

UNA PROPUESTA DE ÍNDICE DE ACTIVIDAD CIENTÍFICA Y SU VALIDACIÓN EN EL CAMPO DE CIENCIAS SOCIALES Y HUMANIDADES

Julieta Barrenechea

Javier García Fronti

Javier Castro Spila

Andoni Ibarra

INTRODUCCIÓN

Los Ranking académicos internacionales que permiten comparar universidades a nivel global se han generalizado en los últimos años (Buena-Casal et al., 2007). Uno de los más destacados es el publicado por The Institute of Higher Education, Shanghai Jiao Tong University (Liu and Cheng, 2005) siendo su criterio metodológico predominante la publicación en revistas científicas.

Asimismo se han realizado clasificaciones a nivel regional europeo (Merisotis, 2002) y también a nivel de país como por ejemplo el CHE en Alemania (Federkeil, 2002) o el RAE en Reino Unido (Eccles, 2002). En las mencionadas clasificaciones, las publicaciones académicas tienen un rol cada vez más importante a la hora de evaluar la calidad de investigación en diferentes contextos académicos. Por ejemplo, (Bence and Oppenheim (2004) describen la evolución de su importancia en la evaluación gubernamental de la investigación en universidades en Reino Unido (RAE). Si bien formalmente se trata de un proceso de evaluación de pares, los autores remarcan el rol hegemónico de las publicaciones en el proceso de evaluación.

La empresa *Thomson Reuters*¹⁹, por medio de su base de datos "Science", se ha transformado en el árbitro de las revistas científicas a nivel mundial. Su sitio Web describe el Informe de citas de revistas que ellos

¹⁹ <http://thomsonreuters.com/>

producen como “un medio sistemático y objetivo para evaluar de manera crítica las revistas más importantes del mundo”²⁰. Su uso extendido otorga a la empresa un increíble poder sobre la comunidad académica y sobre los organismos de financiación. Isagani Cruz (2008) es uno de los exponentes de quienes proponen deconstruir esta situación hegemónica y critica tres conceptos fundacionales de su lógica argumental: a) sostiene que la metodología no es objetiva, b) que su evaluación de impacto no es correcta y c) que no es global pues tiene en cuenta principalmente las publicaciones en idioma inglés.

En el caso de España por ejemplo, Moya-Anegón *et al* (2004) realizaron un informe en el cuál sistematizan información sobre la producción científica del país con el fin de ofrecer un instrumento de política pública. Sin embargo su metodología utiliza como información base solamente las publicaciones internacionales indexadas por la empresa *Thomson Reuters* y como consecuencia gran parte de la producción científica ha quedado fuera de su consideración.

En el caso específico de las ciencias sociales y humanidades (CSH) la evaluación con foco en las publicaciones indexadas en Thomson Reuters es mucho más problemática. Estudios como los de (Becher, 1993), (Kivik, 2003); (Leeuwen, 2006); (Archambault, et al. 2006), ponen de relevancia la heterogeneidad que existe entre las disciplinas de CSH en sus modos de organizar la investigación y en los patrones de producción y distribución de resultados.

A partir del VII Programa Marco Europeo se asigna una mayor importancia relativa a las CSH ya que se incluyen líneas de investigación específicas para estas disciplinas que dejan así de ser subsidiarias de las Ciencias Básicas y Tecnológicas, como ha venido sucediendo en los anteriores Programas Marco.

En este nuevo escenario europeo, es un tema de agenda el problema de contar con indicadores para evaluar la actividad de investigación en CSH, puesto que no existen claros consensos sobre cuáles son los parámetros

²⁰ <http://ip-science.thomsonreuters.com/es/productos/jcr/>

más indicados para captar y valorar las actividades de investigación en estos campos disciplinarios. A pesar de este déficit, existen amplios acuerdos en relación con lo insuficiente y a veces inadecuado, que resulta aplicar al caso de las CSH criterios de evaluación elaborados para las Ciencias Básicas. Las principales dificultades conceptuales a la hora de evaluar las ciencias sociales han sido recogidas en (Ibarra et al: 2006) y se refieren a tres temas: a) las CSH tienen pautas de funcionamiento distintas en su actividad científica respecto de las ciencias experimentales (trabajo individual, mayor orientación hacia ámbitos locales, prácticas de colaboración menos frecuentes), b) la interpretación de sus resultados es dificultosa porque no se cuenta con criterios consolidados para valorar sus publicaciones, visibilidad, modalidades de colaboración y relevancia según sus ámbitos de contribución (local); c) los patrones o estándares de resultados y difusión son poco reconocido dentro de la lógica del status científico (exposiciones, artículos de revista locales, recensiones, exposiciones, etc.)

En función de lo anterior, si se reduce la complejidad de estos campos disciplinarios a unos pocos criterios de evaluación, se corre el riesgo de no evaluar lo que se debe evaluar o, simplemente, evaluar transfiriendo criterios desarrollados en otros campos científicos y aplicarlos al *modo de hacer* de las disciplinas de las CSH. Asumir esta heterogeneidad disciplinaria y su consiguiente complejidad evaluadora implica asumir el desafío de diseñar un modelo de evaluación que sea *integral*.

Con *integral*, en este trabajo nos referimos a dos planos: el primer plano se vincula con la idea de que a partir de una definición más inclusiva de *actividad científica*, será posible captar una mayor variedad en las prácticas y tipo de resultados que permitan procesar de manera analítica el funcionamiento de nuevos *patrones disciplinarios de investigación* (enfoque descriptivo), y al mismo tiempo, establecer de manera sintética niveles o estándares de productividad y calidad (enfoque evaluativo). El segundo plano está relacionado con el hecho de que la calidad de la actividad científica depende cada vez más de la capacidad de interactuar y mantener vinculaciones efectivas con una gama más amplia y compleja de esferas y agentes sociales. Una evaluación *integral* supone entonces que a la hora de evaluar un centro de conocimiento, se necesita tener en cuenta un escenario más diverso que el estrictamente científico-académico. Es

necesario desarrollar e implantar estrategias metodológicas que integren de forma complementaria las valoraciones que sobre la actividad científica tiene un conjunto más amplio de agentes sociales. El concepto de *calidad relacional* que articula esta propuesta integral de evaluación, está basado en la idea de que, en las condiciones actuales y frente a las exigencias que se plantean al sector científico-académico, la producción, distribución y uso de conocimiento científico alcanzará mejores estándares de calidad en la medida en que se garanticen buenas condiciones relacionales para estos procesos.

El Índice Directo de Actividad Científica (IDAC) que se presenta en este artículo forma parte de los resultados de un estudio desarrollado durante el año 2005 referido a la evaluación y el papel de las Ciencias Sociales y Humanidades de la UPV/EHU y a su inserción en el contexto de innovación de la Comunidad Autónoma del País Vasco (CAPV) ²¹ y cuyos principales resultados se han sistematizado en Barrenechea et al (2008) .

En el marco de dichos estudios se ha desarrollado una estrategia evaluativa que plantea un modelo de *Programa Integral de Evaluación de la Actividad Científica en CSH* organizado en cuatro fases consecutivas y articuladas que suponen triangulación metodológica.

Green et al (2001: 30) señalan que la triangulación metodológica en evaluación permite reducir la incertidumbre y comprender mejor los fenómenos estudiados. Tomando como referencia los trabajos de Cook (1985) y Greene et al. (1989) señalan 4 dimensiones en las que la triangulación metodológica es un aporte importante para la evaluación científica: a) mejora la validez y la credibilidad de las inferencias, b) ofrece una visión más comprensiva y contextual de los resultados y sus impactos, c) permite un entendimiento más profundo sobre la actividad que se está evaluando, d) aumenta la conciencia respecto de la diversidad de valores que están en juego en la actividad científica y es más inclusivo.

²¹ El estudio citado ha sido desarrollado por el equipo del Área de Investigación en Gestión de Redes de Ciencia, Tecnología e Innovación, de la Cátedra Sánchez-Mazas de la Universidad del País Vasco/Euskal Herriko Unibertsitatea.

La primera fase, incluye un proceso de *evaluación por indicadores* que tiene el valor de un autodiagnóstico sobre la base de un modelo de evaluación integral de la actividad científica en CSH. En la segunda fase, se propone un proceso de evaluación por panel de expertos (*peer review*), que es complementario de la evaluación por indicadores, y que tiene el valor de una evaluación externa. Lo que se plantea para esta segunda fase está a medio camino entre una evaluación institucional (que valora condiciones académicas e institucionales de la investigación), y una evaluación de la calidad de la investigación (que valora los resultados obtenidos por los procesos de investigación según parámetros disciplinarios). La tercera fase, incluye el diseño y aplicación de indicadores relacionales que den cuenta de la conectividad académica y social de los investigadores. Se desarrolla a partir de una estrategia de consulta a investigadores y agentes sociales, y apunta a identificar en términos descriptivos, evaluativos y prospectivos la oferta académica y la demanda social efectiva y potencial, con el fin de fortalecer dinámicas de conectividad y colaboración entre los centros de conocimiento y su entorno social relevante (mediato e inmediato). La cuarta fase está orientada a la implementación de mecanismos que favorezcan acuerdos institucionales al interior de la comunidad científica en relación con la mejora de condiciones para las actividades de investigación.

El calificativo de *integral* por el que se apuesta en este trabajo, expresa dos niveles de análisis que hacen a la producción, distribución y uso del conocimiento científico. Por una parte, se vincula con la definición de actividad científica y expresa la necesidad de considerar de manera sistémica un conjunto más amplio de dimensiones que forman parte de la actividad científica, y que complementan a aquéllas que se centran estrictamente en la obtención de productos y en los mecanismos y soportes que se ponen en juego para la difusión de los resultados de investigación. En cambio en la noción integral de actividad científica, se incorporan elementos como la formación de recursos humanos, la participación en redes académicas, la trayectoria del personal investigador, la consideración de productos menos ortodoxos, etc.

En el caso de las CSH asumir una perspectiva integral y aplicar como enfoque la triangulación metodológica se hace muy necesario en tanto

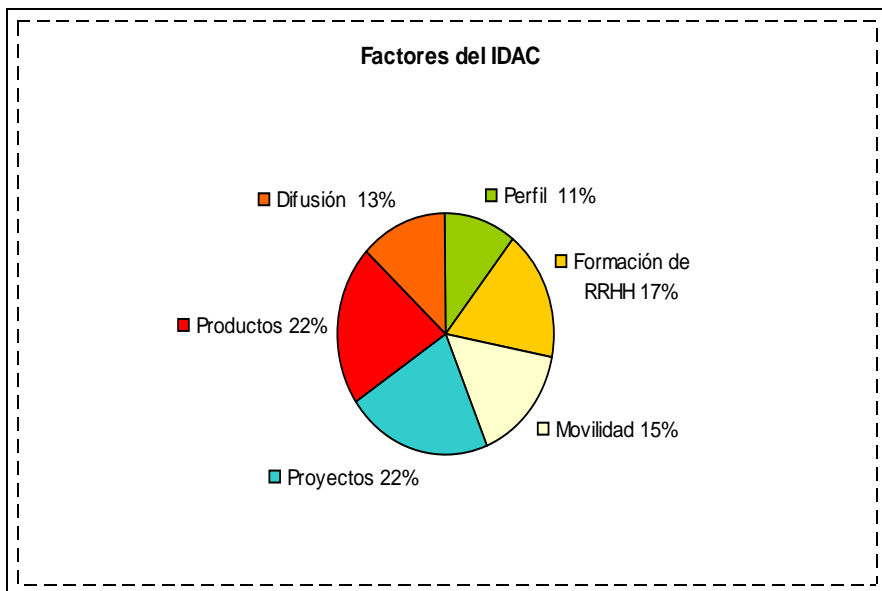
que como se ha mencionado, bajo la etiqueta de las CSH se aglutinan un conjunto heterogéneo de disciplinas, cuyos modos de producir y distribuir conocimiento son diferentes y por ello resulta difícil que se reduzcan a unos cuantos criterios estandarizados de evaluación. A esta situación de heterogeneidad disciplinaria se suma el hecho de que el debate y las técnicas que se aplican a la evaluación de la actividad científica de las CSH está lejos aún de lograr consensos y de contar con herramientas sistematizadas. En este contexto paradigmático y técnico, una aproximación integral para la evaluación de la actividad científica tiene también objetivos exploratorios: se propone captar el funcionamiento de patrones disciplinarios de investigación y, en ese marco, reunir elementos que permitan en el mediano plazo establecer niveles de calidad.

El Índice IDAC que se presenta a continuación, forma parte de las técnicas que se aplican en la fase II del Programa Integral de Evaluación desarrollado en el marco del estudio mencionado y para su cálculo se utiliza la base de datos que resulta de la aplicación de la encuesta de evaluación de la actividad científica elaborada también para este estudio. El índice ha sido diseñado para ser calculado en función de la ponderación de los factores e indicadores que componen el modelo de indicadores de evaluación. En este artículo se incluyen también los resultados de la consulta de validación con expertos que se ha aplicado en el marco del estudio mencionado y que recoge el análisis de las propuestas de ponderación para los indicadores del modelo

El IDAC es una medida síntesis que permite cuantificar en un solo valor la actividad científica de investigadores o grupos para un período determinado. Sin embargo, el IDAC se compone de 6 índices parciales que permiten medir la actividad científica en los factores que lo componen: Perfil, Formación de Recursos Humanos, Movilidad, Proyectos, Productos y Difusión de la Investigación. En este sentido, ofrece la posibilidad de analizar perfiles investigadores así como identificar patrones de actividad científica según la combinatoria que resulte de estos 6 factores. El índice puede también ser utilizado como herramienta de gestión por objetivos.

1 MODELO CONCEPTUAL DEL IDAC

En el marco de la Fase I del Programa Integral de Evaluación desarrollado en el estudio mencionado, se incluye la aplicación y cálculo de un índice de actividad científica que ha sido diseñado en función de la ponderación de los indicadores de evaluación que componen el modelo de evaluación por indicadores. El Índice Directo de Actividad Científica (IDAC) permite cuantificar en un solo valor la actividad científica de investigadores o grupos para un período determinado. El IDAC se compone a su vez de 6 índices parciales que permiten medir la actividad científica en los factores: Perfil, Formación de Recursos Humanos, Movilidad, Proyectos, Productos y Difusión de la Investigación. En el siguiente gráfico se presentan los seis factores y sus ponderaciones que expresan su contribución relativa a la actividad científica considerada globalmente:



Los seis factores están compuestos por un conjunto de indicadores que integran los distintos niveles ponderados en los que se desagrega el índice, hasta llegar al nivel más empírico compuesto por datos que se

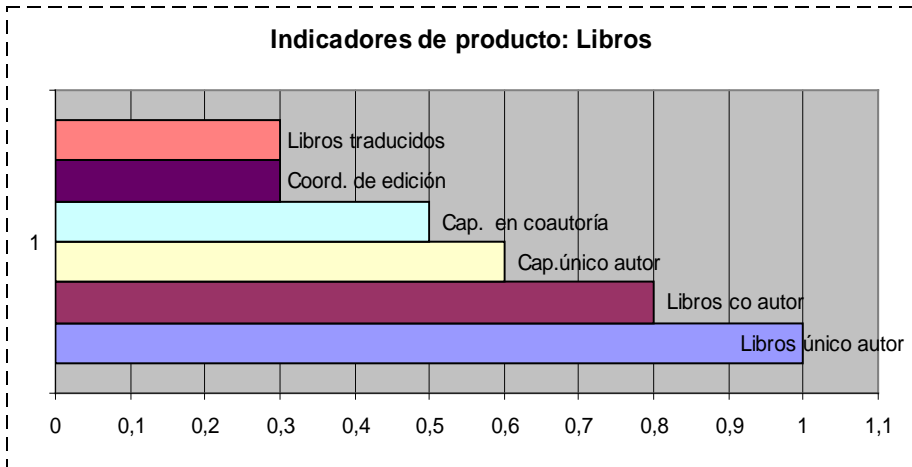
recogen mediante una medición basada en la aplicación de un cuestionario a los investigadores individuales. El número de niveles de desagregación varía según cada factor siendo el mínimo 2, como en el caso de *movilidad académica*, y 4 el máximo, como en el caso del factor *productos* cuyo indicador revistas se desagrega hasta un cuarto nivel al que corresponde una ponderación P4.

El sistema de ponderaciones del *modelo conceptual* del índice responde en todos sus niveles (factores, indicadores, ítem) a conocimientos y supuestos teóricos previos, y combina dos modalidades:

1. Ponderaciones porcentuales.
2. Valoración por equivalencias.

1 – En el caso de las ponderaciones porcentuales el valor atribuido representa la aportación de cada uno de los factores o indicadores al total. Por ejemplo, en el caso de los 6 factores, el porcentaje atribuido a cada uno de los factores equivale a la aportación de cada factor a la actividad científica considerada como un todo. Se representa en el Gráfico 1.

2 – En el caso de la valoración por equivalencias el indicador que es valorado con "1" es el referente. El resto de los indicadores del grupo son traducibles, es decir, guardan equivalencia con este indicador. Por ejemplo, en el caso de los productos *libros* el concepto *libros único autor* = 1 y libros coautor = 0.8, de modo que un libro coautor equivale a 0.8 libros único autor. Conceptualmente lo que se calcula en ese grupo de indicadores y que se obtendrá como resultado, es un valor que represente la *cantidad de equivalentes de libro único autor* que el investigador o grupo ha producido en el período estudiado. El gráfico de barras permite visualizar las equivalencias.



Cada una de las modalidades de ponderación supone una exigencia distinta respecto de la actividad investigadora que desarrollan los investigadores o grupos. En el caso de la ponderación porcentual, para poder alcanzar la máxima puntuación los investigadores deben cumplir con todos los factores o conceptos que conforman el grupo ponderado. En el caso de la valoración por equivalencia cumplir con todos los indicadores del grupo no es un requisito para obtener la máxima puntuación. Esta última modalidad de ponderación se aplica a indicadores de dos tipos: (a) indicadores acumulativos y/o excluyentes por definición, como en el caso del nivel de formación donde si se es doctor no se necesita ser master y se es indefectiblemente licenciado y (b) indicadores en los que no se considera una exigencia razonable que un investigador tenga que cumplir con todos los indicadores del grupo, por ejemplo, no se considera obligatorio que un investigador produzca traducciones o reseñas además de capítulos de libro como único autor. Aunque aquel investigador que tenga una buena producción y en conceptos diversos seguramente alcanzará una mejor puntuación en el grupo de indicadores.

2 METODOLOGÍA PARA EL CÁLCULO DEL IDAC

La información que se toma en consideración para el cálculo del IDAC proviene de la aplicación de un cuestionario a investigadores individuales. Los valores que se toman como datos se llamarán “v” y pueden ser cualidades (ej. ser o no doctor), cantidades (ej. número de artículos en coautoría) o porcentuales (ej. intensidad investigadora). En el caso de los indicadores con ponderación porcentual, estos valores deben estandarizarse en un número que varía entre 0 y 1 antes de ser ponderados, de modo que para estos casos al valor más empírico o que proviene de un nivel de cálculo anterior se lo llamará “q”, y se llamará “v” al valor estandarizado. Se estandariza en función del valor máximo (M) obtenido en ese indicador por la población bajo estudio diferenciada según áreas disciplinarias. El algoritmo de cálculo del IDAC se basa entonces en la suma ponderada de valores estandarizados en función de máximos que resultan de la muestra o población estudiada.

El IDAC resulta de la suma ponderada de los subíndices de sus factores y oscila entre 0 y 100. Los subíndices de cada uno de los factores en que se descompone la actividad científica resultan de la suma ponderada de las distintas variables convenientemente estandarizadas que lo integran y también oscilan entre 0 y 100. La ponderación de los valores obtenidos en cada indicador constituye la forma canónica del término que integra la fórmula de cálculo del índice IDAC y de cada uno de los subíndices que lo componen, y es igual a $(v^{ni} P^{ni})^{22}$.

La estandarización de las variables se realiza en distintos momentos, según se trate de indicadores con ponderación porcentual o indicadores con valoración por equivalencia, y toma en cuenta el valor máximo (M) obtenido por la población bajo estudio diferenciada según áreas disciplinarias.

Dependiendo del número de niveles en que se desagrega cada uno de los factores de actividad científica, se necesitará realizar distinto número de pasos para obtener cada subíndice de factor. El cálculo se inicia multiplicando el “v” obtenido en los indicadores de mayor nivel de desagregación por su valor o ponderación, siendo 4 el máximo nivel y 2 el

²² Más abajo puede verse la definición completa de la fórmula.

mínimo. Todos los pasos del cálculo del IDAC respetan el algoritmo del índice de modo que se obtienen subíndices para cada grupo de indicadores de distinto nivel de desagregación.

$$S^n = \begin{cases} \sum_{i=1}^{kn} v^{ni} P^{ni} \\ \sum_{i=1}^{kn} \frac{v^{ni} P^{ni}}{M^n} \end{cases}$$

1) El subíndice (S) del indicador n es igual a la sumatoria de los términos de la forma $V \cdot P$.

2) El subíndice (S) del indicador n es igual a la sumatoria de los términos de la forma $v \cdot P$, dividida por el máximo valor de sumatoria obtenido por la población bajo estudio y perteneciente a la misma área disciplinaria.

Donde:

n es el número entero positivo que representa el nivel (Ej. Perfil $n=1$)

S^n es la puntuación (expresada en porcentajes) obtenida en el nivel n

i es el número entero que varía entre 1 y k (Ej. para el indicador Grado académico: Doctor 1.1, Master 1.2, Licenciatura 1.3, Diplomatura 1.4) $i= 1$ y $k= 4$

k define el número de términos que componen el indicador n

M es el máximo valor obtenido por los investigadores de la disciplina en el nivel n .

2.1 INTERPRETACIÓN DEL ÍNDICE

Por su construcción el índice IDAC mide la posición relativa de un investigador o un grupo (departamento, instituto, etc) con respecto al rango de la población estudiada y para el período medido, y adoptará un valor igual a 100 en el caso en que el investigador o grupo presente el valor máximo de la población y el período estudiado, y 0 si presenta el valor mínimo.

2.2 PROCEDIMIENTO DE CÁLCULO DEL IDAC

Tanto en el caso de indicadores con ponderación porcentual como en los que tienen valoraciones por equivalencia, el método de ponderación es directo. Sin embargo el cálculo de los subíndices para los grupo de indicadores de uno y otro tipo difiere en algunos pasos que se describirán a continuación.

2.3 PROCEDIMIENTO DE CÁLCULO PARA INDICADORES CON PONDERACIÓN PORCENTUAL

El subíndice (o puntuación) obtenido por un investigador para un grupo de indicadores con ponderación porcentual es el resultado de realizar los siguientes pasos:

- a) Obtener $v =$ para cada indicador que compone el grupo, estandarizar los valores obtenidos en el nivel empírico o en el nivel de desagregación anterior en números entre 0 y 1.
- b) Ponderar $v =$ multiplicar el v obtenido en cada indicador por su ponderación.
- c) Sumar los términos del indicador: sumar el total de términos (v^{ni} P^{ni}) que componen el indicador.

a) Obtención de v por estandarización

Para los valores que se han llamado q y que se expresan en cantidades o porcentajes obtenidos por un investigador en un indicador con ponderación porcentual, el procedimiento de estandarización consiste en tomar como referencia el máximo valor obtenido en ese indicador para el área disciplinaria que corresponda y en el caso de la aplicación/evaluación que se está analizando, y dividir el valor por el máximo obtenido en el área disciplinaria. Aun cuando las mediciones se realicen para períodos superiores a 1 año (ej 5 años), los máximos serán los promedios anuales.

$$v^{ni} = \frac{q^{ni}}{Mx^{ni}}$$

2.4 Mx^{ni}

b) Ponderación de v

La ponderación de los valores obtenidos en cada indicador constituye la forma canónica del término que integra la fórmula de cálculo del índice IDAC y de cada uno de los subíndices que lo componen. Deberá realizarse este cálculo para cada indicador n desde su máximo nivel de desagregación hasta el mínimo (el máximo nivel posible es 4 y el mínimo 1); y para todos los términos de cada indicador, es decir, para i desde 1 hasta k .

$$(v^{ni} P^{ni})$$

c) Sumatoria de los términos

Para cada grupo de indicadores con ponderación porcentual que integra el modelo conceptual del IDAC es posible calcular un subíndice realizando la sumatoria de todos los términos $(v^{ni} P^{ni})$ que lo componen.

kn

$$S^n = \sum_{i=1} v^{ni} P^{ni}$$

Cada subíndice calculado será llamado v cuando se sube un nivel de indicador, y se comporta como un valor a ser ponderado para el cálculo de un subíndice de menor nivel de desagregación de un indicador con valoración por equivalencia.

Cada subíndice calculado será llamado q cuando se sube un nivel de indicador y se comporta como un valor a ser estandarizado para el cálculo de un subíndice de menor nivel de desagregación de un indicador con ponderación porcentual.

2.5 PROCEDIMIENTO DE CÁLCULO PARA INDICADORES CON VALORACIÓN POR EQUIVALENCIA

El subíndice obtenido por un investigador para un grupo de indicadores con valoración por equivalencia es el resultado de realizar los siguientes pasos:

- a) Ponderar v : multiplicar el v obtenido en cada indicador por su ponderación.
- b) Calcular las unidades (U): realizar la sumatoria del total de términos ($v^{ni} P^{ni}$) que componen el indicador.
- c) Estandarizar las unidades: dividir el resultado de la sumatoria de los términos, sobre el valor del máximo valor en unidades obtenido en el mismo indicador por la población bajo estudio perteneciente a la misma área disciplinaria.

a) Ponderación de v .

La ponderación de los valores obtenidos en cada indicador constituye la forma canónica del término que integra la fórmula de cálculo del índice IDAC y de cada uno de los subíndices que lo componen.

$$(v^{ni} P^{ni})$$

b) Cálculo de unidades (U)

El resultado de la sumatoria de los términos que componen un grupo de indicadores con valoración por equivalencia representará el total de la producción científica realizada por el investigador o grupo en relación con ese ítem medido en la unidad de equivalencia correspondiente. Se denominará unidad U seguido por el número del indicador correspondiente.

kn

$$U^n = \sum_{i=1}^n v^{ni} p^{ni}$$

c) Estandarización de Unidades.

Para todos los grupos de indicadores con valoración por equivalencia, el subíndice se obtiene una vez que se han estandarizado los resultados de la sumatoria de términos o “unidades” respecto del máximo valor en unidades obtenido en ese mismo grupo de indicadores por la población bajo estudio que pertenece al mismo campo disciplinario.

$$S^n = \frac{U^n}{M^n}$$

Cada subíndice calculado será considerado v cuando se sube un nivel de indicador y se comporta como un valor a ser ponderado para el cálculo de un subíndice de menor nivel de desagregación.

Cada subíndice calculado será llamado q cuando sube un nivel de indicador y se comporta como un valor a ser estandarizado para el cálculo de un subíndice de menor nivel de desagregación de un indicador con ponderación porcentual.

3 RESULTADOS DE LA CONSULTA DE VALIDACIÓN

En el marco del estudio desarrollado para el caso de la UPV/EHU, con el fin de validar el modelo de indicadores y las ponderaciones asignadas para el cálculo del IDAC, se ha realizado una consulta a expertos e investigadores de CSH tanto del ámbito de la Comunidad Autónoma del País Vasco (CAPV), como del ámbito estatal y europeo. El fin ha sido recoger sus opiniones sobre los indicadores diseñados y evaluar los valores de ponderación asignados a los distintos factores e indicadores que se incluyen en el índice de actividad científica.

La consulta ha consistido en el envío de un cuestionario electrónico a una muestra de expertos en evaluación y temas de educación superior y a un pequeño número de investigadores de CSH en su calidad de usuarios de procesos de evaluación.

Los consultados debían establecer un valor de ponderación en porcentaje para cada uno de los indicadores que componen el IDAC. La ponderación de cada indicador debía guardar relación con la ponderación del resto de los indicadores del mismo grupo, de modo que la suma de las ponderaciones de los indicadores del mismo grupo debía ser 100²³. El cuestionario ha incluido la posibilidad de realizar comentarios abiertos por escrito y si fuera necesario señalar la *no aplicabilidad* o la *no pertinencia* de los indicadores propuestos. Los consultados no tuvieron acceso a los valores de ponderación asignados en el modelo que se estaba validando.

Se han obtenido 28 respuestas válidas que corresponden a una población conformada por: a) 11 investigadores de universidades españolas y del Reino Unido no expertos y vinculados a las áreas de educación, economía, ciencias políticas, historia, derecho, biblioteconomía, periodismo y psicología; b) 17 expertos de los cuáles: 9 forman paneles de evaluación para las áreas de CSH de la FECYT (Fundación Española para la Ciencia y la Tecnología), 4 participan de la RESUP (*Projet de Réseau d'Étude et de Recherche sur l'Enseignement Supérieur-Francia*); 1 pertenece al TASTI (*Research Group for Science, Technology and Innovation Studies – Finlandia*), 2 son especialistas en educación superior de la Universidad de Bath (Reino Unido) y, por último, un especialista en filosofía de la ciencia de la Universidad Autónoma de Barcelona (Cataluña).

La consulta ha tenido un carácter exploratorio. No se trata de una muestra representativa en relación con la pertenencia disciplinaria de los consultados, pero aún así el número de expertos que han respondido es importante (17) si se tiene en cuenta que han sido seleccionados en

²³ Un grupo de indicadores está conformado por todos los indicadores en los que se desagrega un indicador o factor del nivel de desagregación anterior.

función de su experiencia en evaluación²⁴. De este modo, la técnica aplicada y el tamaño y la composición de la población que ha respondido al cuestionario permiten extraer conclusiones y realizar análisis cualitativamente significativos a partir de las respuestas obtenidas.

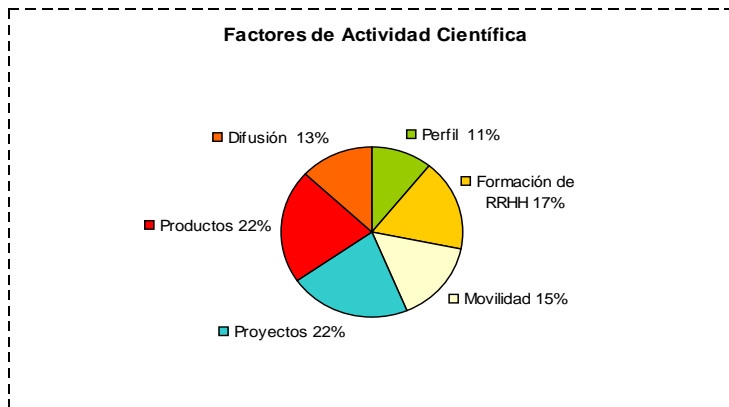
A continuación se analizan los resultados obtenidos para el nivel principal de indicadores que son los 6 factores que componen la actividad científica de acuerdo a la forma en que se la ha definido para el cálculo del IDAC. Se toma como referencia el valor de ponderación que se ha asignado a cada factor en el modelo del IDAC, y se presentan para cada caso: a) todos los valores de ponderación asignados por los consultados, b) la frecuencia de respuestas que se han obtenido para cada valor y c) el porcentaje que representan las frecuencias sobre el total de respuestas. Además se ha realizado el cálculo de la media, la mediana y el desvío estándar que se analizan junto con los resultados obtenidos.

Factores de Actividad Científica

En el Gráfico 14 se presentan los seis factores de actividad científica que componen el cuerpo del modelo de evaluación por indicadores y del modelo de cálculo del índice de actividad científica (IDAC). Los valores de ponderación que se les ha asignado expresan la contribución relativa de

²⁴ Cuando se utiliza la técnica de evaluación por pares académicos el criterio principal para la selección de los expertos es su perfil académico profesional y su experiencia en la temática a evaluar, esto significa que la representatividad de la muestra se basa en criterios cualitativos más que cuantitativos. La evaluación basada en el juicio de pares corresponde al modelo de investigación académica disciplinaria; la interpretación de los resultados de este ejercicio de validación debe encuadrarse en este contexto. Teniendo en cuenta que el consenso académico no es el único marco de relaciones sociales en que se constituye el conocimiento, y asumiendo una visión más interactiva de la ciencia, si se quisiera realmente ponderar incorporando la valoración y las prioridades sociales respecto de los factores de actividad científica, debería complementarse este estudio con otros procedimientos que sometan la validación del modelo al juicio de otros agentes sociales.

cada uno de los factores a la actividad científica de un investigador, o de una unidad de investigación considerada globalmente.



Análisis de resultados

Factor 1: Perfil del PDI

Tabla 1: Medidas estadísticas Factor 1	
1 - Ponderación del modelo	11%
2 - Media	14,96 %
3 - Mediana	10
4 - Desvío estándar	9,92

Tabla 2: Resultados obtenidos para el Factor Perfil

Ponderaciones asignadas	Cantidad de respuestas	Porcentaje de respuestas	Porcentaje acumulado
5%	4	14%	14%
6%	1	4%	18%
8%	1	4%	21%
10%	9	32%	54%
15%	4	14%	68%
20%	5	18%	86%
25%	1	4%	89%
30%	2	7%	96%
50%	1	4%	100%
Total	28	100%	

Con respecto al factor de actividad científica "Perfil del PDI" la media de las ponderaciones asignadas (14,96) supera en casi un 5% a la ponderación propuesta en el modelo del IDAC (11%). Si bien el mayor porcentaje de respuestas 32,14% corresponde a la asignación de 10% como valor de ponderación (la más cercana a la asignada en el modelo del IDAC), se debe tener en cuenta que el desvío estándar es significativo 9.92 y que existe una diferencia de casi 5% entre los valores de la media y la mediana. El segundo lugar en el orden de las ponderaciones asignadas

corresponde al valor de ponderación 20% con 17,86 % de las respuestas y el tercer lugar a los valores de ponderación 5% y 15% con un 14% de respuestas cada uno. De acuerdo a los datos obtenidos y a la aplicación de medidas estadísticas, es posible concluir que existe un bajo nivel de consenso respecto de la ponderación que debe asignarse a este factor para la evaluación de la actividad científica.

Factor 2. Formación de recursos humanos.

Tabla 3: Medidas estadísticas Factor 2	
1 - Ponderación del modelo	17%
2 - Media	11,29%
3 - Mediana	10
4 - Desvío estándar	5,52

Tabla 4: Resultados obtenidos para el Factor Recursos Humanos

Ponderaciones asignadas	Cantidad de respuestas	Porcentaje de respuestas	Porcentaje acumulado
0%	2	7%	7%
5%	4	14%	21%
8%	1	4%	25%
10%	9	32%	57%
15%	8	29%	86%
18%	1	4%	89%
20%	3	11%	100
Total	28	100	

En este caso tanto la media 11,29% como la mediana 10% se distancian del valor de ponderación propuesto en el modelo por 6 y 7 puntos respectivamente. Aunque sólo representan un 7% de las respuestas, es notable la existencia de casos en que se ha asignado ponderación nula a este factor de actividad científica. Si se agregan con los que han asignado el menor valor 5%, se concluye que el 21% de los consultados otorgan un peso muy débil a la formación de recursos humanos como factor de evaluación de la actividad científica. Sin embargo, el desvío estándar no es alto y si se suman los porcentajes, puede observarse que el 65% de las respuestas se concentran en un rango de valores comprendido entre 10% y el 18%, además de que el máximo valor propuesto es 20%. Se podría concluir que, para evaluar la actividad científica de las CSH, es aceptable asignar al factor *Formación de Recursos Humanos* una ponderación intermedia entre 10% y 18% .

Factor 3 Movilidad

1 - Ponderación del modelo	15%
2 - Media	12,79%
3 - Mediana	12,5%
4 - Desvío estándar	6,05

Ponderaciones asignadas	Cantidad de respuestas	Porcentaje de respuestas	Porcentaje acumulado
0%	1	4%	4%
5%	4	14%	18%
8%	1	4%	21%
10%	8	29%	50%
15%	7	25%	75%
20%	6	21%	96%
25%	1	4%	100%
Total	28	100%	

Los datos indican que la valoración de los consultados sobre el peso del factor *Movilidad* se acerca bastante a la ponderación que se ha propuesto en el modelo del IDAC. La media y la mediana casi no difieren (12,79% y 12,50% respectivamente) y el desvío estándar no es muy alto, lo que indica que no hay una importante dispersión respecto del valor medio asignado por los consultados. Sumando los porcentajes puede verse que el 54% de las respuestas se acumulan entre los valores 10% y 15% (29% y 25% respectivamente de porcentaje de respuestas). Por otra parte, la amplitud del rango de las valoraciones propuestas es menos amplia ya que varía entre 0% y 25%. De modo que, de acuerdo con esta consulta, sería aceptable asignar a la movilidad un valor de ponderación intermedio entre 10% y 15% para evaluar la actividad científica de las CSH.

Factor 4: Proyectos

Tabla 7: Medidas estadísticas Factor 4	
1 - Ponderación del modelo	22%
2 - Media	18,36%
3 - Mediana	15%
4 - Desvío estándar	10,96

En el caso del factor *Proyectos*, los resultados obtenidos difieren en forma significativa del valor de ponderación propuesto en el modelo del IDAC. La media se acerca un poco más 18%, pero la mediana indica que el 50% de los casos han asignado un valor de ponderación igual o menor al 15%. Por otra parte, el desvío estándar es alto, llegando casi al 11%. El 81% de los casos se acumulan en el rango que abarca el 10% y el 20%, pero se trata de una importante amplitud de rango. Al mismo tiempo, aunque son pocos

casos, considerados en forma agregada resulta que el 11% de los consultados han otorgado valores de ponderación muy altos al factor *Proyectos* 40% y 50% . Del mismo modo que en el caso del factor *Perfil* es posible afirmar que existe un muy bajo consenso con respecto al valor de ponderación que debe asignarse al factor *Proyectos* para evaluar la actividad científica de las CSH.

Factor 5: Productos

Tabla 9: Medidas estadísticas Factor 5	
1 - Ponderación del modelo	22%
2 - Media	29,54%
3 - Mediana	25%
4 - Desvío estándar	14,05

Tabla 10: Resultados obtenidos para el Factor Productos			
Ponderaciones asignadas	Cantidad de respuestas	Porcentaje de respuestas	Porcentaje acumulado
10%	3	11%	11%
15%	1	4%	14%
20%	7	25%	39%

22%	1	4%	43%
25%	3	11%	54%
30%	2	7%	61%
35%	3	11%	71%
40%	4	14%	86%
50%	2	7%	93%
60%	2	7%	100%
Total	28	100%	

El factor *Productos* muestra ser el más polémico a la hora de asignar un valor de ponderación para evaluar la actividad científica de las CSH. De hecho, el debate más importante en el campo de la evaluación científica se sitúa en este dominio. Se trata del factor con la mayor amplitud de rango de los valores propuestos ya que varían entre 10% y 60%. Por otra parte, el desvío estándar también es muy alto 14,05% por lo que, según esta consulta, queda expresado que no hay consenso en la valoración de este factor para evaluar la actividad científica de las CSH.

Si bien la media (29,54) reflejaría que existe una valoración alta respecto de este factor como medida de evaluación de la actividad científica (supera casi en un 8% al valor de ponderación propuesto en el modelo, 22%), el hecho de que la amplitud de rango y el desvío estándar sean tan importantes, desmerece el peso interpretativo de la media. De todos modos, el valor de ponderación "20%" acumula el mayor porcentaje de respuestas (25%), que resulta aún más significativo si se tiene en cuenta la dispersión de las respuestas obtenidas. Se puede destacar que este valor de ponderación se acerca bastante al propuesto en el modelo (22%),

aunque la principal conclusión con respecto a este factor es la falta de consenso para establecer un criterio de ponderación del factor *Productos*.

Factor 6: Difusión

Tabla 12: Resultados obtenidos para el Factor Difusión

Ponderaciones asignadas	Cantidad de respuestas	Porcentaje de respuestas	Porcentaje acumulado
0%	2	7%	7%
5%	2	7%	14%
8%	1	4%	18%
10%	9	32%	50%
15%	6	21%	71%
18%	1	4%	75%
20%	6	21%	96%
30%	1	4%	100%
Total	28	100%	

Tabla 11: Medidas estadísticas Factor 6	
1 - Ponderación del modelo	13%
2 - Media	13,07%
3 - Mediana	12,5%
4 - Desvío estándar	6,68

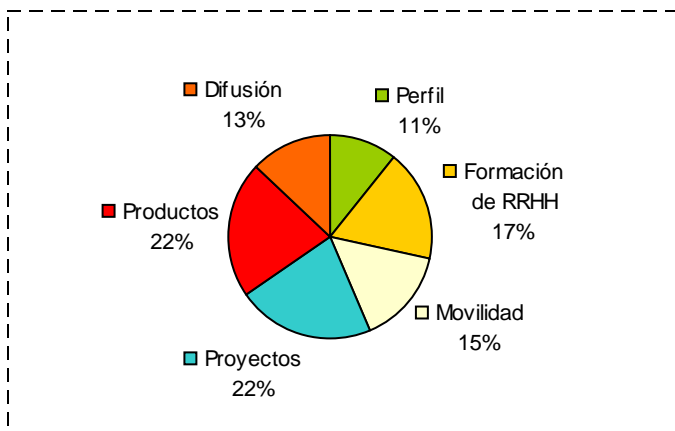
En el caso del factor *Difusión* la media y la mediana se acercan mucho al valor de ponderación asignado en el modelo del IDAC. El desvío estándar no es muy alto y la amplitud del rango de respuestas es moderada. Si se agregan, puede verse que el 53% de las respuestas se acumulan en los valores 10% y 15%, y que los porcentajes de respuestas en el resto de las categorías de ponderaciones asignadas no superan el 7%. El desequilibrio está dado por la acumulación del 21% de respuestas en el valor de ponderación 20%. A partir de estos resultados podría pensarse que el valor asignado en el modelo (13%) goza de un consenso significativo para ser aplicado a la evaluación de la actividad científica de las CSH, aunque podría ajustarse hacia un valor un poco mayor para dar respuesta a esta acumulación de frecuencias en el valor 20%.

Síntesis de resultados

A continuación se presentan en forma comparada el Gráfico 15, que representa el modelo que se ha sometido a validación y muestra los valores de ponderación que se han asignado a los 6 factores de actividad científica para el cálculo del IDAC, y el Gráfico 16, que representa las medias de los valores de ponderación asignados por los expertos e investigadores consultados para cada uno de los 6 factores de actividad científica.

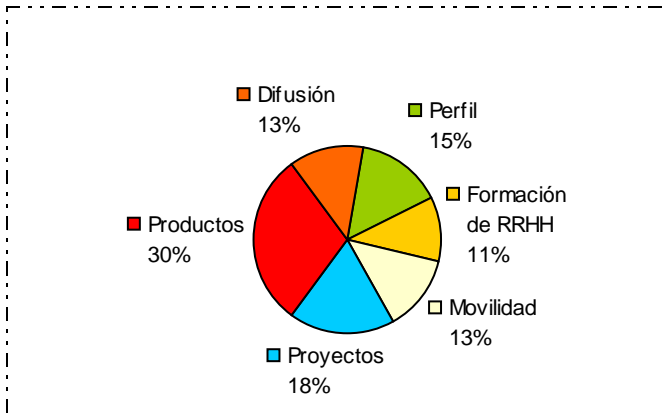
El análisis que se propone en función de las medias obtenidas en la consulta aporta nuevos elementos. Sin embargo, para mayor riqueza y precisión en las decisiones metodológicas que puedan tomarse, el mismo debe ser complementado con los análisis expuestos con anterioridad y que reflejan los resultados que se han obtenido para cada factor donde las medias son relativizadas en función de las frecuencias acumuladas y el desvío estándar.

Ponderación de factores del modelo de cálculo del IDAC



Ponderación de Factores según consulta de validación

Orden de factores y diferencia porcentual interna en cada modelo					
Modelo IDAC	Ponderación	Diferencia porcentual	Resultados de validación	Ponderación	Diferencia Porcentual
1º Proyectos /Productos	22%	+5%	1º Productos	30%	+12%
2º Formación de RRHH	17%	+2%	2º Proyectos	18%	+3%
3º Movilidad	15%	+2%	3º Perfil	15%	+2%
4º Difusión	13%	+2%	4º Movilidad/Difusión	13%	+2%
5º Perfil	11%		5º Formación de RRHH	11%	



En el caso de los factores *Perfil* y *Productos*, las medias de los valores de ponderación asignados por los consultados superan a los valores asignados en el modelo del IDAC siendo más significativa esta diferencia en el caso del factor *Productos*. La ponderación del factor *Difusión* coincide en ambos casos y para los factores *movilidad*, *formación de recursos humanos* y *proyectos* se les ha asignado valores de ponderación más bajos que los del modelo, donde la diferencia en el caso de *Formación de RRHH* es la más notoria.

Por otra parte, si se ordenan los factores según el orden de importancia asignada y se calcula la diferencia porcentual de los puntajes asignados con respecto al factor inmediatamente anterior, se obtiene que las principales diferencias que se han planteado entre el modelo del IDAC y el esquema que surge de la validación son: a) la ordenación de los factores especialmente en el caso de *Perfil* y *Formación de recursos humanos*, y b) la notoria diferencia en la valoración porcentual asignada al factor *Productos* (30%). La diferencia porcentual con respecto a *Proyectos* (12%) señala un orden de prioridad que contrasta con las diferencias porcentuales planteadas en el modelo. Para el resto de los factores aunque varien los valores propuestos, las diferencias porcentuales se mantienen en el mismo orden, entre 2% y 3%.

3.1 CONCLUSIONES DE LA VALIDACIÓN

Si bien la muestra de expertos e investigadores consultados no es representativa en términos de disciplinas de CSH, es posible extraer conclusiones generales de este estudio que permitan revisar y analizar con mayor profundidad los valores de ponderación que se han propuesto para el cálculo del Índice Directo de Actividad Científica. Sería por tanto necesario realizar estudios específicos que permitan revisar los indicadores del modelo de evaluación y ajustar las ponderaciones del IDAC según las particularidades que presenta la actividad científica en las distintas disciplinas de CSH.

La consulta sobre el valor de ponderación de los indicadores ha aportado elementos interesantes en función de dos ejes principales: a) analizar y revisar el valor de ponderación que se ha asignado en el modelo a cada uno de los 6 factores que componen la actividad científica, y b) analizar y visualizar el grado de consenso que existe entre expertos e investigadores de CSH en relación con el peso que se debe asignar a cada uno de los factores para evaluar la actividad científica de las CSH.

De acuerdo a los resultados obtenidos, el factor *Productos* muestra ser el más polémico a la hora de asignar un valor de ponderación para evaluar la actividad científica. De hecho, el debate más importante en el campo de la evaluación científica se sitúa en este dominio. También en el caso del factor *Perfil* aunque la media se aproxima al valor propuesto, ha habido valoraciones altas por parte de algunos consultados. Ambos factores representan los aspectos en los que tradicionalmente se ha focalizado la evaluación científica. La propuesta metodológica que se ha diseñado en el marco del proyecto de investigación, tanto para la evaluación de la actividad científica como para la ponderación de los factores en el IDAC, asume la existencia de este debate y pretende contrarrestar la habitual focalización en los productos de la actividad científica a la hora de evaluarla. Para ello, tal como puede observarse en el Gráfico 3 (pp.51), se propone un modelo que incorpora otros factores y que distribuye de una forma menos disonante los valores de ponderación de dichos factores. Con esta propuesta se pretende dotar de una mayor complejidad al concepto de actividad científica y dar valor y estatus a un conjunto más amplio de actividades y productos que los que tradicionalmente se han valorado. La sobrevaloración del *Perfil* conlleva una concepción un tanto más

individualista de la actividad científica. En el caso de los *Productos*, si bien se admite su importancia, se le asigna el mismo valor de ponderación que a *Proyectos* como una forma de valorizar los procesos y, al mismo tiempo, dar cuenta de que hay muchos aspectos de los procesos que se llevan a cabo en la práctica científica que no son plausibles de sistematización en publicaciones indizadas especialmente en el caso de las CSH.

Finalmente, este ejercicio de validación ha permitido reflexionar en torno a propuestas concretas de valorización de la actividad científica para el caso de las CSH. Se trata de una temática que requiere continuar con la reflexión, al mismo tiempo que con la acumulación de casuística y resultados que permitan estabilizar criterios y estándares de referencia para la medición y evaluación de las disciplinas específicas de CSH.

CONSIDERACIONES FINALES

El punto de partida de este trabajo ha sido considerar que la evaluación de la actividad científica es una actividad que crecientemente compromete a las instituciones y a los grupos de investigación bajo un número más diverso de modalidades, instancias y entornos. La ampliación de las misiones y de los contextos de interacción de las instituciones científicas, ha ampliado al mismo tiempo el campo y los públicos frente a los cuales éstas deben rendir cuentas. Ello conlleva una mayor complejidad, debido a la diversidad de los criterios que participan en la valorización de la producción científica. Por otra parte, las exigencias actuales de los sistemas de Ciencia y Tecnología suponen alcanzar estándares de calidad en un contexto donde la competitividad necesita combinarse con la integración, la colaboración y la coordinación a fin de garantizar la circulación de conocimiento, investigadores y tecnologías.

La propuesta de evaluación que se ha desarrollado en el proyecto en el cual se inscribe el desarrollo de este índice asume que la triangulación de metodologías es un requisito ineludible, si se tiene en cuenta la diversidad de exigencias y problemáticas con las que se enfrenta la producción científica en la actualidad. Combinar metodologías, técnicas y modelos permite complementar enfoques y tipos de resultados como ocurre en el modelo aquí se propone y que incluye enfoques como evaluación por

indicadores, evaluación por pares, análisis de pertinencia y consenso institucional.

De este modo, el calificativo de *integral* utilizado en este trabajo se ha referido a varios planos: en primer lugar, al modo en que se ha definido y operacionalizado el concepto de actividad científica; en segundo lugar, a la propuesta de incorporar elementos relacionales en el tratamiento de la información y en el estudio de la calidad científica y, finalmente, se ha referido también al proceso evaluativo que se propone: un Programa de Evaluación diseñado con el múltiple objetivo de captar, evaluar y apoyar las dinámicas que organizan patrones de investigación.

La propuesta de evaluación integral ha sido desarrollada teniendo como objeto a las CSH y, en particular, teniendo presente la necesidad que existe en estos campos disciplinarios de identificar estándares de producción, patrones de funcionamiento y dinámicas de vinculación. En este sentido, la heterogeneidad y la escasa estandarización de criterios de evaluación que caracterizan la situación actual de las CSH, más que como un obstáculo, en este caso se han tomado como una ventana de oportunidad para reflexionar y elaborar propuestas que, por su carácter exploratorio y por el conjunto de aproximaciones que integra, no deja de tener interés para ser adaptada a la evaluación de otros campos disciplinarios. Hoy en día la importancia de la vinculación con el contexto y la conectividad social, no son una preocupación exclusiva de las CSH que han sido más clásicamente asociadas a la idea de *pertinencia*. La fórmula *calidad-productividad* resulta cada vez más insuficiente en términos evaluativos y por supuesto también, en términos explicativos.

El concepto de *calidad relacional* tal como hemos querido presentarlo aquí, incluye como dimensiones intrínsecas a la calidad científica un conjunto amplio de actividades pero principalmente, incluye en su definición a los modos de organización y entre ellos, los modos de establecer vinculaciones efectivas con una gama más amplia de esferas y agentes sociales en el marco de los procesos de producción científica. Con el fin de incorporar aspectos relacionales en el estudio de la calidad científica, el modelo de evaluación integral que aquí se ha presentado avanza en tres planos: a) en el plano *metodológico* introduce mecanismos de consulta y validación del propio modelo que incluyen criterios de valoración de la actividad científica de distintos agentes sociales (expertos y usuarios de los

sistemas de evaluación), b) en el plano *analítico*, incorpora un conjunto amplio de dimensiones e indicadores para dar cuenta de la integralidad de la actividad científica, y propone un tratamiento relacional de la información que resulta más exigente en términos explicativos que su tratamiento como indicadores aislados; y finalmente, c) en el plano de la *pertinencia*, donde a partir de la consulta a investigadores de CSH y a potenciales usuarios de la investigación, se busca testar su posible receptividad, expectativas y necesidades con el fin de orientar estrategias y ofrecer elementos que permitan estructurar la demanda y la oferta de conocimientos científicos.

En definitiva, la hipótesis que se mantiene detrás de esta propuesta metodológica integral, es que la producción, distribución y uso de conocimiento científico alcanzará buenos estándares de calidad, en la medida en que se garanticen buenas condiciones relacionales para estos procesos, especialmente en un contexto en que la búsqueda de la excelencia académica está asociada a lograr altos grados de integración y conectividad del sistema científico entre sí y con sus entornos. En este sentido, se espera que el esfuerzo realizado por diseñar una propuesta integradora de enfoques y aproximaciones, resulte una aportación que de lugar a intercambios productivos en el entorno diverso de agentes que se ven involucrados en procesos de valoración de la actividad científica.

REFERENCIAS BIBLIOGRÁFICAS

ARCHAMBAULT, E. , VIGNOLA-GAGNE, E., COTE, G., LARIVIERE, V. , GINGRAS, Y. (2006). "Benchmarking scientific output in the social sciences and humanities : The limits of existing databases". *Scientometrics* 68 (3) 329-342.

BARRENECHEA, J., CASTRO SPILA, J. E IBARRA, A. (2008) Calidad relacional y evaluación integral de la actividad científica en ciencias sociales y humanidades: propuesta metodológica e indicadores. Estudios en Ciencia, Tecnología e Innovación nº 3. Cátedra Sánchez Mazas. Servicio Editorial de la UPV/EHU. País Vasco.

- BECHER, T. (1993). "Las disciplinas y la identidad de los académicos". *Pensamiento Universitario*, 1, 56-77.
- BENCE, V. & OPPENHEIM, C. (2004) The role of academic journal publications in the UK Research Assessment Exercise . Learned Publishing, 17, 53-68.
- BUELA-CASAL, G., GUTIÉRREZ-MARTÍNEZ, O., BERMÚDEZ-SÁNCHEZ, M. P. & VADILLO-MUÑOZ, O. (2007) Comparative study of international academic rankings of universities. *Scientometrics*. 71, 349-365.
- COOK, THOMAS D. (1985) 'Postpositivist Critical Multiplism', in R.L. Shotland and M. M. Mark (eds) *Social Science and Social Policy*, pp. 21–62. Thousand Oaks, CA: Sage.
- CRUZ, I. R. (2008) Challenging ISI Thomson Scientific's Journal Citation Reports: Deconstructing "Objective," "Impact," and "Global". *portal: Libraries and the Academy*, 8, 7-13.
- ECCLES, C. (2002) The Use of University Rankings in the United Kingdom. *Higher Education in Europe*, 27, 423-432.
- FEDERKEIL, G. (2002) Some Aspects of Ranking Methodology—The CHE-Ranking of German Universities. *Higher Education in Europe*, 27, 389-397.
- GREENE, J. BENJAMIN, L. Y GOODYEAR, L. (2001). The Merits of Mixing Methods in Evaluation. *Evaluation*, 7 (1), pp.25-44.
- GREENE, J., CARACELLI, V. Y GRAHAM, W. (1989) 'Toward a Conceptual Framework for Mixed-Method Evaluation Designs', *Educational Evaluation and Policy Analysis* 11: 255–74. Greenwood, Davydd and Morten Levin
- IBARRA, A., CASTRO, J. Y BARRENECHEA, J. (2006) "Acta del taller: Indicadores para la evaluación de la actividad de investigación en Ciencias Sociales y Humanidades" en Ibarra, A., Castro, J. y Barrenechea, J. *La Evaluación de la Actividad Científica en Ciencias Sociales y Humanidades*.

Estudios en Ciencia, Tecnología e Innovación nº 1. Cátedra Sanchez Mazas. Servicio Editorial de la UPV/EHU. País Vasco

KYVIK, S. (2003). "Changing trends in publishing behaviour among university faculty, 1980-2000". *Scientometrics*, 58 (1) 35-48.

VAN LEEUWEN, T. (2006). The application of bibliometric analyses in the evaluation of social science research. Who benefits from it, and why it is still feasible. *Scientometrics*, 66 (1), pp.133-154.

LIU, N. C. & CHENG, Y. (2005) The Academic Ranking of World Universities. *Higher Education in Europe*, 30, 127-136.

MERISOTIS, J. P. (2002) On the Ranking of Higher Education Institutions. *Higher Education in Europe*, 27, 361-363.

MOYA-ANEGÓN, F., CHINCHILLA-RODRÍGUEZ, Z., CORERA-ÁLVAREZ, E., MUÑOZ-FERNÁNDEZ, F. J. & NAVARRETE-CORTÉS, J. (2004) Indicadores bibliométricos de la actividad científica española:(ISI, Web of science, 1998-2002)