

*Herramientas e indicadores
de medición de la calidad
docente en rankings.*

*Estado del arte y
recomendaciones para la
selección de indicadores.*

Endika Bengoetxea

2 diciembre 2016

Resumen

Los rankings universitarios son un tema de actualidad y a la vez polémico por la metodología empleada en los rankings existentes y por las consecuencias de utilizarlos para fines por los que no fueron concebidos.

La mayoría de rankings universitarios no han sido diseñados para presentar una comparación de la calidad docente de los estudios universitarios, si bien en los últimos años algunos de ellos incorporan indicadores con este propósito. Esto contribuye a la tendencia de que los rankings se utilicen para que nuevos estudiantes seleccionen una universidad o para inspirar decisiones tanto en los rectorados como en las agencias o gobiernos que financian las universidades.

Este informe presenta estas últimas tendencias en rankings internacionales, centrándose en las metodologías e indicadores que se utilizan para la medición de la calidad docente de las distintas universidades. Se presentan sus aspectos positivos y sus limitaciones, el modo en que se recopilan los datos, así como las últimas tendencias internacionales en cuanto a la calidad docente y su medición. Se revisan estos aspectos en el contexto internacional, sobre todo en el europeo, ya que las iniciativas que se desarrollan en la Comisión Europea y por los países firmantes de la Declaración de Bolonia son las que inspiran la mayoría de tendencias en evaluación de la calidad en el contexto del estado y del País Vasco.

Este documento presenta asimismo algunas recomendaciones para definir indicadores de la calidad docente que pudieran servir a Unibasq para preparar un conjunto de indicadores a difundir al público en general.

Palabras clave

Ranking universitarios, educación superior, transparencia en educación superior, análisis de decisiones multi-criterio, excelencia en educación superior

Herramientas e indicadores de medición de calidad en la actividad docente en rankings

Introducción

Cada vez hay mayor consenso sobre que la educación superior es clave para el futuro de Europa. Es necesario que los sistemas de educación superior ofrezcan mayor calidad, dar respuesta a la demanda de formación a un cada vez más diverso grupo de personas (incluyendo la formación a lo largo de la vida), adaptando la oferta formativa a los cada vez más diversos requisitos en materia de competencias y habilidades que son necesarios. Esto es cada vez más relevante ya que los tipos de empleos del mercado de trabajo mundial evolucionan muy a menudo. En este contexto, la mayoría de las instituciones de educación superior se ven cada vez más impulsadas a diversificar sus perfiles, de forma que cada institución pueda centrarse en sus puntos fuertes (European Comisión, 2011). La crisis económica ha acrecentado esta tendencia, de forma que las instituciones han pasado de competir con el resto a centrarse en impulsar sus propias fortalezas y buscar alianzas en las demás para poder dar respuesta más eficazmente.

En Europa, la Comisión Europea apoya esta estrategia de diversificación de perfiles de las instituciones de educación superior como parte integrante de la estrategia de modernización de las universidades europeas (European Comisión, 2011). El fortalecimiento de sus diversas misiones en materia de educación, investigación e innovación, es una condición sine qua non para el éxito de la estrategia Europa 2020 que busca establecer a Europa como líder mundial en la economía del conocimiento. En la iniciativa emblemática *Juventud en Movimiento* (European Comisión, 2010a) la Comisión subraya la necesidad de modernización de la educación superior europea en diversos ámbitos: mayor excelencia en la docencia, una mejor correspondencia con las demandas del mercado laboral y más excelencia en investigación e innovación.

Reformas como las impulsadas por el Proceso de Bolonia han traído consigo una creciente diversidad de la educación superior, y esto se ha traducido además en una creciente tipología de instituciones de educación superior con muy diversas misiones. Si bien esta diversidad se considera una de las ventajas del Espacio Europeo de Educación Superior, cobra mayor relevancia la necesidad de transparencia (European Comisión, 2010a; 2010b; 2011): es necesario disponer de información más clara sobre cómo las instituciones desempeñan sus funciones y responden a sus misiones cara a estudiantes, empleadores y responsables políticos. Los responsables universitarios también son beneficiarios directos de la transparencia, ya que les facilita disponer de información adecuada en la toma de decisiones estratégicas: no es exagerado señalar que muchos rectorados desconocen muchas de las áreas en las que destaca su institución sobre las demás (salvo quizás en ciertos casos, principalmente en lo referente a investigación).

En aras de encontrar información sobre la calidad docente de las universidades, tenemos sobre todo dos referencias: las revisiones realizadas por agencias de calidad independientes, y los rankings. En ambos casos se utilizan distintas metodologías para medir la calidad docente, y sobre todo en el caso de los rankings se utilizan indicadores que intentan medir la complejidad de la calidad docente y el aprendizaje de forma puramente cuantitativa para poder hacer comparaciones entre instituciones de educación superior. Los informes de evaluación de las agencias de calidad independientes se centran en aspectos más cualitativos y suelen tener información más detallada, pero son a menudo muy técnicos para el público general y no se suelen presentar de forma que faciliten la comparación del desempeño de cada institución.

Cabe destacar que en cuanto a la medición de la excelencia no hay teoría científica que defina qué es la mejor universidad, y por lo tanto los rankings comparan las universidades de acuerdo a unos criterios definidos con unos objetivos muy concretos. Tampoco hay autoridades que definan las reglas de juego en los rankings. Esto conlleva que una organización no tiene más autoridad que otra cuando crea un ranking y se ciñe a un conjunto de reglas e indicadores.

A pesar de ello cada ranking tiende a presentar la información como si sus indicadores reflejaran por sí solos la calidad de una institución, como si hubieran sido diseñados en base a una (inexistente) teoría de calidad de la educación superior [REF multirank], con la consiguiente confusión entre el público que además se amplifica por la búsqueda de titulares impactantes y sesgados de los medios de comunicación.

En este contexto, la medición de la calidad docente es un aspecto que cada vez cobra más relevancia ya que la docencia es una misión de la universidad de especial interés para el público en general y para los estudiantes en particular. Los rankings se utilizan a menudo para tomar decisiones sobre la docencia, pero no está claro que los indicadores que se utilicen sean los más adecuados para ese propósito. Este informe presenta el estado del arte, tendencias futuras, y presenta recomendaciones sobre cómo podría definirse un conjunto de indicadores de la calidad docente de las universidades en el País Vasco como herramienta de transparencia de la calidad docente para el gran público.

Tendencias en transparencia de la educación superior

El objetivo de las llamadas políticas de transparencia de la educación superior es el de asegurar la existencia de herramientas que faciliten y aseguren la transparencia en la forma de operar y de medir el rendimiento de las instituciones de educación superior. Mayor transparencia con respecto a cómo las diferentes instituciones realizan su trabajo facilita a los estudiantes e investigadores elegir dónde estudiar y trabajar. En el caso de muchas instituciones, estas herramientas se ven como una oportunidad de acceso a información más accesible y comparable, que les ayuda a posicionarse mejor y mejorar sus estrategias institucionales de desarrollo, calidad y rendimiento.

Las herramientas de transparencia aplican metodologías para apoyar la rendición de cuentas y evaluar la calidad de la educación superior. Los poderes públicos están especialmente interesados en promover este tipo de herramientas para poder poner en marcha políticas más adecuadas.

Las herramientas para promover la transparencia en educación superior en la enseñanza universitaria incluyen premios y concursos a la calidad docente o al mejor profesor, mecanismos de aseguramiento de la calidad y acreditación, y los rankings (y ratings). La diferencia entre los rankings y los ratings es la siguiente:

- Los RANKINGS comparan universidades una a una y las pone en orden según una medida de performance en base a diversos indicadores. El orden en el que aparecen las universidades se calcula agrupando (habitualmente, ponderando) diversos aspectos en un único indicador. Ejemplos incluyen el ranking de Shanghai (*ARWU*) y el de *Leiden*.
- Los RATINGS evalúan las universidades según su desempeño en diversas áreas, de forma similar a los rankings. Sin embargo, en vez de comparar instituciones una a una, estas se comparan con un conjunto de estándares (si bien estos estándares son en muchos casos relativos, es decir, contabilizan el desempeño de las demás universidades en base a aquella con los mejores resultados en cada indicador). Los ratings son útiles para informar a los gestores de las instituciones sobre cómo podrían posicionarse, además de para identificar debilidades y fortalezas de cada universidad. Ejemplos incluyen *U-multirank* y el *QS stars*

Los rankings (y ratings) se consideran una herramienta más de transparencia, complementaria a otras tales como las existentes en el contexto del Sistema Europeo de Garantía de Calidad. Los rankings proporcionan una comparativa general que puede servir de orientación inicial, pero la información debe ser complementada con otras herramientas antes de utilizarla para la toma de decisiones. Los rankings no pueden sustituir a otras herramientas como el sistema de garantía de calidad articulado a través de agencias. Aunque se acuda más a los rankings que a la información de agencias de calidad por su facilidad de uso e interpretación (si bien estas listas de universidades no se concibieron

para el uso que algunos usuarios le otorgan), esto simplemente muestra la necesidad de mejorar cómo se comunica al público los resultados de la acreditación y la evaluación de la calidad de las agencias. En cualquier caso, los rankings son útiles para ofrecer una visión global inicial.

Lamentablemente, los rankings existentes (tales como los más conocidos como los de Shangai, o Times Higher Education) adolecen de importantes desventajas:

- A pesar de que las instituciones de educación superior pueden ser excelentes en varios ámbitos, estos rankings tienden a analizar una sola misión (o dimensión), y se centran casi exclusivamente en la investigación, ignorando otras áreas como la calidad de la enseñanza, la internacionalización, la innovación y el compromiso con el entorno.
- Los indicadores se agrupan para dar un valor único de forma que se asocian según ponderaciones decididas unilateralmente por los creadores del ranking (ej. El 40% a reputación).
- Los rankings se aplican a instituciones en su conjunto en lugar de a nivel de disciplina. Una aproximación disciplinar sería sin duda más interesante para la mayoría de los estudiantes y para la búsqueda de alianzas con otros socios institucionales.
- Muchas veces se ignoran las diferentes culturas de publicación y citación en las disciplinas, haciendo un sesgo hacia algunas de ellas (sobre todo las ingenierías y las ciencias)
- Favorecen de forma especial a las instituciones más anglófonas y disciplinas como las ingenierías, resultando en una penalización a universidades con formación en otras lenguas y a las generalistas sobre las politécnicas.
- No cubren la diversidad de la educación superior y, en la práctica, no incluyen en sus listas más de alrededor del 3-5% de las instituciones de educación superior de todo el mundo.

Otras críticas se centran en la metodología adoptada (por ejemplo, muchos rankings incluyen el prestigio entre los indicadores), el tipo de indicadores elegidos (a menudo se centran en evaluar lo que se puede medir en vez de lo que deberían medir) y, en muchos casos, en la falta de transparencia que ciertos rankings muestran a la hora de explicar la metodología para obtener la tabla final que constituyen los rankings (Buela-Casal et al., 2007; EUA, 2011; Salmi, 2009).

Sin embargo, uno de los mayores riesgos es el de utilizar de formas para las que no han sido concebidos, y los efectos que esto provoca en las propias instituciones de educación superior. Por ejemplo, ciertos rankings tienen en consideración el número de estudiantes o profesores de la institución que han obtenido un premio Nobel. La reacción de algunas universidades ha sido la de contratar (aunque sea de forma parcial) a profesores que hayan obtenido un Nobel para subir posiciones, y esto evidentemente no está claro que esté en línea ni con la agenda de modernización de las universidades que es necesaria, ni siquiera que aporte directamente un plus a los estudiantes de la universidad. En (EUA, 2011; Salmi, 2009) se muestran ejemplos de utilización de rankings para los que no están concebidos y que resultan en efectos imprevistos

A pesar de estas críticas y potenciales desventajas, los rankings están para quedarse y son hasta ahora las únicas herramientas de transparencia que ofrecen una comparación (si bien muy sesgada) del grado de eficiencia de las instituciones de educación superior, tanto a nivel nacional como internacional. Es más, algunos gobiernos están impulsando el establecimiento de rankings nacionales de sus universidades para condicionar aspectos clave tales como la financiación.

Con el objetivo de mejorar el estado del arte y definir un conjunto de directrices sobre cómo deberían de ser los rankings, el *International Ranking Expert Group* desarrolló los denominados "Principios de Berlín sobre Rankings de Instituciones de Educación Superior" (IREG, 2006). Estos principios de Berlín, aceptados por los países firmantes de la Declaración de Bolonia y por la Comisión Europea, destacan que cada ranking tiene su propio propósito y grupo(s) objetivo(s), y que esta finalidad condiciona los supuestos modelados, indicadores seleccionados y pesos de cada indicador. Es por ello que la finalidad de cada ranking debe explicarse claramente para poder interpretar los resultados correctamente. Posteriormente en el año 2008, bajo la Presidencia francesa de la Unión Europea, se celebró la conferencia "Comparación internacional de sistemas educativos ¿Un modelo europeo?", en la que se subrayó la necesidad de una metodología nueva en materia de rankings que tenga en consideración todas las misiones (dimensiones) en cuanto a la excelencia de los centros de educación superior y centros de investigación, tanto en Europa como en el contexto internacional. (French Government, 2008)

La complejidad de la calidad docente y aprendizaje

La *calidad docente o del aprendizaje* es un concepto complejo, que además carece de una definición clara. Esto se debe a que el concepto de calidad se puede entender como un resultado, una propiedad, o incluso como un proceso [REF]. Además, este concepto es dependiente del grupo de interlocutores, de manera que estudiantes, profesores y las agencias de calidad por ejemplo tienen interpretaciones diferentes de qué es y por lo tanto cómo se debería de medir su impacto y eficiencia.

En la literatura se pueden distinguir varios niveles en los que la calidad docente tiene distintos componentes y metas: Sistema, Institucional, Disciplina y/o titulación, y Profesor individual.

Sin embargo, a pesar de estos condicionantes, hay consenso en la literatura sobre el hecho de que ni el tamaño ni las especificidades de las instituciones son un obstáculo para desarrollar políticas universitarias para promover la calidad docente en la forma que sea más adecuada a cada uno de estos niveles [REF]. Además, también hay acuerdo en la necesidad de considerar tanto los procesos top-down como bottom-up a la hora de medir la calidad (y sobre todo cuando queremos promover la innovación en la docencia). Esto último es además una de las críticas más habituales a los procesos de evaluación de la calidad de las agencias independientes, donde algunos autores contraponen la promoción de la innovación docente y los procesos de acreditación, ya que los procesos de evaluación de la calidad son a veces tan rígidos que no reconocen o valoran suficientemente aspectos innovadores con la consiguiente desmotivación del profesorado.

Además de la necesidad de encontrar un equilibrio en la promoción de la innovación docente y la existencia de criterios de acreditación robustos, a menudo se considera que las metodologías de aseguramiento de la calidad docente no abordan suficientemente la complejidad de la docencia. Las críticas más habituales son que estos procesos están muy enfocados al input y no a los resultados del aprendizaje, cambios en la actitud y comportamiento de los estudiantes (que se consideran lo que idealmente debería de hacerse). Sin embargo, en la práctica es muy difícil medir resultados del aprendizaje o las competencias mejoradas de los estudiantes, y esto resulta aún más complicado si se pretende hacerlo con criterios únicamente cuantitativos como usando sólo indicadores, como es el caso de la mayoría de rankings universitarios.

Metodologías para medir y mejorar la calidad docente

Aunque hay muchas formas de evaluar la calidad docente, destacan sobre todo tres:

1. Cuestionarios a los estudiantes

Este tipo de herramientas se han generalizado en los últimos años debido a que la calidad docente está necesariamente centrada en los estudiantes.

A pesar de ello, la literatura demuestra que hay varias debilidades de este método a tener en cuenta cuando se analizan los resultados, entre la que destaca que los estudiantes tienden a culpar a los profesores de más problemas de los que les corresponde, olvidando el rol de las infraestructuras institucionales. Por otro lado, los profesores tienden a tener muy poca fe en los cuestionarios a estudiantes, y muchas veces estos cuestionarios provocan que los profesores adopten estrategias a corto plazo inadecuadas e incluso indeseadas.

La literatura coincide en que los cuestionarios de estudiantes son muy útiles cuando el objetivo de la evaluación es la mejora de la calidad docente y no para penalizar al profesorado. Es por ello que este tipo de encuestas se utilizan muy a menudo también en premios de calidad educativa o concursos para premiar al mejor docente.

2. Evaluación en-clase entre iguales

Este mecanismo se está popularizando porque se centra sobre todo en el proceso y no solamente en el resultado.

Sus detractores señalan sin embargo que si se utiliza mal, puede desmotivar la innovación docente y provocar la estandarización en la forma e impartir docencia según los patrones establecidos y no necesariamente de mayor impacto.

3. Teaching portfolios

Los portfolios proporcionan evidencia documentada de la calidad docente desde fuentes muy diversas, de forma que pretenden analizar toda la carrera docente del profesorado.

Si bien puede haber diferentes formas de presentar un portfolio docente, los siguientes elementos se consideran más habitualmente:

- La filosofía o planteamiento personal sobre la docencia
- Documentación sobre el historial docente personal (asignaturas impartidas, material de formación...)
- Efectividad de la docencia (resumen de encuestas de estudiantes sobre la docencia del profesor, otras formas de feedback de estudiantes y de ya titulados...)
- Material que demuestre el aprendizaje de los estudiantes (puntuaciones en test estandarizados o de otro tipo, trabajos presentados por los estudiantes, comparaciones del trabajo antes y después de la docencia...)
- Actividades para la mejora continua del profesor
- Contribuciones a la profesión de la docencia en su institución o en otra
- Premios, felicitaciones o reconocimientos

Todos estos métodos de evaluar la calidad docente no son aplicables en la mayoría de rankings universitarios globales, ya que las evaluaciones más completas y robustas requieren de mucho más esfuerzo que la mera recopilación de datos numéricos en forma de indicadores.

Sin embargo, muchos rankings hablan de la calidad de las universidades de forma global incluso cuando asignan más importancia a la evaluación de la calidad investigadora y a veces apenas contienen indicadores en calidad docente. Hay varias razones que explican por qué la calidad investigadora predomina en los rankings sobre la docente:

- Los indicadores internacionales de la calidad investigadora están consolidados y son comúnmente aceptados

- Es difícil encontrar datos comparables internacionalmente en cuanto a educación, lo que conlleva grandes dificultades para crear indicadores y que suelen provocar grandes diferencias en los resultados de los diferentes países.

Algunos justifican la utilización de indicadores de investigación únicamente asumiendo que la calidad investigadora es en realidad un buen proxy de la calidad docente, ya que las universidades con gran calidad investigadora también lo son en calidad docente. En la literatura hay tanto estudios que avalan la relación bidireccional como los que concluyen que de facto están en conflicto directo debido a cómo se incentiva las dos misiones. Tampoco faltan defensores de las metodologías de ranking existentes que muestran correlaciones entre los indicadores de los rankings de calidad investigadora y docente, pero la realidad muestra que las instituciones muy orientadas a la investigación tienen serias dificultades para marcar su perfil de calidad docente con respecto a la investigadora (por ejemplo, algunas instituciones de este tipo apenas tienen estudiantes de grado y existen rankings que muestran el *ratio doctorando vs. estudiantes de grado* como un indicador de calidad de la universidad).

En este debate, es necesario tener en cuenta que actualmente en muchas áreas geográficas del mundo la sociedad espera de las universidades que tomen partido por extender su misión docente de forma que cubra una dimensión más social y no se centren en ser únicamente universidades de élite, sino que atiendan también al gran público. Por ejemplo, se espera de las universidades no sólo que faciliten el acceso a sus jóvenes graduados al mercado laboral, sino que también sean ciudadanos responsables. Por ello en algunos rankings se empiezan a incluir incluso indicadores que contemplen dimensiones tales como su contribución al desarrollo regional o la inclusión social. Y lo más importante de todo: estos aspectos afectan cada vez más la reputación de las universidades.

Marco de medición del impacto de una intervención aplicado a la docencia y el aprendizaje

Excelencia e impacto son dos conceptos íntimamente unidos cuando se pretende medir una intervención (e.g. una política de promoción de la calidad docente o la docencia de una titulación) en un sistema (e.g. una sociedad o una universidad). Sin embargo, la medición de resultados e impacto constituye a menudo uno de los mayores retos que a menudo conducen a que las mediciones de la calidad y efectividad de una intervención se centren más en aspectos que se pueden medir en vez de en los resultados que motivaron la intervención.

Así, por ejemplo, si quisiéramos medir la efectividad de una formación para fomentar el emprendizaje, más allá de centrarnos en el número de start-ups creadas lo ideal sería centrarse en medir el éxito a largo plazo en términos de impacto socio-económico o creación de empleo. Sin embargo, en algunos casos lo que se tiende a medir en muchas metodologías de calidad docente no son resultados a corto o largo plazo, sino aspectos más bien de input como por ejemplo el número de estudiantes matriculados, calidad de las instalaciones docentes, número de profesores, la inversión realizada a priori, etc. (cuando lo ideal sería centrarse en perspectivas de carrera, éxito y desempeño en el área de estudio por parte de los estudiantes).

A la hora de medir los efectos de una intervención de cualquier tipo, es necesario distinguir distintos niveles en cuanto a *outputs*, *outcomes* e impactos.

- **Outputs** son condiciones generadas directamente como resultado de una actividad – a menudo cuantificable. Ejemplos son en nuestro caso el número o porcentaje de graduados y la satisfacción de los estudiantes. Típicamente estos se refieren a los productos directos de las actividades realizadas (a menudo no abordando más de dos años vista) que se espera producirán los outcomes deseados.

- **Outcomes** son los efectos obtenidos o plausibles de ocurrir en el corto o medio plazo cuando se ponen a disposición de los beneficiarios de la intervención. El término outcomes es típicamente asociado a impactos, resultados, y consecuencias externas. Ejemplos incluyen por ejemplo la tasa de empleo de los titulados en la formación. En el contexto educativo, el término outcome es a menudo utilizado como sinónimo de resultado.
- **Impacto** es lo que ocurre a un nivel superior y engloba todos los efectos a largo plazo producidos por una política o proyecto, ya sean primarios y secundarios, y esperados o no.

Además del input, output, *outcome* e impacto, también se suelen considerar en algunos modelos el Proceso (ej. Calidad de la selección de estudiantes, seguimiento de los estudiantes, facilidades para conectar con el mercado laboral, etc.) y el Contexto (aspectos demográficos, situación económica del entorno, facilidades e incentivos para la carrera de los buenos profesores, etc.)

Un ejemplo para ilustrar estas definiciones es la formación de *líderes reconocidos en una profesión* como el impacto de alto nivel que se desea alcanzar, donde elementos como el sueldo inicial o el tiempo en encontrar el primer empleo encajado son un *outcome/resultado* de la titulación, y donde *output* es el número de titulados.

Debido a la complejidad de la calidad docente, con multitud de formas de crear impacto a nivel sistémico y temático, en este marco conceptual la evaluación debería de tener en cuenta cómo una titulación o Universidad contribuye a lograr un impacto, prestando atención a las necesidades de todos los interlocutores relevantes, desarrollando nuevas metodologías docentes y en definitiva contribuyendo a la mejora de la docencia de forma eficiente y efectiva.

Indicadores de calidad docente en los rankings internacionales

La posibilidad de que la calidad docente pueda medirse por medio de indicadores de forma robusta es algo que se cuestiona en la literatura científica, si bien algunos proyectos como por ejemplo el denominado AHELO (desarrollado por la OCDE) concluyó tras un estudio de varios años y en varios países de todo el mundo que efectivamente es posible medir los *learning outcomes* de forma comparable independientemente de los condicionantes culturales de cada país haciendo por ejemplo encuestas a estudiantes.

Sin embargo, debido a que las encuestas a estudiantes son uno de los métodos más habituales para obtener información sobre calidad docente para algunos de estos rankings, otro elemento distorsionador a considerar es el de la lealtad de los estudiantes a su universidad, y el hecho de que sus respuestas no sean del todo objetivas debido a que el resultado pueda tener consecuencias en la percepción del público sobre la calidad de su propio título. Hay tres modos principales a la hora de utilizar indicadores:

- Como un indicador de clase (ej. indicadores estándares y cuantitativos tales como presupuesto, universidad generalista vs. Politécnica, etc). Estos indicadores son sobre todo de ‘input’ y en rankings como U-Multirank se utilizan para hacer tipologías de instituciones y así no comparar directamente universidades que tengan condicionantes básicos muy diferentes (ej. no tendría mucho sentido comparar las universidades del País Vasco con el MIT en muchos aspectos, ya que el presupuesto del MIT comparado con el de nuestras universidades hace que estas comparaciones sean injustas).
- Como indicador de performance ó rating
- Como un indicador de satisfacción, habitualmente por medio de una encuesta ya sea de estudiantes, profesorado o empleadores, y que habitualmente se recoge por medio de encuestas de algún tipo.

Además del marco conceptual que describe los criterios y su contenido, es esencial que los indicadores estén basados en estándares metodológicos de investigación empíricos. Esto se traduce esencialmente en que los indicadores deben ser:

- Válidos: deben aportar evidencia sobre si el indicador representa adecuadamente lo que se pretende evaluar.
- *Reliable*: se refiere a la consistencia de un conjunto de mediciones o de instrumentos de medición.
- Comparables: en un ranking se comparan instituciones entre ellas con indicadores numéricos, por lo que los datos o mediciones en los que se basan deben ser comparables entre instituciones y deben medir el mismo tipo de calidad entre ellas.

Todos estos son aspectos teóricos, pero la realidad de los rankings universitarios actualmente no sigue en muchos casos estas pautas. De hecho, los indicadores de calidad docente más habituales en los rankings son los siguientes:

- Graduados que hayan obtenido un premio Nobel
- Ratio profesor/estudiante
- Encuestas de reputación (a académicos, estudiantes, y empleadores)
- Inversión en educación (tanto en términos absolutos como relativos)
- Porcentaje de personas que finalizan los estudios
- Tiempo en acabar los estudios

Algunos rankings miden la calidad docente en ocasiones por medio de encuestas a estudiantes en el aula, sin embargo, esta práctica no está muy extendida en los rankings globales ya que cada institución suele tener un proceso de evaluación diferente y los resultados no son siempre accesibles fuera de la universidad. Por ello, los rankings tienden a basarse en métricas que (se asume por los creadores del ranking) están correlacionados con la calidad docente (ej.: encuestas a empleadores, que hacen de *clientes* sustituyendo las encuestas a los estudiantes)

Como se ha señalado, uno de los mayores problemas para definir indicadores de calidad docente es la ausencia de datos internacionales y, para aquellos que existen, su validez para realizar comparaciones. En algunos países existen encuestas que se realizan a los estudiantes de todas las universidades y que se utilizan por algunos rankings. Algunos ejemplos son el *National Survey on Student Engagement* en los EEUU, y el *College Student experience Questionnaires* de Australia. En Europa no existe un cuestionario de este tipo, si bien la Comisión Europea puso en marcha en 2012 un proyecto junto con EURSTAT para la recogida de datos sistemática y comparable de las universidades, y que se denomina *European Project For Tertiary Education Register* (ETER). Éste último se va mejorando cada año y no contiene resultados de encuestas a estudiantes, pero es una fuente de información muy interesante para rankings en un futuro muy próximo.

Conclusiones. La medición de la calidad docente en rankings internacionales y tendencias

Se presenta en el Anexo una lista exhaustiva de indicadores de calidad docente utilizados por los rankings universitarios globales más relevantes. Es importante destacar que los rankings actualmente están evolucionando y que algunos incluyen datos estadísticos e indicadores que no se utilizan para hacer la lista ordenada de las universidades, pero debido a que tienen un interés informativo para el público se incluyen en la tabla del anexo.

Lo siguiente es un resumen de los aspectos más relevantes en forma de tendencias:

- En vista de los resultados de los indicadores de calidad investigadora y docente de los rankings actuales, se puede concluir que el supuesto de que buenos investigadores son automáticamente buenos docentes es cuando menos cuestionable;

- Los estudiantes universitarios son cada vez más diversos y además su número va en aumento. Esto viene acompañado de la expectativa de que las universidades tengan en consideración cada vez más una dimensión social, que se ve reflejada en los indicadores más novedosos en algunos rankings tales como:
 - Tasa de abandono,
 - Éxito en grupos más desfavorecidos,
 - Indicadores (de input) sobre la condición social de los estudiantes
 - ...
- No es posible medir únicamente la docencia, y la tendencia es utilizar métodos para medir también el aprendizaje, que se traduce en algunos indicadores que empiezan a verse o a debatirse tales como:
 - La relación y accesibilidad del profesorado con los estudiantes
 - La promoción y oportunidades de la vida de estudiante en los campus
 - Atención en las *comunidades de aprendizaje*
 - Promover liderazgo y la capacidad de crear "*change agents*"
 - Un buen profesor ayuda a sus estudiantes, para la contribución al mundo de la docencia sería muy pobre si éste no comparte sus experiencias; por ello la difusión de métodos docentes innovadores es un valor en alza

En cuanto a las conclusiones sobre los indicadores analizados:

- Llama la atención que, tal y como se esperaba, los rankings tienden a centrarse en indicadores de input, y no hay muchos de outcome e impacto que son en definitiva los que más interés suscitan e el público en general. Se confirma también para el caso de la calidad docente que los indicadores que se utilizan no muestran toda la complejidad y resultados de la docencia:

- Los indicadores de performance se eligen actualmente por las universidades y los rankings por ser criterios cuantificables y accesibles, y no porque evalúen la calidad docente de forma robusta.
- Las encuestas de reputación favorecen a las universidades que sean bien conocidas, independientemente del mérito:
 - Las encuestas de reputación no son "Peer reviews", aunque se denominen de esa forma;
 - La reputación está relacionada con el conocimiento a priori, no con performance
- El porcentaje profesor-alumno, al usarse como proxy para la calidad docente, no sirve como medida de aspectos pedagógicos, del aprendizaje o de su entorno. Podría incluso dar pie a tomar medidas para la mejora de la calidad en la dirección equivocada.
- El porcentaje de abandono de los estudios tiene el riesgo de ser mal interpretado: los abandonos pueden ser debidos a circunstancias y problemáticas externas. Por ejemplo, una universidad que sigue una política más inclusiva entre grupos de estudiantes menos favorecidos tiende a tener más abandonos.
 - Lo que sí es útil es la evolución de la tasa de abandonos a lo largo del tiempo, ya que un aumento en esta tasa puede ser debido a que los estudiantes no se adaptan adecuadamente a la vida universitaria.
- Indicadores como '*inclusion of work experience*' son muy interesantes y en las titulaciones debe incluir aspectos como prácticas en empresa, profesores con experiencia profesional fuera de la academia, etc. Otro que puede ser atractivo a priori es el de '*international orientation of programs*', que también se compone de otros sub-elementos. Sin embargo, este tipo de indicadores suele ser muy complicado porque si uno de los datos para

uno de estos elementos que componen el indicador no existen el indicador no podría ser calculado.

- Algunos indicadores como el de la empleabilidad son muy dependientes del Contexto y no son atribuibles en gran medida a las universidades: un ejemplo es el de la empleabilidad en situaciones económicas como la crisis actual con la complejidad de un mercado laboral muy desfavorable.
- Hay que tener cuidado con la definición de los indicadores cuando se intenta aplicar a todo tipo de disciplinas. Por ejemplo, la definición de *post-doc* es diferente entre disciplinas: los post-docs son habituales en carreras de ciencias, pero el concepto puede llevar a confusión en áreas como ciencias sociales, y por ejemplo no está claramente definido en las humanidades. Otro aspecto a menudo problemático al aplicarse a disciplinas y titulaciones es el de *Transferencia de Conocimiento* (si bien éste último no suele ser problemático al aplicarlo al nivel de la universidad). Debe prestarse cuidado a la definición de los indicadores y a cómo estos pueden ser interpretados entre las disciplinas.
- A menudo el aspecto internacional puede ser complicado, sobre todo en cuanto se refiere a un título doble o conjunto internacional. Es un problema de interpretación que suele necesitar cuidado en la definición.
- Para evitar confusiones en indicadores como *Porcentaje de profesorado internacional*, debe señalarse si “profesorado internacional” se refiere a aquellos que no tienen la nacionalidad del país o aquellos que se han contratado desde el extranjero.
- Los indicadores sobre prácticas en empresa suelen ser a menudo complicados de recopilar. Esta problemática es además diferente entre disciplinas, por ejemplo, en empresariales esto suele ser más sencillo que en ingenierías. A pesar de las dificultades, las prácticas en empresa y las tesis desarrolladas conjuntamente con empresas son indicadores muy

recorridos para analizar el *regional engagement*, ya que se consideran en cierta forma experiencias de transferencia de conocimiento, en particular para PYMES. Al mismo tiempo, éstas crean enlaces con potenciales empleadores de los estudiantes y tienen un rol importante en la empleabilidad de los estudiantes.

Recomendación para utilizar un conjunto de indicadores de ranking para transparencia al público

Como se ha indicado anteriormente, una característica esencial en la elaboración de cualquier ranking o de indicadores con vocación de ser aceptados es que en su definición y desarrollo estén representados todos los interlocutores relevantes.

En [CHERPA-Network (2011)] tenemos algunas buenas ideas a considerar a la hora de diseñar una herramienta con indicadores que vayan a ser públicos:

- Tal y como se señala en los principios de Berlín, es necesario definir los grupos objetivo a los que se dirigen los rankings e indicadores, así como cómo se presentaran los resultados. Esto evitará en cierta medida malinterpretaciones de los indicadores.
- Los rankings y los mecanismos de aseguramiento de la calidad son instrumentos complementarios. Los rankings muestran a la institución desde una perspectiva exterior y cuantitativa; los mecanismos tradicionales de aseguramiento de calidad institucionales, tanto los internos como externos, se centran en la rendición de cuentas y mejora de la calidad. Aunque los rankings no son similares a los instrumentos de aseguramiento de la calidad, pueden ser útiles para identificar las preguntas adecuadas en los procesos de mejora internos y externos.
-

- Para algunos grupos objetivos a los que van dirigidos los indicadores, en particular los estudiantes y los investigadores, la información debe ser basada en la disciplina/titulación; para otros (ej. gestores universitarios y poderes públicos) la información prioritaria será la de nivel institucional (relacionada con la orientación estratégica de las universidades). Un conjunto de indicadores debe reflejar estas necesidades tan diferentes.
- Salvo en contadas excepciones, los rankings no deberían utilizar tablas ordinales de 1 a N sino marcar diferencias claras y robustas en niveles de desempeño. La decisión sobre el número óptimo de categorías de performance debe definirse teniendo en cuenta el número de universidades incluidas en el ranking y la distribución de los datos.
- Los rankings habitualmente utilizan varias bases de datos para mostrar diferentes perspectivas del desempeño de las universidades. Deben utilizarse tanto como sea posible fuentes de información accesibles, pero su disponibilidad suele estar limitada. Por ello, la calidad del proceso de recogida de datos es esencial.
- Los rankings deben ser definidos teniendo cuidado con las potenciales consecuencias imprevistas o efectos perversos que puedan generar.
- Si bien la calidad en los procesos de asegurar la calidad de los rankings es importante, también es crucial que haya transparencia sobre el proceso y los instrumentos que se utilizan en el ranking. Esta transparencia debe incluir también las limitaciones del ranking.

Además de estas recomendaciones, deberían de tenerse en cuenta las siguientes consideraciones:

- Seguramente será más efectivo mostrar los indicadores como un *rating*, y no como un ranking, donde se haga una clasificación en 4 ó 5 clases por cada indicador o grupo de indicadores, poniendo un color a cada clase (como si fuera un semáforo). Hacer un ranking al uso con un orden uno a uno no portaría mucha información y podría dar lugar a malentendidos.
- Ya que no se pretende hacer un ranking sino elegir un conjunto de indicadores como herramienta de transparencia para el público, debería evitarse elegir únicamente aquellos que sean fácilmente medibles pero que no sean buenos *proxy* de lo que se pretende medir. En algunos casos podría ser que esto requiera de las universidades un esfuerzo para recoger datos que no tengan actualmente, pero a la larga es mejor un indicador bueno que un que sea demasiado simplista
- Los indicadores deberían seleccionarse teniendo en cuenta que cubran la mayor parte de los siguientes apartados:
 - Idealmente los indicadores deberían de cubrir aspectos como el impacto y el *outcome*, y no centrarse demasiado en el input.
 - Habrá casos en los que escuelas y facultades no puedan ser comparados de forma justa dependiendo de algunos condicionantes. Estos condicionantes pueden ser indicadores de clase y podría plantearse que dependiendo de ello no siempre se compararan universidades y titulaciones. Por ello se sugiere que los indicadores se separen por su naturaleza de clase, performance o de satisfacción.

- Relacionado de forma muy cercana a la calidad docente hay varias dimensiones que contribuyen a esta calidad de una u otra manera, y que suele tener diferente importancia para el público que va a leer el ranking. Se recomienda que los indicadores se consideren en base a si se refieren a Infraestructura/recursos, docencia y aprendizaje, internacionalización, transferencia de conocimiento, y dimensión social. Como sub-apartados se incluirían la calidad del profesorado, la de la infraestructura docente o de campus, la empleabilidad etc.

Todas estas consideraciones se han recogido en la lista de indicadores que se presenta en el Anexo, si bien esta es una clasificación por indicador que es una opinión personal del autor y está abierta a discusión. En definitiva, lo más relevante sería que los indicadores que se elijan abarquen la mayor parte de estas consideraciones, y si además éstos también abordan todos los aspectos de los procesos de calidad que realiza Unibasq a las titulaciones y universidades aún mejor.

- Los indicadores seleccionados deberían tener en cuenta aquellos que sean más utilizados por los rankings internacionales, sin embargo, algunos de estos indicadores si bien parecen similares no son siempre los mismos. Es posible que al final sea mejor incluso definir un indicador nuevo en estos casos. Ej: la empleabilidad tras la graduación es uno de los indicadores que más se repiten en los rankings, pero algunos lo recogen a los 3 meses de la graduación otros a los 12 meses, y otros a los 18. En Euskadi la encuesta del Gobierno Vasco recoge los datos del empleo a los 3 años tras la graduación, además del empleo encajado. Se recomienda seguir con la encuesta del gobierno vasco si se eligiera este indicador.
 - Se sugiere que se consideren otros indicadores que, si bien actualmente aún no forman parte de los rankings internacionales, la tendencia actual muestra que lo serán pronto. Esto incluye sobre todo información de seguimiento de los titulados, como por

ejemplo aquellos que trabajan en el extranjero o que prosiguen sus estudios (en esto último, idealmente se consideraría el haber sido admitido en instituciones que sean conocidas por procesos de selección de estudiantes muy restrictivos).

- Al margen de aspectos metodológicos, un elemento importante a considerar es el del costo de obtener los datos para alimentar el indicador. No se debería elegir indicadores que no se justifiquen por el coste de obtención de datos
- La información sobre desarrollo regional podría incluir el número de programas de formación continua que ofrece cada Universidad. Esta información no debería ser difícil de obtener para las universidades y este tipo de aspectos son importantes para que la sociedad sea consciente de la importancia e impacto de las universidades en el a región.
- Es importante que los indicadores no se utilicen fuera de contexto. Por ejemplo, la EHU al ser una universidad pública no puede decidir en gran manera sobre la cantidad de estudiantes en primero que debe aceptar, ya que tiene una función social que lo solicita. Por ello indicadores como la tasa de abandonos son de facto peores que en las otras universidades.
- De forma similar, las universidades necesitan posicionarse en base a sus fortalezas y debilidades, y necesitan definir estrategias que pueden conllevar alguna penalización en algunos indicadores pero que de por sí son positivas. Por ejemplo, sino de las universidades es muy fuerte en internacionalización y ya aparece por encima de las demás en indicadores, se debe seguir incentivando que siga siendo fuerte en internacionalización (no olvidemos que la competición no es entre universidades vascas, sino con todas las universidades en general) y no desviar demasiado su atención hacia otras áreas donde no puede ser líder.

Por todo esto, una propuesta inicial de indicadores a consideración incluiría los siguientes:

Tabla 1: Propuesta inicial de indicadores para su consideración para elaborar un conjunto de indicadores por Unibasq sobre calidad docente.

El símbolo (*) señala indicadores que o tienen una viabilidad menor o que, aunque no sean utilizados actualmente por los rankings la tendencia muestra que lo serán en un futuro próximo.

<i>Infraestructura/Recursos</i>
Inversión por estudiante (de total del presupuesto institucional)
Ratio de ingresos de la institución por profesor
<i>Docencia y aprendizaje</i>
Ratio profesor/ alumno
Tasa de graduación/ abandono
Tiempo medio para la obtención del título (contrastada con la duración teórica del título)
Acreditaciones nacionales o internacionales de los títulos
Satisfacción del alumnado: valoración general del programa de estudios (encuestas)
Satisfacción de titulados: valoración general del aprendizaje (encuestas) (*)
Ratio de doctorados otorgados por profesor
Porcentaje de doctorados finalizados vs. admitidos
Porcentaje de empleo (encajado) de los titulados
Porcentaje de titulados que prosiguen los estudios tras graduarse (*)
Sueldo medio de los graduados (*)
<i>Orientación Internacional</i>
Porcentaje de personal académico internacional
Porcentaje de estudiantes internacionales
Porcentaje de estudiantes entrantes y salientes en movilidad de créditos (ej. Erasmus)
Grados y masters ofrecidos en idioma extranjero
Titulados que trabajan en el extranjero (*)
<i>Dimensión social</i>
Equilibrio hombres/ mujeres entre los estudiantes
Estudiantes con becas para estudios
<i>Transferencia de conocimiento / Compromiso regional</i>
Prácticas para estudiantes en empresas locales / regionales (*)
Número de graduados que trabajan en la región (*)
Número de spin-offs / start-ups

– **Referencias**

- Buela-Casal, G., Gutiérrez-Martínez, O., Bermúdez-Sánchez, M., Vadillo-Muñoz, O. (2007). Comparative study of international academic rankings of universities, *Scientometrics*, Vol. 71, No. 3, pp. 349-365.
- CHERPA-Network (2011). U-Multirank Final Progress Report, from: <http://www.u-multirank.eu>
- EUA (2011). *Global university Rankings and Their Impact*. ISBN: 9789078997276. From the official web of the European University Association at <http://www.eua.be>
- European Commission (2010a). *Youth on the Move – An initiative to unleash the potential of young people to achieve smart, sustainable and inclusive growth in the European Union*, COM(2010) 477 final.
- European Commission (2010b). *Europe 2020 Flagship Initiative Innovation Union*, SEC(2010) 1161.
- European Commission (2011). *Supporting growth and jobs – an agenda for the modernisation of Europe's higher education systems*, COM(2011) 567 final.
- EUMIDA-consortium (2010). *Feasibility Study for Creating a European University Data Collection*. From: <http://ec.europa.eu/research/era/docs/en/eumida-final-report.pdf>
- French Government (2008). EU Presidency conference on “International comparison of education systems: a european model ?”. From: <http://www.ue2008.fr>
- IREG (2006). *Berlin Principles on Ranking of Higher Education Institutions*. IREG-2 Conference in Berlin, 18-20 May, 2006. From: <http://www.ireg-observatory.org>
- OECD (2008). *Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value-Added of Schools*. ISBN: 9789264050228.
- OECD (2010). *Assessment of higher education learning outcomes (AHELO)*. From: <http://www.oecd.org/edu/ahelo>

- Salmi, J. (2009). The of challenge establishing World class universities, The World Bank, Washington. From: <http://www.worldbank.org/>
- van Vught, F., Ed. (2009). Mapping the Higher Education Landscape: Towards a European Classification of Higher Education. Springer. ISBN 978-90-481-2248-6
- van Vught, F., Kaiser, F., File, J., Gaethgens, C., Peter, R. and Westerheijden, D.F. (2010). U-Map: The European Classification of Higher Education Institution. Center for Higher Education Policy Studies (CHEPS), University of Twente. Final report of the U-map project, funded by the European Commission's Lifelong Learning Programme. From: <http://www.u-map.eu>
- van Vught, F., Ziegele F., Ed. (2012). Multidimensional ranking - the making of U-multirank. Springer. ISBN 978-94-007-3004-5