluar la calidad del proceso de enseñanza/aprendizaje, sea más o menos tecnológico. Aquí, también, la literatura científica ha producido una innumerable cantidad de materiales, de teorías, de modelos, de prácticas. De hecho, se puede decir que es uno de los temas más explorados en el mundo de la enseñanza. Sin embargo, aún hoy, no se ha alcanzado una teoría unificadora. Son demasiados los elementos que concurren en el proceso de aprendizaje; es demasiado compleja la rigurosa experimentación científica clásica de tipo positivista en un entorno «de campo» como es el de un aula. Se puede sugerir una publicación de un estudioso ampliamente conocido que afronta el problema de la calidad de la enseñanza (y que también sigue siendo válido para el *e-learning*) precisamente desde la perspectiva de la eficacia del aprendizaje, que ya ha sido previamente citado y del que se propone la lectura: Gibbs (2010).

9. Herramientas para la calidad del *e-learning*

Y, finalmente, ya se puede abordar el tema principal de este módulo, es decir: ¿cuáles son las herramientas para evaluar la calidad del *e-learning*, y cuáles son sus puntos fuertes y sus debilidades?

Cabe especificar que *herramienta*, aquí, hace referencia a «aquello que se necesita para alcanzar un objetivo» (De Mauro, 2009). En este caso, una herramienta es el medio que se utiliza para obtener informaciones, es decir, para recopilar datos sobre las características de un producto de *e-learning* con el fin de evaluar su calidad. Como ya se ha comentado en la introducción se hará referencia, también en este capítulo, a un producto de *e-learning* destinado a la enseñanza académica, aunque muchas de las herramientas que se presentarán se pueden utilizar también para un ámbito de evaluación más amplio.

Se pueden identificar tres grandes categorías de herramientas:

- Herramientas que se utilizan **durante** la realización de un *e-learning*, o herramientas para la evaluación *in itinere*.
- Herramientas que se utilizan **después** de un *e-learning*, o herramientas para la evaluación *ex post*.
- Herramientas que se utilizan **durante o después** de la realización de un *e-learning*.

9.1. Herramientas utilizadas durante el *e-learning*

En esta sección se tratarán las herramientas que normalmente se utilizan para productos *off-the-shelf* (listos para usar), es decir, elaborados por una empresa que tiene los recursos para dedicar espacio a la evaluación de sus productos, generalmente destinados a una distribución bastante amplia. Sin embargo, también pueden ser útiles para evaluar la calidad de un producto de *e-learning* «a medida», que a menudo es elaborado por el mismo docente titular del curso, con el objetivo de mejorarlo en los años académicos sucesivos.

Muchas de las herramientas que aquí se explorarán están muy relacionadas con el diseño de la experiencia del usuario (UX) (Nielsen, 1999), que las utiliza para obtener el conocimiento de los expertos en contenido, y de hecho las dos áreas (la de evaluación de productos *in itinere* y la obtención del conocimiento de expertos) a menudo son adyacentes o están superpuestas.

Se pueden enumerar las siguientes herramientas para la evaluación *in itinere* de un producto de *e-learning off-the-shelf*:

- Entrevistas con los *stakeholders*.
- Prototipos.
- Revisión con el grupo de trabajo o con expertos.
- Test de usabilidad con los usuarios.
- *Checklist*.
- Test de aceptación.

A continuación, se ofrece una breve descripción de cada uno.

9.1.1. Entrevistas con los *stakeholders*

Como dice Adam Fard (2019), un *stakeholder* es una persona que considera el éxito de un producto como beneficio personal o profesional, no solo económico, sino también en términos de prestigio. Puede ser una persona de alto nivel, pero también una persona que, pese a no ocupar posiciones elevadas en la escala jerárquica, igualmente utilizará con frecuencia el producto que se está elaborando. Estas personas pueden hacer sugerencias acerca del producto para simplificar y mejorar su diseño.

«El análisis de estas perspectivas te ayudará a entender mejor lo que los *stakeholders* quieren y necesitan del producto».¹¹

⁽¹¹⁾Fard (2019).

Por ello, la entrevista con los *stakeholders* tiene gran importancia.

9.1.2. Prototipos

Como todo el mundo puede imaginar, el prototipo es el primer modelo de un producto, ya sea una obra intelectual o un artefacto. Se utiliza para dar una idea, al cliente, del trabajo final, de lo que se hará, y también para obtener la financiación oportuna. También se utiliza para mostrarlo a todos los *stakeholders*, con el fin de recabar sugerencias que puedan mejorar o modificar lo que está haciendo (Wikipedia, 2020b). Este es el motivo por el que, con razón, se halla entre las herramientas más utilizadas para la evaluación de la calidad del *e-learning*, en fase de diseño o *in itinere*.

En arquitectura, en modelismo, etc., el prototipo se llama *maqueta*, del francés (*maquette*), o *modelo a escala*. Es una representación a escala bidimensional, o más a menudo tridimensional, del producto que se quiere realizar. Cada vez más, el modelo se realiza con gráficos por computadora o con inteligencia artificial.

El prototipo es, pues, la realización visible del *design thinking* (Brown, 2008).

9.1.3. Revisión con el grupo de trabajo

En la vida de un investigador, la revisión entre iguales o *peer review* es un método clásico de trabajo. Todos los estudiosos facilitan sus artículos a sus colegas mientras los están escribiendo, para intercambiar ideas o para una mejora recíproca. Muy a menudo sus estudios se someten, al terminar el trabajo, a estudiosos de la misma materia de las revistas a las que han sido enviados para tener un juicio sobre la calidad del artículo, con el fin de poderlo publicarlo evitando el plagio, las falsedades, las distorsiones o, peor aún, las estafas científicas (Brown, 2008).

Un método derivado de esto es el que se utiliza durante el diseño de un módulo de *e-learning*. Es decir, se presenta un prototipo de lo que se ha elaborado a los colegas del grupo de trabajo para que puedan dar consejos, plantear críticas o aprobar lo que se ha realizado.

9.1.4. Test de usabilidad con los usuarios

Stefano Bussolon (2020) afirma que:

«El test de usabilidad con los usuarios es una metodología empírica de evaluación que consiste en observar el comportamiento de las personas ajenas al proyecto de evaluación, mientras interactúan con la interfaz objeto de investigación».

Nielsen asegura que es inútil realizar más de cinco test con los usuarios porque, como demuestra en un interesante artículo (Nielsen, 2020), la curva del número de los problemas de usabilidad que se encuentran tiende a aplanarse completamente una vez superados los cinco usuarios.

Según Rubin y Chisnell (2008), se pueden distinguir cuatro tipos de test:

- Test explorativo. Finalidad: evaluar la eficacia del proyecto inicial, ayudar al grupo de trabajo en la elección de soluciones de diseño alternativas.
- Test de evaluación. Finalidad: profundizar en los aspectos evidenciados con el test explorativo.
- Test de validación. Finalidad: evaluar la usabilidad efectiva del producto, medida en términos de precisión, velocidad, preferencia.
- Test de comparación. Finalidad: comparar dos o más proyectos alternativos, o dos o más productos competidores. Se recopilan datos sobre el rendimiento y sobre la preferencia, y se comparan los resultados.

9.1.5. Checklist

Todo el mundo está familiarizado con las listas, ya que se utilizan a menudo para hacer la compra (¡y también para limitar la adquisición de cosas que no son realmente necesarias!). Una lista de verificación, o *checklist*, es cualquier lista exhaustiva de cosas para hacer o verificar para realizar una determinada actividad (Rubin y Chisnell, 2008). Surgió para ayudar a la memoria y la atención humanas, que si no se les ayuda, especialmente cuando el protocolo o proceso a seguir es complejo, tienden a olvidar algún elemento, o a distraerse con elementos extraños. Así pues, ayuda a realizar una tarea de un modo más completo, siguiendo todos los pasos que se consideran correctos para llegar a un resultado. La idea nació en el mundo de la aeronáutica, parece que en 1935, después de un accidente mortal causado por la imposibilidad de que los pilotos realizaran en la secuencia correcta todas las maniobras necesarias para pilotar un nuevo modelo de aeronave. Ahora la lista de verificación está muy extendida también en el mundo del diseño de calidad, sobre todo durante un proceso *in itinere*. Sin embargo, también se utiliza al final del proceso, para un control de la calidad *ex post*.

Algunos estudios sobre la *checklist* son los de Kim, Lee y Ryu (2013), y Lim y Lee (2007).

9.1.6. Test de aceptación

El test de aceptación por parte del cliente concluye el proyecto y determina que se han cumplido los objetivos establecidos. También se considera una herramienta para la calidad *in itinere*, aunque se podría incluir entre las herramientas para la calidad *ex post*.

Es un test que debe ser realizado por el cliente o comprador. Certifica que el producto encargado funciona, satisface las especificaciones del contrato y es de calidad. Es un test que tiene lugar inmediatamente antes de la entrega y, si es positivo, concluye con la satisfacción del pago final (Wikipedia, 2020c).

9.2. Herramientas utilizadas después de la realización del *e-learning*

Entre las herramientas utilizadas después del proceso de realización del *e-learning*, es decir, para una evaluación *ex post*, se pueden identificar las herramientas «clásicas» para la recopilación de datos, dispuestas por consolidadas metodologías de investigación.

La actividad de evaluación es, de hecho, una actividad de investigación, pero aplicada. Así, para evaluar la calidad del *e-learning* (y cualquier otro producto/proceso), es necesario, como en toda actividad de investigación, recopilar los datos que, una vez analizados, podrán ofrecer información sobre lo que se está evaluando. En este trabajo, la atención se centrará en herramientas pa-

ra la recopilación de datos, y no en herramientas y métodos para el análisis de datos. ¿Qué datos se deben recopilar? Aquellos que sugieren las hipótesis (investigación cuantitativa) y las preguntas de la investigación (investigación cualitativa) que indican los parámetros del modelo para la calidad que se ha adoptado. Sobre qué parámetros, entre los propuestos por el modelo, llevar a cabo la investigación será una decisión que tomará el responsable de la investigación.

No es posible aquí hacer un tratado sobre metodología de la investigación, por lo que solo se ofrecerán indicaciones básicas. Sin embargo, es muy aconsejable el texto de Creswell (2003), donde se analizan de un modo ejemplar diversas metodologías: cuantitativa, cualitativa y mixta. Y también un sitio (SAGE, 2020) que intenta apoyar a la investigación facilitando materiales para guiar a los usuarios a través de todos los elementos del proceso de investigación, proporcionando un diccionario, algunos estudios de caso, y la posibilidad de consultar conjuntos de datos o títulos dedicados a los distintos tipos de investigación.

Las herramientas más importantes que ofrece el mundo de la investigación y que se pueden utilizar en la recopilación de datos para analizar la calidad de un producto de *e-learning*, son:

- Entrevista.
- Grupo focal.
- Análisis de documentos.
- Observación.
- Cuestionario.

En la investigación cuantitativa, donde se consolidan protocolos bastante detallados, para recopilar los datos en primer lugar hay que identificar una muestra. La muestra se da mediante la selección de un subconjunto de casos (esto es, un número equis de casos entre todos los casos posibles del entorno de referencia) que sea de algún modo representativo del universo (el conjunto de todos los casos posibles).

En la investigación cualitativa, en lugar de seleccionar un muestreo basado en conceptos, se seleccionan los comportamientos, las actitudes o las situaciones que son relevantes para la pregunta que se investiga, y se entrevista a los individuos que asumen esos comportamientos/actitudes o que se encuentran en las situaciones de referencia. Las muestras serán necesariamente de dimensiones reducidas, estudiadas de un modo intensivo, dado que todo el análisis de datos se realiza manualmente. Es decir, no es el ordenador el que extrae las estadísticas, sino que es el investigador quien analiza los significados que sugieren los textos recopilados. A menudo, el muestreo no se define completamente al inicio de la recopilación, sino que se especifica progresivamente sobre la base de los resultados de los análisis preliminares de los datos recopilados.

lados. Finalmente, las investigaciones cualitativas no ofrecen generalizaciones estadísticas, sino más bien ideas detalladas sobre la calidad de un contexto, de un proceso, de un producto.

Las herramientas para la recopilación de datos que se han enumerado anteriormente se pueden utilizar tanto en la investigación cuantitativa como en la cualitativa, aunque algunas de ellas son más comunes en el paradigma cuantitativo y otras en el paradigma cualitativo.

A continuación, se analizarán las herramientas de recopilación de datos antes enumeradas.

9.2.1. Entrevista

La entrevista es una conversación (profesional) entre dos personas dirigida a un objetivo. Puede ser directa o mediada por algún sistema telemático o multimedia. Puede ser estructurada, semiestructurada o en profundidad. En el primer caso, en todas las entrevistas se hacen las mismas preguntas y en el mismo orden, para garantizar la máxima comparabilidad entre los datos recogidos. En el segundo caso, la entrevista semiestructurada, las preguntas son las mismas, pero se pueden hacer en un orden distinto. Y también se hace alguna concesión a la improvisación de preguntas. La entrevista en profundidad es un coloquio libre entre el entrevistador y el entrevistado, aunque esté conducida por el primero.

Para más información

Para los que quieran profundizar en el tema de las entrevistas, se recomienda consultar «SAGE Research Methods» (SAGE, 2020).

9.2.2. Grupo focal

El grupo focal es un método basado en la discusión entre un pequeño grupo de personas, en presencia de un moderador, centrada en un tema para explorar en profundidad (Brinkmann y Kvale, 2018). Las discusiones se planifican cuidadosamente y se llevan a cabo en un ambiente permisivo y no amenazante, incluso con una investigación de opiniones y actitudes relativas a temas sensibles, observando también la dinámica de grupo durante la discusión, con el fin de entender los procesos de construcción de la realidad social.

Las ventajas del grupo focal respecto a la entrevista son:

- La posibilidad, para los participantes, de formarse una opinión incluso durante la discusión, o modificar la que se expresa inicialmente.
- Los bajos costes.
- La rapidez y la eficiencia de la recogida de datos, así como su riqueza.

El grupo focal también tiene algunas desventajas:

- Dificultad para crear condiciones de homogeneidad en la recopilación de datos.
- Necesidad de un moderador experto, no demasiado directivo.
- Dificultad para reunir sujetos y hacerlos coincidir en horarios, lugares, etc.
- Los datos producidos reflejan la naturaleza del proceso conversacional (discusiones interrumpidas, cambios de tema, contradicciones, malentendidos).
- Es difícil recoger el punto de vista de las personas menos asertivas, especialmente si algunos participantes dominan la discusión.

9.2.3. Análisis de documentos

Uno de los métodos casi exclusivamente cualitativos para recopilar datos es pedir, a aquellos que disfrutan del curso de *e-learning*, que expresen su opinión sobre el propio curso. La documentación que elaboren los estudiantes puede ser completamente libre, o puede seguir una serie de cuestiones a considerar, más o menos rigurosamente.

También se pueden hallar modos de garantizar que la opinión de quien escribe sea anónima, con el fin de no influir en la expresión de su opinión.

9.2.4. Observación

La observación es una descripción, intencional y precisa, de lo que sucede a las personas en una determinada situación (Richards y Morse, 2009, pág. 149). La observación es el acto de percibir la actividad y las relaciones de las personas sobre el terreno, a través de los cinco sentidos del investigador (Angrosino, 2007, pág. 37). La **observación** es el acto de adquirir información de una fuente primaria, en seres vivos, obtenida a través de los órganos de los sentidos. También puede implicar el registro de datos mediante el uso de herramientas (Wikipedia, 2020a).

Figura 4. Las dimensiones de la observación

Grado de implicación	Nivel de conciencia	Grado de sistematicidad	Contexto
<ul style="list-style-type: none"> • Observador completo • Participante completo 	<ul style="list-style-type: none"> • Cubierta • Descubierta 	<ul style="list-style-type: none"> • Estructurada • Libre 	<ul style="list-style-type: none"> • Artificial • Natural
Enfoque	Notas/vídeo	Duración	Feedback a los participantes
<ul style="list-style-type: none"> • Estrecho • Amplio 	<ul style="list-style-type: none"> • Notas • En mente 	<ul style="list-style-type: none"> • Larga • Corta/urgente 	<ul style="list-style-type: none"> • Puesta en común de las observaciones • Sin <i>feedback</i>

Fuente: adaptado de Cicognani (2002).

Implicación y participación del observador

El grado de participación puede variar con el tiempo, desde la inmersión completa en el contexto hasta el total distanciamiento. Puede haber:

- Participante completo.
- Participante como observador.
- Observador como participante.
- Observador completo.

Observación cubierta/descubierta

Se refiere a los actores, es decir, a la percepción y el conocimiento que tienen de la presencia del investigador. Puede ser:

- Cubierta (problemas éticos): los actores desconocen la identidad del investigador y los objetivos de la investigación.
- Semicubierta: la mayoría de los involucrados desconocen la investigación. Algunos miembros del grupo están al corriente de la identidad del investigador y los objetivos de la investigación.
- Descubierta (peligro de distorsión): es la técnica más difundida y requiere un gran esfuerzo para obtener el consentimiento por parte de los responsables, y para ganarse la confianza de los informadores.

Sistematicidad y estructuración de los procedimientos

Puede ser:

- Observación sistemática (esquema de observación estandarizado).
- Observación no sistemática (observación libre y abierta).

Contexto

Puede ser:

- Situación natural (complejidad del fenómeno, dificultad de aislar los elementos más significativos).
- Situación artificial (riesgo de desnaturalizar el significado del comportamiento observado).

Enfoque de la observación

Ejemplo:

- Enfoque amplio (incluye todos los aspectos del contexto).
- Enfoque reducido (limitado solo a un cierto comportamiento).

Notas

Se pueden tomar apuntes en video o con notas. O se pueden tomar notas mentales de la entrevista, cuando el tema es sensible, aunque haya un cierto riesgo de olvidar algunos elementos.

Escala temporal o duración de la observación

Se pueden realizar observaciones instantáneas o prolongadas (meses, años).

Además, el número de las observaciones es naturalmente variable en función de los recursos y del contexto.

Explicaciones facilitadas a los participantes y *feedback* posterior

Puede haber:

- Completa compartición de las observaciones.
- Ausencia de contactos ulteriores con los participantes.

9.2.5. Cuestionario

El cuestionario es una herramienta utilizada normalmente en la investigación cuantitativa.

Puede estar compuesto por preguntas de respuesta cerrada, o puede incluir también preguntas de respuesta abierta.

Para más información

Para quien desee profundizar en el tema de la observación, se recomienda el estudio de Corrao (2005).

Para más información

Quien desee profundizar en el tema de los cuestionarios, puede consultar el texto de Cellini (2008).

9.3. Herramientas utilizadas durante o después de la realización del *e-learning*

En la categoría de las herramientas que pueden ser utilizadas en el proceso de evaluación de la calidad, tanto durante como después de la realización del *e-learning*, se incluyen en particular aquellas que sugieren algunas buenas prácticas, es decir, que ayudan a traducir a la práctica, durante o después de la realización del *e-learning*, las indicaciones teóricas proporcionadas por los modelos adoptados.

Las herramientas que sugieren buenas prácticas de *e-learning* y que considerarán aquí, son:

- Estándar.
- *Benchmark*.
- Buenas prácticas.
- Líneas guía.
- Rúbrica.

Las herramientas de la lista pueden ser utilizados:

- a efectos de garantía de calidad, para la evaluación del *e-learning* a nivel macro o de sistema (figura 2); o
- a efectos de mejora de la calidad, para la evaluación del curso individual o de un producto o proceso.

También hay que decir que, si bien las herramientas clásicas de investigación están muy bien definidas porque se han utilizado durante muchos años y por muchas personas, los límites entre las herramientas mencionadas anteriormente en realidad son muy relativos. A menudo el *benchmark* es denominado *estándar*, por ejemplo. O se asocian a las buenas prácticas. Así pues, la subdivisión que se presenta aquí es, principalmente, con fines didácticos.

A continuación se describen las herramientas una por una y se citan los ejemplos más conocidos.

9.3.1. Estándar

Es una característica –o un conjunto de características– predefinida, conocida y aceptada, o dada por sentada, de una determinada categoría de objetos o procesos. Su finalidad es armonizar o formalizar productos, servicios o procesos, y simplificar los procedimientos, es decir, hacerlos fácilmente comprensibles para todos aquellos que deben seguirlos. También pretende que los productos de *e-learning* sean reutilizables e interoperables independientemente de quién los haya creado y de quién deba utilizarlos. Un estándar puede presentarse ba-

jo la forma de reglas, directrices, especificaciones o normas, y debe indicar cuál es su objetivo, en qué contextos puede ser aplicado, qué perspectiva adopta y en qué modo se debe evaluar su aplicación.

Un estándar es una herramienta rígida y prescriptiva. Si se establece que algo debe hacerse según un estándar determinado, lo ideal es que no haya dos maneras distintas de hacerlo y que los resultados finales sean muy parecidos. Los estándares se utilizan normalmente durante un proceso o para la evaluación de un producto.

En lo que respecta a la calidad de un curso *online*, los estándares son necesarios en un nivel de análisis detallado. Por ejemplo, cuando hay que ponerse de acuerdo sobre el formato que debe tener un texto wiki colaborativo, para que todos los participantes puedan contribuir con sencillez y claridad. O cuando hay que decidir el tipo de formato a utilizar para preparar los materiales, como en el caso del estándar SCORM –*Shareable Content Object Reference Model* (modelo de referencia para los objetos de contenido compartible–, que es una colección de estándares y especificaciones que hacen posible la interoperabilidad, la accesibilidad y la reutilización de contenido para el aprendizaje basado en la web (De Vaus, 2013; Philip, 2000).

Existen dos tipos de estándar (Philip, 2000):

- *De iure*.
- *De facto*.

Los estándares *de iure* son promovidos por una organización responsable de ratificarlos y aprobarlos. Los estándares *de facto* son adoptados por una gran parte del mundo de la industria, que constituye gradualmente una masa crítica que ratifica la adopción del propio estándar. Estos estándares pueden convertirse en *de iure* si las organizaciones pertinentes los adoptan y certifican.

9.3.2. **Benchmark**

El *benchmark*, o análisis competitivo, consiste en analizar y evaluar productos concurrentes de *e-learning*, para identificar las soluciones funcionales y de contenido adoptadas por los competidores, en términos de características generales, estructura y funcionalidad.

De Wikipedia (Bartleson, 2012):

«Por *benchmark* [...] en economía se entiende una metodología basada en la comparación sistemática que permite a las empresas que la aplican compararse con las mejores y sobre todo aprender de ellas para mejorar. [...] Se remonta a finales de los años setenta, cuando Robert Camp, en nombre de la Xerox Corporation, decide mejorar radicalmente las prestaciones de la función logística mediante una comparación de productos y procesos operativos de las empresas con más autoridad presentes en el mismo sector de mercado».

El *benchmarking* es, en esencia, una herramienta de mejora organizativa basada en el análisis de la competencia y, por lo tanto, la comparación. La idea básica es encontrar una referencia externa para evaluar la calidad y la relación coste-eficacia de la propia actividad y los propios procesos. Una vez que se establece cuál es el proceso, el producto o el servicio a mejorar, se toman cada una de sus fases y se compara, aspecto por aspecto, con el mismo proceso, producto o servicio ofrecido por los competidores considerados líderes en el sector. Además de las empresas del mismo sector, se pueden realizar *benchmarkings* entre empresas de sectores diferentes pero con procesos similares, ya sea dentro de la misma empresa, entre distintos departamentos o en momentos distintos.

«El *benchmarking* se realiza eligiendo un conjunto de indicadores clave, cada uno con nombre y métrica, que sean objetivos, comprensibles, poco costosos de medir, representativos de un proceso crítico. Una unidad de medida común (porcentaje, metro, número puro...) para todas las empresas comparadas en el *benchmarking* asegura la conmensurabilidad de los resultados».¹²

(12)Wikipedia (2019).

La empresa puede elegir si estudiar el mejor elemento en todas sus prestaciones o el mejor en su clase, es decir, el elemento que tenga un mayor número de prestaciones con la máxima puntuación. El mejor producto en su clase es elegido entre empresas del mismo sector. Como se puede adivinar por lo que se acaba de decir, el *benchmarking* nació en el ámbito económico.

Se proponen aquí tres ejemplos de *benchmarking* para el *e-learning*:

- BENVIC (Benchmarking of Virtual Campuses).
- Quality on the Line («Benchmarks for success in Internet-based distance education»).
- Excellence in e-Learning: un *benchmark* para el *e-learning* propuesto por la EADTU (European Association of Distance Teaching Universities).

El primer ejemplo de *benchmarking* para el *e-learning* es el proyecto BENVIC (2002). BENVIC es un proyecto con el objetivo de trasladar el *benchmarking* al ámbito del *e-learning*. Nacido bajo el patrocinio de la UE, se desarrolló entre los años 1999 y 2002 (la última actualización del sitio del proyecto se remonta a febrero de 2002) y ha conducido a la construcción de una herramienta capaz de tener en cuenta el tipo de organización que se está evaluando, de modo que pueda compararse con instituciones similares en fines, tamaño, estructura, etc. Los ciento dos criterios identificados se dividen en ocho categorías:

- *Learner Services*.
- *Learning Delivery*.
- *Learning Development*.
- *Teaching Capability*.
- *Evaluation*.
- *Accessibility*.

- *Technical Capability*.
- *Institutional Capability*.

La primera autoevaluación se realiza mediante una *checklist*, donde se da un valor a cada indicador:

- 0 = no implementado.
- 1 = parcialmente implementado.
- 2 = totalmente implementado para indicadores de estructura o práctica, como porcentaje para los indicadores de desempeño.

Esto permite al equipo de BENVIC realizar, posteriormente, un análisis comparativo que será la base sobre la que trabajar para obtener mejoras. Se puede encontrar información más detallada en el documento «Evaluation Methodology Report», disponible en el sitio web de BENVIC (2002).

El segundo ejemplo facilitado aquí es Quality on the Line.

El documento «Benchmarks for success in Internet-based distance education», de Jamie Merisotis y Ronald Phipps, del año 2000, se puede descargar gratuitamente desde el sitio web del Institute for Higher Education Policy (Merisotis y Phipps, 2000).

Y, finalmente, el proyecto «Excellence in e-Learning, a European Quality Benchmark» realizado por la EADTU (European Association of Distance Teaching Universities) está bien descrito en numerosos documentos disponibles (IHEP, 2020; EADTU, 2020; Williams *et al.*, 2011; Rosewell y Jansen, 2014).

Existen también algunos estudios sobre *benchmark*, como los propuestos aquí: Kear, Williams y Rosewell (2012), Ossiannilsson y Landgren (2012), Choy (2007).

9.3.3. Buenas prácticas

Por *buenas prácticas* se entienden, en general, las experiencias más significativas o con mejores resultados en distintos contextos. Una buena práctica es una técnica, un método, una actividad, un proceso, un incentivo, un modo de hacer algo que es más eficaz que cualquier otro para lograr un cierto resultado. La idea es que, con los procesos, los análisis y los controles adecuados, se pueda lograr un resultado con menos problemas y complicaciones. Las buenas prácticas también se pueden definir como el modo más eficiente (menor esfuerzo) y más eficaz (mejores resultados) de realizar una tarea, basado en procedimientos repetibles que se han demostrado válidos y útiles con el tiempo y para un gran número de personas o empresas. Las buenas prácticas (Sae-Khow, 2014; Buzzetto-More y Alade, 2006) se pueden definir como una colección de ejemplos debidamente formalizados en reglas.

9.3.4. Líneas guía

Las líneas guía son cualquier documento que tiene como objetivo hacer más eficientes los procesos particulares, de acuerdo con una rutina dada. Consisten en una lista de reglas o principios que derivan de la experiencia, o de la evaluación de los resultados y de las mejores prácticas. Por definición, seguir las líneas guía no es estrictamente obligatorio (como sucede con los estándares) (Edmonds, 2006).

Ejemplos de líneas guía que están fácilmente disponibles en la web, pueden ser los trabajos de Bussolon (2020), Ryan y Brown (2012), Stracke (2009), De-Rouin, Fritzsche y Salas (2004), y Nilsson y Palmér (2002).

9.3.5. Rúbrica

El término *rúbrica* proviene del latín *rubrica* y se refería a la tierra roja (*ruber*) utilizada para teñir la tira de pergamino (índice) que normalmente colgaba del rollo de papiro, y que contenía el nombre del autor y el título del texto. Además, la tierra roja también se utilizaba para escribir las primeras letras de un texto, para resaltar el inicio, o incluso el prospecto (sumario) que, al principio o al final del volumen, reunía y reúne los títulos de los capítulos individuales. Este uso se mantuvo en los manuscritos medievales y luego en obras impresas (Instituto Treccani, 2020).

El término *rúbrica* ha sido, recientemente, prestado a las áreas evaluativas, con la intención de evidenciar los principales parámetros en los que se basa una evaluación.

«La rúbrica es un documento que articula las expectativas para una tarea enumerando los criterios, o aquello que recuenta, y describiendo niveles de calidad de excelente a pobre. Las rúbricas pueden tanto enseñar como evaluar.

⁽¹³⁾Goodrich (1996).

Cuando se usan como parte de un enfoque de evaluación formativa y centrada en el estudiante, las rúbricas tienen el potencial de ayudar a los estudiantes a desarrollar comprensión y habilidad, así como a emitir juicios fiables sobre la calidad de su propio trabajo. Los estudiantes deben poder usar las rúbricas de muchos de los mismos modos en que las usan los docentes –para aclarar los estándares para el desempeño de la calidad, y para guiar un *feedback* continuado sobre el progreso hacia esos estándares–.¹³

La rúbrica se utiliza particularmente en el ámbito escolar para evaluar la consecución de los objetivos de aprendizaje por parte de los estudiantes. Estos parámetros normalmente se comparten, por lo que una rúbrica permite (por ejemplo, en el sector universitario) tanto a profesores como a estudiantes evaluar del mismo modo todo criterio complejo o subjetivo. Por lo tanto, una rúbrica (o una cuadrícula de evaluación) es una herramienta que tiene como objetivo hacer lo más objetiva posible la evaluación del objeto al que se aplica. También puede proporcionar una base para la autoevaluación, la reflexión y la revisión entre iguales.

Todo esto lleva a quienes utilizan las rúbricas a reflexionar sobre las prácticas y las metodologías de formación, y también proporciona una visión general de la naturaleza del proceso analizado.

El uso de rúbricas, normalmente creadas para medir un objetivo determinado (de desempeño, comportamiento o calidad) se está extendiendo, ya que una rúbrica es a menudo más útil que *checklists* y los juicios basados exclusivamente en una impresión general sobre el objeto a juzgar (Beetham y Sharpe, 2007).

Según Bernie Dodge y Nancy Pickett (2007), y Joan L. Herman (1992, pág. 40), se pueden distinguir un cierto número de características comunes en las rúbricas, en particular:

- Una o más dimensiones mediante las cuales se mide el desempeño.
- Definiciones o ejemplos en base a los que se mide el desempeño.
- Una escala de evaluación para cada dimensión.
- Modelos y ejemplos que acompañan a cada nivel de desempeño.

Normalmente, las rúbricas se presentan en forma de tabla. Realizando una búsqueda en Google se pueden encontrar muchos ejemplos (Google, 2020).

Una de las rúbricas para el *e-learning* más conocidas es la propuesta por Quality Matters (QM). Las rúbricas QM han sido desarrolladas como parte del proyecto Quality Matters (2020a), patrocinado por MarylandOnline, el consorcio para la promoción de la excelencia en el *e-learning* que une escuela y universidad en el estado de Maryland (Quality Matters, 2020b; MarylandOnline, 1999).

Las rúbricas QM, por el modo en el que están estructuradas, son más bien un punto medio entre la rúbrica y la *checklist*, aunque a veces son llamadas estándares (Quality Matters, 2020a). Como ya se ha tenido ocasión de comentar anteriormente, los límites entre estas herramientas no siempre están bien definidos. Más allá de los términos utilizados, el conjunto de rúbricas QM consiste en una serie de cuarenta elementos subdivididos en ocho áreas basadas en la literatura científica, y en una investigación que ha involucrado a docentes y otros *stakeholders* que se ocupan del *e-learning*. Las ocho áreas enumeran las características que debe tener cada una de las partes de un formulario o producto de *e-learning*. A continuación se ofrece su descripción, recuperada del sitio web de QM.

- Introducción al curso. Da las instrucciones que normalmente se proporcionan el primer día de curso en el aula: requisitos mínimos, cómo navegar por el curso, introducción y política.
- Objetivos de aprendizaje. Determina cómo los objetivos de los módulos y del curso pueden ser algo medible.

- Evaluación. Indica cómo será realmente medida la consecución de los objetivos y facilita los criterios mediante los cuales los estudiantes recibirán las calificaciones.
- Materiales didácticos. Especifica el modo en el que el material didáctico garantiza a los estudiantes la posibilidad de alcanzar los objetivos de aprendizaje.
- Actividad e interacción entre los estudiantes. Determina si la actividad prevista garantiza que los estudiantes alcancen los objetivos didácticos. También si la actividad y el proyecto del docente promueven un aprendizaje activo.
- Tecnología del curso. Se ocupa de cómo la tecnología del curso apoya los objetivos de aprendizaje, y promueve el aprendizaje activo y la participación.
- Apoyo a los estudiantes. Se ocupa de cuatro tipos distintos de apoyo ofrecido durante el curso: técnico, a la accesibilidad, al estudio y a los estudiantes (sobre ayudas financieras y de asesoramiento).
- Accesibilidad y usabilidad. Determina si la navegación y las características del curso facilitan el aprendizaje.

Los cuarenta ítems de las ocho áreas se revisan año tras año sobre la base del *feedback* proporcionado por quien utiliza rúbricas QM para la evaluación del curso. De este modo las rúbricas QM, que aparecieron en 2004, mejoran constantemente.

Otras rúbricas de utilidad:

- California State University - Chico Rubrics for Online Instruction (McAleese, 2013).
- iNACOL Standards and Rubrics to Measure Quality e-Learning Course Design (California State University Chico, 2003; Aurora Institute, 2020). El Aurora Institute, anteriormente iNACOL, es una organización sin ánimo de lucro con sede en Viena, una ciudad del condado de Fairfax, Virginia, Estados Unidos. Tiene la misión de guiar la transformación del sistema educativo, y de acelerar el avance de políticas y prácticas innovadoras para garantizar un aprendizaje de alta calidad para todos.

También es interesante señalar el trabajo de Yuan y Recker (iNACOL, 2011), disponible en su web, que presenta un análisis de catorce rúbricas para la evaluación de los OER (Open Educational Resources).

Cabe señalar que la falta de escalas de clasificación, guías de clasificación, pruebas empíricas y revisiones iterativas plantea muchas dudas sobre la fiabilidad y la validez de todas estas rúbricas. Hay que decir que su usabilidad real también depende del apoyo que se asegure a los usuarios.

Puede ser útil concluir el resumen de estas herramientas con una comparativa (tabla 5) que confronte *checklist*, líneas guía y rúbrica en función de algunos parámetros (Yuan y Recker, 2015).

Tabla 5. Comparativa entre las características de la *checklist*, las líneas guía y la rúbrica

Característica	Checklist	Líneas guía	Rúbrica
Lista de los parámetros a monitorizar	√	√	√
Descripción de los parámetros a monitorizar	X	√	√
Espacio para comentarios por parte del usuario	X	X	√
Peso de los elementos a monitorizar	X	X	√
Espacio para ejemplos y demostraciones	X	√	√

Fuente: Yuan y Recker (2015).

10. Ejemplo de herramienta: la rúbrica adAstra

De las rúbricas se ha hablado un poco más arriba. Pero aquí se quiere presentar un proyecto de rúbrica llevado a cabo en la Universidad de Trento y que se ha estado desarrollando desde hace muchos años. Si se quisiera atender rigurosamente a la tabla presentada con anterioridad, sería impropio utilizar el término *rúbrica* para el proyecto que aquí se propone. Pero como se ha dicho, los límites entre estas herramientas de evaluación son algo difusos y, por lo tanto, se seguirá haciendo uso de este término para el proyecto adAstra, como se hecho en muchas otras publicaciones. Además, se han incluido las rúbricas entre los productos que, en general, se suelen utilizar *ex post* para la evaluación final de un producto de *e-learning*. La rúbrica adAstra, por otro lado, fue creada principalmente para ayudar a los docentes *durante* el diseño de cursos de *e-learning*, además de al concluir el trabajo. Así pues, a continuación, se verán sus principales características.

Las rúbricas adAstra –cuyo proyecto tiene su inicio en 2007– han sido llevadas a cabo por el grupo de investigación dirigido por la autora de este texto. Todas las rúbricas adAstra están disponibles en Zenodo.¹⁴ Zenodo es la plataforma de acceso abierto de la Unión Europea, lanzada en mayo de 2013. Toma el nombre de Zenodotus, el primer bibliotecario de la mítica biblioteca de Alejandría y el primero en usar «metadatos» *ante litteram*, piedra angular de la evolución de las bibliotecas.

⁽¹⁴⁾<https://zenodo.org/communities/adastra/>

Han participado en el proyecto adAstra muchos estudiosos, becarios, docentes e investigadores. Sobre el proyecto, se han publicado numerosos artículos, muchos de los cuales están disponibles en Zenodo (CERN Data Centre *et al.*, 2020; Ghislandi, 2012, 2014; Ghislandi y Cumer, 2011, 2012, 2013; Ghislandi y Pedroni, 2009a, 2009b, 2011; (Ghislandi *et al.*, 2008a; Ghislandi y Raffaghe-lli, 2012a, 2012c; Ghislandi, Raffaghe-lli y Cumer, 2012).

El objetivo de la rúbrica adAstra es proporcionar a los docentes que tengan la intención de diseñar y realizar un *e-learning*, una herramienta que les ayude a lo largo de todo el proceso, es decir, que les permita disponer de una guía durante el diseño, la erogación y el uso de un producto de *e-learning*, y que les ayude en la evaluación de la calidad del producto realizado. La idea de fondo de adAstra es la conciencia de que los docentes no siempre tienen competencias de diseñadores instruccionales, y las rúbricas tienen precisamente el objetivo de ayudar a los docentes a desarrollar esta competencia.

Los supuestos teóricos del proyecto adAstra son los siguientes:

- La facilitación de los procesos de autoevaluación y evaluación entre iguales mediante herramientas en línea mejora la *quality literacy* o alfabetización

de calidad (Ehlers, 2007). Las indicaciones no son suficientes para garantizar la calidad. Hay que desarrollar prácticas y habilidades que orienten toda la institución hacia una cultura *formativa de calidad*.

- *Minimally Invasive Education*, o educación mínimamente invasiva (Mitra, 2015): los sujetos están capacitados para tomar decisiones adecuadas para la autoformación, solo se requiere el entorno adecuado.
- Calidad mediada (Ghislandi y Raffaghelli, 2012b): la cultura formativa de calidad *está apoyada (mediada) por herramientas adecuadas* que permiten la reflexión, la negociación, la colaboración entre actores.
- Para mejorar la enseñanza, se debe adoptar principalmente una actitud *de mejora de la calidad*, así como la más tradicional de *aseguramiento de la calidad*, un constructo procedente del mundo de la industria.

Las características peculiares de adAstra son las siguientes:

- Apoyar el diseño, la erogación y el uso de una enseñanza *e-learning ex ante, in itinere y ex post*: mientras la casi totalidad de las rúbricas prevé que se lleve a cabo una monitorización al final del curso (*ex post*), adAstra permite la evaluación como intervención dirigida a lo que acaece *durante* el desarrollo del curso, desde el mismo momento del diseño.
- Mejorar la calidad del *e-learning* a través de un correcto diseño.
- Definir las opciones principales que están en la base del diseño y la realización de una enseñanza *e-learning*, en particular de tipo constructivista.
- Apoyar a docentes y a diseñadores instruccionales. Los docentes que son expertos en diseño instruccional tienen el apoyo de las dimensiones y los parámetros de las rúbricas, mientras que los diseñadores instruccionales cuentan con una cómoda *checklist* de control.

En el futuro, las rúbricas se ampliarán para apoyar las diversas estrategias y situaciones educativas que tienen un enfoque pedagógico participativo e innovador.

La *suite* adAstra está actualmente compuesta por tres rúbricas:

- Diseño y realización.
- Erogación.
- Satisfacción.

La rúbrica de diseño y rúbrica de erogación son de autoevaluación (el docente se evalúa a sí mismo durante el trabajo de construcción de un curso *online*). La rúbrica de satisfacción es una rúbrica de heteroevaluación, es decir, el estudiante evalúa el producto realizado por el enseñante.

Cada rúbrica está compuesta por siete dimensiones, que a su vez agrupan un cierto número de parámetros:

- Organización.
- Materiales didácticos.
- Didáctica.
- Comunicación.
- Colaboración.
- Tecnología.
- Verificación del aprendizaje.

El proyecto adAstra nace teniendo presentes los principales modelos para la calidad del *e-learning*, de los que ya se ha hablado, como por ejemplo el de OLC, o de los «cinco pilares» (eficacia del aprendizaje, satisfacción de los estudiantes, satisfacción del docente, control de los costes, accesibilidad).

Posteriormente, las rúbricas adAstra se han utilizado y validado en muchos proyectos:

- Seis ediciones del curso TIA –Tecnología de la Educación y del Aprendizaje (por sus siglas en italiano, Tecnologie dell’Istruzione e dell’Apprendimento)–, en la Universidad de Trento (desde el curso académico 2003-2004, hasta el 2008-2009).
- Entrevistas semiestructuradas con los diseñadores senior de Dol-Didattica OnLine, de la Universidad de Trento (aproximadamente, en el curso académico de 2012).
- El curso ELFO (formación para docentes de *e-learning*), un curso organizado por la Provincia de Trento y la Universidad de Trento para preparar diseñadores de *e-learning* (curso académico 2011-2012).
- El proyecto FIRB, con seis tutores (2009-2012).
- Cuatro ediciones del curso eLD-eLearning Design de la Universidad de Trento (2016-2017, 2017-2018, 2018-2019, 2019-2020).

Las rúbricas adAstra se utilizan normalmente por separado. A continuación, se reporta la rúbrica adAstra en formato comparativo, con el fin de apreciar algunas características adicionales de la propia rúbrica.

La principal es que las preguntas de cada rúbrica individual son similares a las de las otras rúbricas. Es decir, las preguntas, en la medida de lo posible, se readaptan para poder comparar las respuestas de las distintas rúbricas relativas a un mismo parámetro.

Esta elección se hizo para que fuera posible, mediante un análisis estadístico de los datos recabados con la compilación de rúbricas, una comparativa entre la autoevaluación de la rúbrica de erogación y la heteroevaluación de la rúbrica de satisfacción.

A continuación se muestran algunas figuras.

Tabla 6. Rúbrica adAstra: ejemplo de dimensiones (en el ejemplo: la enseñanza) y sus parámetros

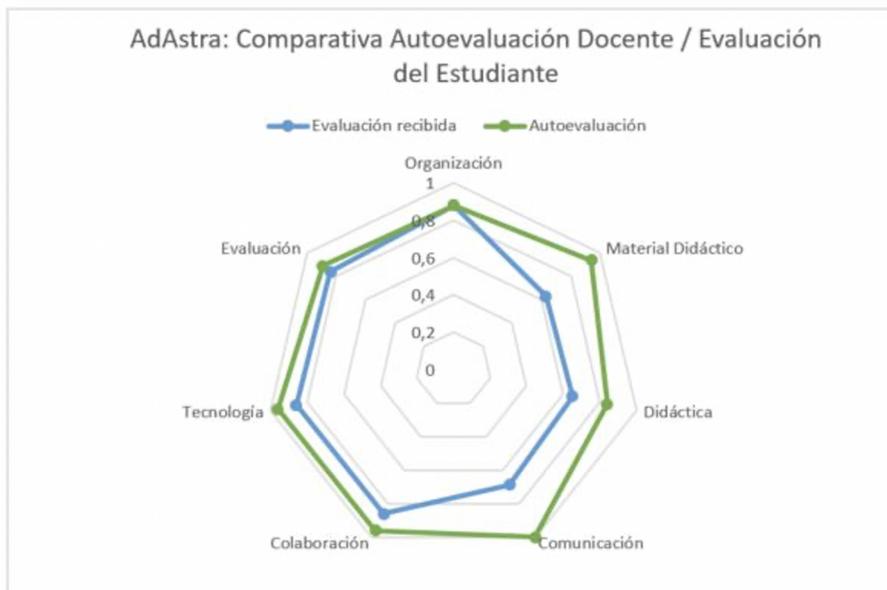
Didáctica					
Diseño instruccional		Implementación docente		Participación	
DR/d01	En el <i>syllabus</i> /plan docente están descritas de un modo claro las características y las actividades del curso.	I/d01	En el <i>syllabus</i> /plan docente están descritas de un modo claro las características y las actividades del curso.	P/d01	En el <i>syllabus</i> /plan del curso están descritas de un modo claro las características y las actividades del curso.
DR/d02	Los objetivos del curso son adecuados a los contenidos, estrategias y modalidades de evaluación del aprendizaje.	I/d02	Los estudiantes pueden expresar dudas y preguntas respecto a los objetivos del curso.	P/d02	Como estudiante, sientes que has alcanzado los objetivos descritos en el <i>syllabus</i> /plan del curso.
DR/d03	Se especifican todos los prerrequisitos necesarios.	I/d03	Los estudiantes pueden expresar dudas y preguntas respecto a los prerrequisitos del curso.	P/d03	Los prerrequisitos declarados en el curso son apropiados respecto a tus conocimientos iniciales.
DR/d04	Las actividades y tareas pedidas a los participantes en el curso están relacionadas con los objetivos.	I/d04	Las actividades y tareas pedidas a los participantes en el curso están relacionadas con los objetivos.	P/d04	Las actividades y tareas que te han pedido están relacionadas con los objetivos.
DR/d05	Están claramente definidas las características, tiempos y modalidades de entrega de las tareas.	I/d05	Están claramente definidas las características, tiempos y modalidades de entrega de las tareas.	P/d05	Te ha resultado clara la definición de características, tiempos y modalidades de entrega de las tareas.
DR/d06	Las actividades <i>online</i> y eventuales actividades presenciales están bien compensadas en relación con los objetivos y los contenidos del curso.	I/d06	Las actividades <i>online</i> y eventuales actividades presenciales están bien compensadas en relación con los objetivos y los contenidos del curso.	P/d06	Según tu experiencia, las actividades <i>online</i> y eventuales actividades presenciales están bien compensadas en relación con los objetivos y los contenidos del curso.

Fuente: tabla no publicada de Stefania Cucchiara, Patrizia Ghislandi y Juliana Raffaghelli.

La tabla 6 muestra una comparativa entre las tres rúbricas sobre una de las siete dimensiones de las propias rúbricas, en este caso la enseñanza, y todos los parámetros relativos a la misma.

Figura 5. Rúbrica adAstra: ejemplo de restitución a los usuarios de la elaboración de los datos de la rúbrica que han compilado

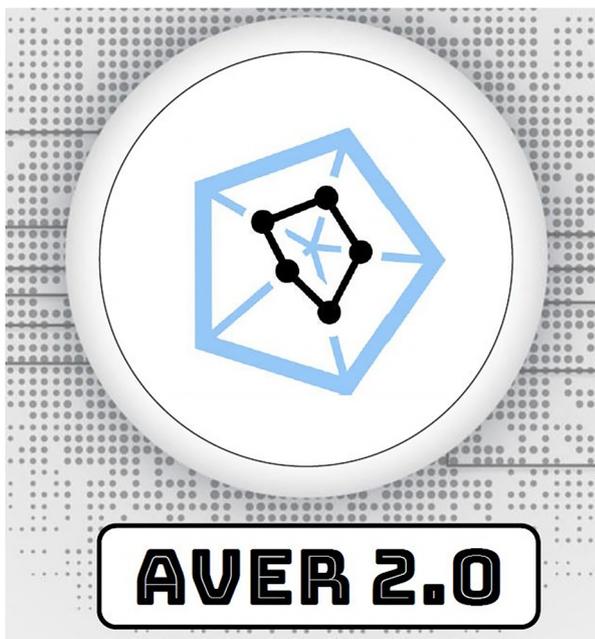
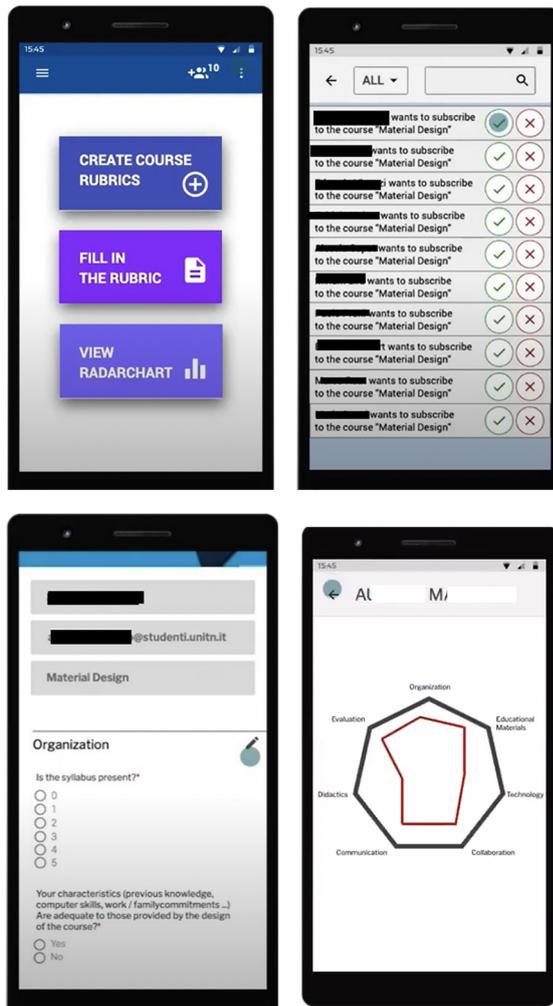
	Organización	Material Didáctico	Didáctica	Comunicación	Colaboración	Tecnología	Evaluación
Evaluación recibida	0,88	0,63	0,65	0,69	0,86	0,86	0,84
Autoevaluación	0,88	0,94	0,84	1	0,96	0,96	0,89



Fuente: imagen no publicada de Stefania Cucchiara, Patrizia Ghislandi y Juliana Raffaghelli.

La figura 5 muestra un ejemplo de retorno a los usuarios del tratamiento de los datos de las rúbricas que han sido compilados. Se puede observar, en la tabla, que los valores de los parámetros individuales de las distintas dimensiones se han promediado y normalizado a 1. También se puede observar, en misma la tabla, que se realiza una comparación entre la rúbrica de erogación, compilada como autoevaluación por parte de quien ha realizado el proyecto sobre el que se está examinando la calidad, y la rúbrica de satisfacción, compilada como heteroevaluación por aquellos que han disfrutado del mismo curso *e-learning*.

Figura 6. Imágenes de *mock-up* de la App for Visual Learning Rubric (AVER; i. e., rúbrica adAstra para móvil)



Fuente: imágenes no publicadas de Aurora Magnago y Aurela Pjeci. Proyecto coordinado por Patrizia Ghislandi y Juliana Raffaghelli.

Por último, en la figura 6 hay un ejemplo de un primer boceto de la interfaz de la aplicación para móvil que transportará la rúbrica adAstra a dispositivos móviles, o lo que es lo mismo: «App for visual eLearning Rubric (AVER)».

En adAstra han contribuido muchas personas, listadas en Zenodo. En particular, hay que destacar la importante contribución en la concepción de la comparativa de datos entre las distintas rúbricas de Juliana Raffaghelli, de la Universitat Oberta de Catalunya, y de Stefania Cucchiara, investigadora independiente y colaboradora de la Universidad de Trento. También hay que destacar, para el proyecto AVER, además de la contribución de Juliana Raffaghelli, la de algunos becarios de la Universidad de Trento, en particular: Aurora Magnago y Aurela Pjeci (2020), Simone Tagliente (2019) y Michele Chistá (2017).

Para más información

Para quien desee profundizar aún más en el funcionamiento y las características de la rúbrica adAstra, se recomienda consultar el sitio Zenodo dedicado a estos asuntos, donde varios autores han publicado sus informes de trabajo.

Abreviaturas

- ADL** Advanced Distributed Learning.
- ARWU** Academic Ranking of World Universities.
- ATAG** Authoring Tool Accessibility Guidelines.
- AVER** App for Visual e-Learning Rubric.
- BENVIC** Benchmarking of Virtual Campuses.
- COI** Committee for Online Instruction.
- DOL** Didattica OnLine.
- DBR** Design Based Research.
- EADTU** European Association of Distance Teaching Universities.
- ECTS** European Credit Transfer System.
- EFQUEL** European Foundation for Quality in e-Learning.
- FIRB** Fondo per gli Investimenti della Ricerca di Base.
- ICDE** International Council for Open and Distance Education.
- IHEP** Institute for Higher Education Policy.
- INACOL** International Association for K-12 Online Learning.
- INDIRE** Istituto Nazionale Documentazione Innovazione Ricerca Educativa.
- IPTS** Institute for Prospective Technological Studies .
- MOOC** Massive Open Online Course.
- OER** Open Educational Resource.
- OLC** Online Learning Consortium.
- QA** Quality Assurance.
- QE** Quality Enhancement.
- QM** Quality Matters.
- QS** Quacquarelli Symonds.
- SCORM** Shareable Content Object Reference Model.
- THE** Times Higher Education.
- TIA** Tecnologie dell'Istruzione e dell'Apprendimento.
- UAAG** User Agent Accessibility Guidelines.
- UNESCO** United Nations Educational, Scientific and Cultural Organization.
- W3C** World Wide Web Consortium.
- WCAG** Web Content Accessibility Guidelines.

Bibliografía

Anderson, T.; Shattuck, J. (2012). Design-Based Research: A Decade of Progress in Education Research? *Educational Researcher* (vol. 41, n.º 1, págs. 16-25). doi: <https://doi.org/10.3102/0013189X11428813>

Angrosino, M. (2007). *Doing ethnographic and observational research*. Thousand Oaks, CA: Sage Publications.

Aurora Institute (2020). *iNACOL* [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.inacol.org/>

Barchechath, E. (1996). «La progettazione dei sistemi formativi a distanza dal punto di vista economico, pedagogico e organizzativo». En: M. A. Garito (Ed.). *La multimedialità nell'insegnamento a distanza*. Roma: Garamond.

Bartleson, K. (2012). *What's The Difference Between De Jure And De Facto Standards?* [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.electronicdesign.com/technologies/embedded-revolution/article/21796209/whats-the-difference-between-de-jure-and-de-facto-standards>

Beetham, H.; Sharpe, R. (2007). *Rethinking pedagogy for a digital age*. Londres: Routledge. doi: <https://doi.org/https://doi.org/10.4324/9780203961681>

BENVIC (2020). *Benchmarking of Virtual Campuses* [fecha de consulta: 23 abril 2020]. Disponible en: <http://www.benvic.odl.org/>

Biggs, J. B. (1993). «From theory to practice: A cognitive systems approach». *Higher education research and development* (vol. 12, n.º 1, págs. 73-85).

Brinkmann, S.; Kvale, D. (2018). *Doing interviews. Second edition*. Thousand Oaks, CA: Sage Publications.

Brown, T. (2008). «Design Thinking». *Harvard Business Review. America Latina* (n.º junio, págs. 1-10).

Buckley, A. (2014). *UK Engagement Survey 2014. The second Pilot Year*. York: The Higher Education Academy.

Bussolon, S. (2020). *Interazione persona-macchina con elementi di comunicazione multimodale –corso avanzato* [fecha de consulta: 23 abril 2020]. Disponible en: https://bussolon.s3.eu-central-1.amazonaws.com/pdf/dispensa_hci.pdf

Buzzetto-More, N. A.; Alade, A. J. (2006). «Best practices in e-assessment». *Journal of Information Technology Education Research* (vol. 5, n.º 1, págs. 251-269).

Cellini, E. (2008). *L'osservazione nelle scienze umane*. Milano: FrancoAngeli.

CERN Data Centre; Invenio; OpenAIRE; Unión Europea (2020). *Zenodo* [fecha de consulta: 23 abril 2020]. Disponible en: <https://zenodo.org>

California State University Chico (2003). «Rubric for Online Instruction» [fecha de consulta: 23 abril 2020]. Disponible en: https://www.csuchico.edu/eoi/_assets/documents/rubric.pdf

Choy, S. (2007). «Benefits of e-Learning Benchmarks: Australian Case Studies». *Electronic Journal of E-Learning* (vol. 5, n.º 1, págs. 11-20) [fecha de consulta: 23 abril 2020]. Disponible en: https://www.researchgate.net/publication/27478172_Benefits_of_e-Learning_Benchmarks_Australian_Case_Studies

Cicognani E. (2002). *Psicologia sociale e ricerca qualitativa*. Roma: Carocci.

Cocuzza, D. (2015-2020). «Benchmarking». *Glossariomarketing.it* [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.glossariomarketing.it/significato/benchmark/>

Corrao, S. (2005). *Il focus group*. Milán: Franco Angeli.

Creelman, A.; Ehlers, U.-D.; Ossiannilsson, E. (2014). «Perspectives on MOOC quality An account of the EFQUEL MOOC Quality Project». *Innoqual. The International Journal for Innovation and Quality Learning* (vol. especial 2, n.º 3, págs. 79-87) [fecha de consulta: 23 abril 2020]. Disponible en: <http://innoqual.efquel.org/>

- Creswell, J. W.** (2003). *Research design. Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage Publications.
- Crotty, M.** (1998). *The foundations of social research. Meaning and perspective in the research process*. Thousand Oaks, CA: Sage Publications.
- Davidson, J.** (2005). *Evaluation Methodology Basics. The nuts and bolts of sound evaluation*. Thousand Oaks, CA: Sage Publications.
- De Mauro, T.** (2009). *Nuovo De Mauro*. Roma: Internazionale [fecha de consulta: 23 abril 2020]. Disponible en: <https://dizionario.internazionale.it/>
- De Nicolao, G.** (2015). «Le classifiche CENSIS-Repubblica sono credibili?» *ROARS. Return on Academic Research* [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.roars.it/online/le-classifiche-censis-repubblica-sono-credibili/>
- De Nicolao, G.** (2018). «Classifica ARWU 2018 delle università: ecco la guida per giornalisti grulli». *ROARS. Return on Academic Research* [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.roars.it/online/classifica-arwu-2018-delle-universita-ecco-la-guida-per-giornalisti-grulli/>
- De Vaus, D.** (2013). *Surveys in social research*. Londres: Routledge.
- DeRouin, R. E.; Fritzsche, B. A.; Salas, E.** (2004). «Optimizing e-learning: Research-based guidelines for learner-controlled training». *Human Resource Management* (vol. 43, n.º 2-3, págs. 147-162).
- Dodge, B.; Pickett, N.** (2007). *Rubrics for web lessons* (23 abril).
- EADTU** (2020). *European Association of Distance Teaching Universities* [fecha de consulta: 23 abril 2020]. Disponible en: <https://eadtu.eu/>
- Edmonds, R.** (2006). «Best practices for e-learning». En: *Handbook on quality and standardisation in e-learning* (págs. 485-500). Berlín / Heidelberg: Springer International Publishing.
- EdX** (2014). «edX MOOC Development Checklist» [fecha de consulta: 23 abril 2020]. Disponible en: https://courses.edx.org/c4x/edx/edx101/as-set/edx_MOOC_Development_Checklist-a11y.pdf
- EFQUEL** (2011). *UNIQUe. European University Quality in e-Learning. Information Package*. Brussels: EFQUEL [fecha de consulta: 23 abril 2020]. Disponible en: https://www.google.it/url?sa=t&rcct=j&q=&esrc=s&source=web&cd=12&ved=0ahUKEwjHgtOx2NHLAhUGng4KHdT1DhoQFghdMAs&url=https://www.efmd.org%2Fprojects-test%3Fdownload%3D6%3A06-unique-guidelines-2011&usq=AFQjCNGwmgif_Jwica7rMSupMrOjJCbvZQ&cad=rja
- EFQUEL** (2015). *European Foundation for Quality in e-Learning* [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.eurashe.eu/about/partners/efquel/>
- Ehlers, U. D.** (2007). «Quality Literacy. Competencies for Quality Development in Education and e-Learning». *Journal of Educational Technology & Society* (vol. 10, n.º 2, págs.96-108).
- Ehlers, U. D.; Ossiannilsson, E.; Creelman, A.** (2013). *Week 1: MOOCs and Quality. Where are we. Where do we go from here...?* [fecha de consulta: 23 abril 2020]. Disponible en: <http://mooc.efquel.org/first-post-of-the-series/>
- Elassy, N.** (2015). The concepts of quality, quality assurance and quality enhancement. *Quality Assurance in Education* (vol. 23, n.º 3, págs. 250-261). doi: <https://doi.org/10.1108/QAE-11-2012-0046>
- Fard, A.** (2019). *The ultimate guide to stakeholder interviews: understand your clients* [fecha de consulta: 23 abril 2020]. Disponible en: <https://uxdesign.cc/the-ultimate-guide-to-stakeholder-interviews-understand-your-clients-a3bcf87b6e8b>
- Filippakou, O.; Tapper, T.** (2008). «Quality Assurance and Quality Enhancement in Higher Education: Contested Territories?» *Higher Education Quarterly* (vol. 62, n.º 1-2, págs. 84-100). doi: <https://doi.org/https://doi.org/10.1111/j.1468-2273.2008.00379.x>
- Ghislandi, P. M. M.** (2005). *Didattiche per l'Università*. Trento: Editrice Università degli Studi di Trento Dipartimento di scienze della Formazione e della Cognizione.

Ghislandi, P. M. M. (2012). «adAstra: A Rubrics' Set for Quality eLearning Design». En: Patrizia M. M. Ghislandi (Ed.). *eLearning: theories, design, software, applications* (págs. 91-106). Rijeka: InTech – Open Access Publisher.

Ghislandi, P. M. M. (2014). «Quality teaching by design for learning... or about the Brunelleschi's egg. Qualità dell'insegnamento e progetto per apprendere... o dell'uovo di Brunelleschi». *Formazione & Insegnamento. Rivista Internazionale di Scienze Dell'educazione e Della Formazione. European Journal of Research on Education and Teaching, Scholarship of Teaching and Learning per una Didattica Universitaria di Qualità* (vol. 12, n.º 1, págs.197-210). ISSN 2279-7505. doi: <https://doi.org/ISSN 1973-4778>

Ghislandi, P. M. M. (2016a). «“The fun they had” or about the quality of MOOC». *Journal of E-Learning and Knowledge Society* (vol. 12, n.º 3).

Ghislandi, P. M. M. (2016b). «Qualità della didattica accademica ai tempi dell'open education». En M. Rui, L. Messina y T. Minerva (Eds.). *Teach Different! Proceedings della multiconferenza EMEMItalia 2015* (págs. 32–37). Génova: Genova University Press.

Ghislandi, P. M. M.; Cumer, F. (2011). «Le rubric di qualità didattica per il sito Ret@ccesibile». En: T. Minerva y L. Colazzo (Eds.). *Comessi! Scenari di innovazione nella formazione e nella comunicazione*. Reggio Emilia: Ledizioni. The innovative LEDI publishing company.

Ghislandi, P. M. M.; Cumer, F. (2012). «La validazione qualitativa di adASTRA, una suite di rubric per la progettazione/realizzazione di eLearning di qualità». *L'integrazione Scolastica e Sociale* (vol. 11, n.º 3, págs. 265-271). doi: <https://doi.org/ISSN 1720-996X>

Ghislandi, P. M. M.; Cumer, F. (2013). «L'accessibilità per un e-Learning di qualità, Accessibility for eLearning quality». *Tecnologie Didattiche, Edizioni Menabò* (vol. 21, n.º 3). doi: <https://doi.org/ISSN 1970-061X>

Ghislandi, P. M. M.; Pedroni, A. (2009a). «Comunità di apprendimento distribuite e qualità della didattica. Distributed learning communities and quality of the teaching/learning process». En: T. Leo, R. Maragliano, F. **Falcinelli y P. M. M. Ghislandi (Eds.)**. *Digital collaboration: some issues about teachers' functions* (págs. 87-130). Nápoles: Editore ScriptaWEB. ISBN 978-88-6381-067-7.

Ghislandi, P. M. M.; Pedroni, A. (2009b). «Modelli e strumenti per la qualità delle comunità di apprendimento online». En: Convención *Didamatica 2009. Informatica per la didattica* (22, 23, 24 abril 2009). Trento: Facultad de Economía, Jurisprudencia y Sociología. Università degli Studi di Trento. ISBN 978-88-8443-277-3 [fecha de consulta: 23 abril 2020]. Disponible en: <http://services.economia.unitn.it/didamatica2009/>

Ghislandi, P. M. M.; Pedroni, A. (2011). «Progettare eLearning: le rubric “ad Astra” per la valutazione di corsi accademici online». En: *Didattica Universitaria Online: Teorie, Esperienze, Strumenti* (vol. 2, págs. 305-330). Nápoles: Scriptaweb.

Ghislandi, P. M. M.; Pedroni, A.; Paolino, D.; Franceschini, D. (2008a). «Quality in elearning. Some results from a national research program». En: *CEUR Workshop Proceedings* (vol. 398).

Ghislandi, P. M. M.; Pedroni, A.; Pellegrini, A.; Franceschini, D. (2008b). «eLearning e qualità». En: *Il Giornale dell'eLearning* (vol. 2, n.º 3).

Ghislandi, P. M. M.; Raffaghelli, J. E. (2012a). «Implementing quality e-Learning in higher education: change efforts, tensions and contradictions». *ICERI2012 Proceedings* (págs. 1107-1117).

Ghislandi, P. M. M.; Raffaghelli, J. E. (2012b). «La mediación del proceso de Learning Design como aporte a la calidad del aprendizaje en red». En José Antonio Jerónimo Montes (Ed.). *CIAMTE. Congreso Iberoamericano de Aprendizaje mediado por tecnologías* (10-14 septiembre 2012, págs. 433-446). Ciudad de México: UNAM. doi: <https://doi.org/ISBN978-607-02-4148-2>

Ghislandi, P. M. M.; Raffaghelli, J. E. (2012c). «Participatory and constructivist-evaluation as part of eLearning Quality: A case study». En: *Proceedings of the Fifth Annual Edition of ICERI2012: 5th International Conference of Education, Research and Innovation* (19, 20 y 21 noviembre 2012). ISBN: 978-84-616-0763-1.

- Ghislandi, P. M. M.; Raffaghelli, J. E.** (2015). «Forward-oriented designing for learning as a means to achieve educational quality». *British Journal of Educational Technology* (vol. 46, n.º 2). doi: <https://doi.org/10.1111/bjjet.12257>
- Ghislandi, P. M. M.; Raffaghelli, J. E.; Cumer, F.** (2012). «La qualità dell'eLearning: una metodologia per l'analisi qualitativa dei feedback degli studenti e dei docenti». *Ricerche di Pedagogia e Didattica. Journal of Theories and Research in Education* (vol. 2, n.º 7, págs. 25-47). doi: <https://doi.org/ISSN-1970-2221> [fecha de consulta: 23 abril 2020]. Disponible en: <http://rpd.unibo.it/article/view/3218>
- Ghislandi, P. M. M.; Raffaghelli, J. E.; Sangrà, A.; Ritella, G.** (2020). «The street lamp paradox: Analysing students' evaluation of teaching through qualitative and quantitative approaches [Il paradosso del lampione. Analizzare attraverso approcci qualitativi e quantitativi la valutazione di un insegnamento accademico da parte degli studenti]». *Journal of Educational, Cultural and Psychological Studies* (n.º 21). doi: <https://dx.doi.org/10.7358/ecps-2020-021-ghis>
- Ghislandi, P. M. M.; Raffaghelli, J. E.; Yang, N.** (2013). «Mediated Quality: An Approach for the eLearning Quality in Higher Education». *International Journal of Digital Literacy and Digital Competence (IJDLDC)* (vol. 4, n.º 1, págs. 56-63). EISSN: 1947-3508. doi: <https://doi.org/ISSN:1947-3494>
- Gibbs, G.** (2010). *Dimension of quality in higher education*. York. doi: <https://doi.org/ISBN978-1-907207-24-2>
- Gibbs, P.** (2011). «Finding quality in “being good enough” conversations». *Quality in Higher Education* (vol. 17, n.º 2, págs. 139-150).
- Goodrich H.** (1996). «Understanding rubrics». *Educational Leadership* (vol. 54, n.º 4, págs. 14-17).
- Google** (2020). «Rubric Imagini». *Google Search Engine* [fecha de consulta: 23 abril 2020]. Disponible en: https://www.google.com/search?source=univ&tbm=isch&q=rubric+immagini&client=firefox-b-d&sa=X&ved=2ahUKEwje45_hnobpAhVISHUIHS_yDF8QsAR6BAGKEAE&biw=1072&bih=724
- Green, D.** (1994). *What is Quality in Higher Education?* Buckingham: The Society for Research into Higher Education / Open University Press [fecha de consulta: 23 abril 2020]. Disponible en: <https://files.eric.ed.gov/fulltext/ED415723.pdf>
- Guba, E. G.; Lincoln, Y. S.** (1989). *Fourth generation evaluation*. Thousand Oaks, CA: Sage Publications.
- Harvey, L.; Green, D.** (1993). «Defining quality». *Assessment & Evaluation in Higher Education* (vol. 18, n.º 1, págs. 9-34).
- Herman, J. L.** (1992). *A practical guide to alternative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Horton, W.** (2001). *Evaluating e-learning*. Alexandria, VA: ASTD-American Society for Training and Development.
- ICDE** (1938-2020). *International Council for Open and Distance Education* [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.icde.org/>
- IHEP** (2020). *Institute for Higher Education Policy* [fecha de consulta: 23 abril 2020]. Disponible en: <http://www.ihep.org/>
- iNACOL** (2011). «National Standards for Quality Online Courses (v2)». *Aurora Institute* [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.inacol.org/resource/inacol-national-standards-for-quality-online-courses-v2/>
- INDIRE** (s. f.). *Istituto Nazionale Documentazione Innovazione Ricerca Educativa* [fecha de consulta: 23 abril 2020]. Disponible en: <http://www.indire.it/websemantico/news.php>
- Instituto Treccani** (2020). «Rubrica». *Vocabolario on line* [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.treccani.it/vocabolario/rubrica/>
- Kawachi, P.** (2015). «Quality Assurance for OER: Current State of the Art and the TIPS Framework». *ELearning Papers* (n.º 40). doi: <https://doi.org/ISSN1887-1542>.

Kear, K.; Williams, K.; Rosewell, J. (2012). *Quality Assessment for E-learning: a Benchmarking Approach* (2.ª ed.). Heerlen: EADTU [fecha de consulta: 23 abril 2020]. Disponible en: http://e-xcellencelabel.eadtu.eu/images/documents/Excellence_manual_full.pdf

Kim, J.; Lee, A.; Ryu, H. (2013). «Personality and its effects on learning performance: Design guidelines for an adaptive e-learning system based on a user model». *International Journal of Industrial Ergonomics* (vol. 43, n.º 5, págs. 450-461).

Kirkpatrick, D. L. (1998). «The four levels of evaluation». En: *Evaluating corporate training: Models and issues*. Dordrecht: Springer.

Kirkpatrick, J. D.; Kirkpatrick, W. K. (2016). *Kirkpatrick's Four Levels of Training Evaluation*. Alexandria, VA: Association for Talent Development.

Kocdar, S.; Aydin, C. H. (2015). «Quality Assurance and Accreditation of MOOCs: Current Issues and Future Trends». En: *Proceedings of Open Education Global 2015: Innovation and Entrepreneurship*.

Lantero, L. (1996). *Degree Mills: non-accredited and irregular higher education institutions*. Roma: CIMEA. Information Centre on Academic Mobility and Equivalence.

Learningcircuits.org. (s. f.). *Learning Circuits* [fecha de consulta: 23 abril 2020]. Disponible en: <http://www.learningcircuits.org/glossary.html>

Lim, C. J.; Lee, S. (2007). «Pedagogical usability checklist for ESL/EFL e-learning websites». *Journal of Convergence Information Technology* (vol. 2, n.º 3, págs. 67-76).

Lorenzo, G.; Moore, J. (2002). *The Sloan Consortium Report to the Nation. Five Pillars of Quality Online Education*. Nueva York, NY: The Sloan Consortium [fecha de consulta: 23 abril 2020]. Disponible en: <http://www.sloanc.org/effective/pillarreport1.pdf>

Lowenthal, P. R.; Hodges, C. B. (2015). «In Search of Quality: Using Quality Matters to Analyze the Quality of Massive, Open, Online Courses (MOOCs)». *International Review of Research in Open and Distributed Learning* (vol. 16, n.º 5).

Margaryan, A.; Bianco, M.; Littlejohn, A. (2015). «Instructional quality of Massive Open Online Courses (MOOCs)». *Computers & Education* (n.º 80, págs. 77-83).

Mariani, L. (2013). «Il ruolo dell'autovalutazione come competenza da costruire». *Lingua e Nuova Didattica* (vol. XLII, n.º 2) [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.learningpaths.org/Articoli/autovalutazioneroma.htm>

MarylandOnline (1999). *marylandonline.org* [fecha de consulta: 23 abril 2020]. Disponible en: <http://www.marylandonline.org/>

Mathes, J. (2016). «Global quality in online, open, flexible and technology enhanced education. An analysis of strenghts, weaknesses, opportunities and threats. ICDE. *International Council for Open and Distance Education* [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.icde.org/knowledge-hub/report-global-quality-in-online-education>

McAleese, M. (2013). *Report to the European Commission on Improving the quality of teaching and learning in Europe's higher education institutions*. Luxemburgo: Publication Office of the European Union. ISBN978-92-79-30360-9. doi: <https://doi.org/10.2766/42468>

Merisotis, J. P.; Phipps, R. A. (2000). *Quality On the Line: Benchmarks for Success in Internet-Based Distance Education*. Washington, DC: IHEP [fecha de consulta: 23 abril 2020]. Disponible en: <http://www.ihep.org/research/publications/quality-line-benchmarks-success-internet-based-distance-education>

Mitra, S. (2015). «Minimally invasive education: Pedagogy for Development in a Connected World». En: *International Perspectives on Home Education* (págs. 254-277). Londres: Palgrave Macmillan.

Moore, J. C. (2005). *The Sloan Consortium quality framework and the five pillars*. Nueva York, NY: The Sloan Consortium [fecha de consulta: 23 abril 2020]. Disponible en: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.4238&rep=rep1&type=pdf>

Nielsen, J. (2000). *Designing web usability: The practice of simplicity*. San Francisco, CA: New Riders Publishing.

Nilsson, M.; Palmér, M. (2002). «Semantic web meta-data for e-learning: some architectural guidelines». En: *11th World Wide Web Conference*.

OLC (2020). *Online Learning Consortium* [fecha de consulta: 23 abril 2020]. Disponible en: <https://onlinelearningconsortium.org/>

OLC (ed). (2000). «Education Volume 2: Learning Effectiveness, Faculty Satisfaction and Cost Effectiveness». *Online Learning Consortium*. Newburyport, MA: OLC.

Ossiannilsson, E.; Landgren, L. (2012). «Quality in e-learning: a conceptual framework based on experiences from three international benchmarking projects». *Journal of Computer Assisted Learning* (vol. 28, n.º 1, págs. 42-51).

Ossiannilsson, E.; Williams, K.; Camilleri, A.; Brown, M. (2015). *Quality models in online and open education around the globe: State of the art and recommendations*. Oslo: ICDE. International Council for Open and Distance Education.

Pascarella, E. (2001). «Identifying excellence in undergraduate education. Are we ven close?». *Change: The Magazine of Higher Learning* (vol. 3, n.º 33, págs. 18-23). doi: <https://doi.org/10.1080/00091380109601796>

Patton, M. Q. (1990). *Qualitative Evaluation and Research Methods. Second Edition*. Thousand Oaks, CA: Sage Publications.

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3.ª ed.). Thousand Oaks, CA: Sage Publications.

Philip, D. (2000). *Sharable Courseware Object Reference Model (SCORM). Version 1.0*. Alexandria, VA: Institute for Defense Analyses [fecha de consulta: 23 abril 2020]. Disponible en: <https://apps.dtic.mil/dtic/tr/fulltext/u2/a410103.pdf>

Phillips, R.; McNaught, C.; Kennedy, J. (2012). *Evaluating e-Learning. Guiding Research and Practice*. Nueva York / Londres : Routledge.

Quality Matters. (2014). *Quality Matters Rubric Standards. Fifth Edition* [fecha de consulta: 23 abril 2020]. Disponible en: <https://cpb-us-e1.wpmucdn.com/blogs.cornell.edu/dist/f/3782/files/2016/04/Quality-Matters-Standards-with-Point-Values-Fifth-Edition-2015-1om3rq3.pdf>

Quality Matters (2020a). «Quality Matters Rubrics & Standards». *QM. Quality Matters* [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.qualitymatters.org/qa-resources/rubric-standards>

Quality Matters (2020b). «Helping you deliver on your online promise». *QM. Quality Matters* [fecha de consulta: 23 abril 2020]. Disponible en: <http://www.qualitymatters.org/>

Raban, C. (2007). «Assurance versus enhancement: less is more?». *Journal of Further and Higher Education* (vol. 1, n.º 31, págs. 77-85).

Redazione ROARS (2013). «Cosa vogliamo». *ROARS. Return on Academic Research* [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.roars.it/online/chi-siamo/cosa-vogliamo/>

Redazione ROARS (2015). «In Italia qualcuno sta truccando i ranking QS? E che succede se ti beccano?». *ROARS. Return on Academic Research* (n.º noviembre) [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.roars.it/online/in-italia-qualcuno-sta-truccando-i-ranking-qs-e-che-succede-se-ti-beccano/>

Redazione ROARS (2019). «Ranking università USA: Berkeley squalificata per doping sui dati delle donazioni». *ROARS. Return on Academic Research* (n.º agosto).

Richards, L.; Morse, M. J. (2009). *Fare ricerca qualitativa*. Milán: FrancoAngeli.

Rosewell, J.; Jansen, D. (2014). «The OpenupEd quality label: benchmarks for MOOCs». *The International Journal for Innovation and Quality Learning* (págs. 88-100).

Rubin, J.; Chisnell, D. (2008). *Handbook of usability testing: how to plan, design and conduct effective tests*. Hoboken, NJ: John Wiley & Sons [fecha de consulta: 23 abril 2020]. Disponible en: <http://ccftp.scu.edu.cn:8090/Download/efa2417b-08ba-438a-b814-92db3dde0eb6.pdf>

Ryan, Y.; Brown, M. (2012). «Quality assurance policies and guidelines for distance education». En: *Australia and New Zealand. Quality assurance and accreditation in distance education and e-learning: models, policies, and research* (págs. 91-101).

Sae-Khow, J. (2014). «Developing of Indicators of an E-Learning Benchmarking Model for Higher Education Institutions». *TOJET.Turkish Online Journal of Educational Technology* (vol. 13, n.º 2, págs. 35-43).

SAGE (2020). «SAGE Research Methods». *Sage Publishing* [fecha de consulta: 23 abril 2020]. Disponible en: <https://us.sagepub.com/en-us/nam/sage-research-methods>

Schiavone, A. G. (2017). «Accessibilità Web in Moodle: Un'analisi attraverso le esperienze in letteratura scientifica ed un caso di studio». En: *MoodleMoot Italia 2017*. Roma: Consiglio Nazionale delle Ricerche, Sistemi informativi / Sapienza Università di Roma.

Scriven, M. (1991). *Evaluation thesaurus. Forth edition* (4.ª ed.). Thousand Oaks, CA: Sage Publications.

Shaw, I. (1999). *Qualitative evaluation*. Londres: Sage Publications. doi: <https://doi.org/10.4135/9781849209618>

Stake, R. E. (2004). *Standard Based & Responsive Evaluation*. Thousand Oaks, CA: Sage Publications.

Stracke, C. M. (2009). «Quality development and standards in e-learning: Benefits and guidelines for implementations». En: *Proceedings of the ASEM lifelong learning conference: e-Learning and workplace learning*.

THE (2019). «Best universities in the world 2020». *THE World University Rankings* [fecha de consulta: 23 abril 2020]. Disponible en: <https://www.timeshighereducation.com/student/best-universities/best-universities-world>

W3C Accessibility (2018). <https://www.w3.org/standards/webdesign/accessibility>

Wikipedia (2019). «Benchmark». *Wikipedia, l'enciclopedia libera* [fecha de consulta: 23 abril 2020]. Disponible en: [https://it.wikipedia.org/wiki/Benchmark_\(economia\)](https://it.wikipedia.org/wiki/Benchmark_(economia))

Wikipedia (2020a). «Observación». *Wikipedia, la enciclopedia libre* [fecha de consulta: 23 abril 2020]. Disponible en: <https://es.wikipedia.org/wiki/Observaci%C3%B3n>

Wikipedia (2020b). «Prototipo». *Wikipedia, la enciclopedia libre* [fecha de consulta: 23 abril 2020]. Disponible en: <https://es.wikipedia.org/wiki/Prototipo>

Wikipedia (2020c). «Test de aceptación – Acceptance testing». *Wikipedia, la enciclopedia libre* [fecha de consulta: 23 abril 2020]. Disponible en: https://es.qaz.wiki/wiki/Acceptance_testing

Williams, K.; Kear, K.; Rosewell, J.; Ferreira, G. (2011). «Incorporating quality assurance criteria for OER and Social Networking in the E-xcellence QA methodology». En: *24th ICDE World Conference «Expanding Horizons – New Approaches to Open and Distance Learning»* (2-5 octubre 2011). Bali.

Witthaus, G.; Childs, M.; Nkuyubwatsi, B.; Conole, G.; Inamorato dos Santos, A.; Punie, Y. (2015). «An Assessment-Recognition Matrix for Analysing Institutional Practices in the Recognition of Open Learning». *eLearning Papers. In-Depth* (n.º 40, págs. 1-9). doi: <https://doi.org/ISSN:1887-1542>

Wright, C. R. (2003). *Criteria for evaluating the Quality of Online Courses*. Alberta: Alberta Distance Education and Training Association.

Yousef, A. M. F.; Chatti, M. A.; Schroeder, U.; Wosnitza, M. (2014). «What Drives a Successful MOOC? An Empirical Examination of Criteria to Assure Design Quality of MOOCs». En: *2014 IEEE 14th International Conference on Advanced Learning Technologies*. doi: <https://doi.org/109/ICALT.2014.23>

Yuan, M.; Recker, M. (2015). «Not All Rubrics Are Equal: A Review of Rubrics for Evaluating the Quality of Open Educational Resources. International Review of Research in Open and Distributed Learning». *The International Review of Research in Open and Distributed Learning* (vol. 16, n.º 5). doi: <https://doi.org/https://doi.org/10.19173/irrodl.v16i5.2389>