



Gestión y Análisis de Políticas Públicas

ISSN: 1134-6035

ISSN: 1989-8991

Instituto Nacional de Administración Pública (INAP)

Cruz-Castro, Laura

>Diferencias y sesgos de género en la financiación de la investigación: un enfoque dinámico 1

Gestión y Análisis de Políticas Públicas, núm. 26, 2021, pp. 6-19

Instituto Nacional de Administración Pública (INAP)

DOI: <https://doi.org/10.24965/gapp.i26.10909>

Disponible en: <https://www.redalyc.org/articulo.oa?id=281567964001>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en [redalyc.org](https://www.redalyc.org)

UAEM [redalyc.org](https://www.redalyc.org)

Sistema de Información Científica Redalyc

Red de Revistas Científicas de América Latina y el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto



GAPP, número 26, julio de 2021
Sección: ARTÍCULOS
Recibido: 15-02-2021
Modificado: 19-04-2021
Aceptado: 21-04-2021
Publicado: 01-07-2021
DOI: <https://doi.org/10.24965/gapp.i26.10909>
Páginas: 6-19

Diferencias y sesgos de género en la financiación de la investigación: un enfoque dinámico¹

Gender differences and gender bias in research funding: a dynamic approach

Laura Cruz-Castro

Instituto de Políticas y Bienes Públicos – CSIC (España)

ORCID: <https://orcid.org/0000-0002-9786-6251>

laura.cruz@csic.es

NOTA BIOGRÁFICA

Laura Cruz Castro es Investigadora Científica en el Instituto de Políticas y Bienes Públicos del CSIC. Es doctora en Sociología por la Universidad Autónoma de Madrid. Su área de especialización son las políticas de ciencia e investigación, las carreras académicas y las organizaciones de I+D. Ha dirigido y participado en numerosos proyectos de investigación financiados competitivamente y publicado en las revistas internacionales más relevantes de su campo.

RESUMEN

En general, las científicas consiguen menos fondos para investigar y tienen, con respecto a los solicitantes, menores tasas de éxito que los hombres. Este artículo revisa críticamente la literatura sobre las disparidades y sesgos de género en la financiación de la investigación, reconociendo la diversidad de hallazgos, y poniendo el énfasis en la necesidad de clarificación conceptual y rigor metodológico. Se adopta un enfoque dinámico con la premisa de que las diferencias en las tasas de solicitud y éxito en las convocatorias de proyectos de investigación se ven afectadas por diferencias previas y acumulativas y a su vez influyen en la carrera posteriormente. Se revisa también el papel de las preferencias y de las expectativas individuales a la hora de competir, así como el proceso de evaluación y cómo las características de los evaluadores y de los procedimientos de las agencias pueden activar o desactivar mecanismos de sesgo.

PALABRAS CLAVE

Financiación de la investigación; Diferencias de género; Sesgo; Evaluación por pares.

ABSTRACT

In general, women scientists get less funding for research and, compared to applicants, have lower success rates than men. This article critically revises the literature on gender disparities and gender bias in research funding, acknowledging diverse findings and emphasizing the need for conceptual clarification and methodological rigor. A dynamic approach is adopted with the premise that differences in application and success rates in research grants are affected by previous and cumulative differences, and in turn influence the career afterwards. It also reviews the role of individual preferences and expectations in differential application, as well as the evaluation process and how the characteristics of evaluators and procedures can activate or mitigate bias mechanisms.

KEYWORDS

Research funding; Gender differences; Bias; Peer review.

¹ La autora agradece los comentarios y sugerencias de dos evaluadores anónimos de la Revista a una versión previa.

SUMARIO

INTRODUCCIÓN. 1. UNA REFLEXIÓN CRÍTICA SOBRE ARGUMENTOS COMUNES. 2. EL EFECTO DE LAS DIFERENCIAS PREVIAS EN LA ESTRUCTURA DE OPORTUNIDADES. 3. PARTICIPACIÓN EN LAS CONVOCATORIAS: PREFERENCIAS FRENTE A EXPECTATIVAS. 4. LOS PROCESOS DE EVALUACIÓN. 4.1. QUÉ Y CÓMO SE EVALÚA. 4.2. QUIÉN EVALÚA Y QUIÉN DECIDE. 5. EL IMPACTO DE LA FINANCIACIÓN EN LA PRODUCTIVIDAD Y EN LAS CARRERAS. 6. ALGUNOS EJEMPLOS DE POLÍTICAS DE LAS AGENCIAS DE INVESTIGACIÓN. 7. REFLEXIÓN FINAL SOBRE EL CAMBIO INSTITUCIONAL.

INTRODUCCIÓN

La igualdad de oportunidades es un valor y la importancia de dicha igualdad en el acceso a los recursos entre hombres y mujeres en las sociedades democráticas está fuera de toda duda; es más, la igualdad de oportunidades no trata solo de valores políticos; en ciencia, las becas de investigación en general y las subvenciones para proyectos de investigación en particular son un recurso vital para la carrera científica, y algunos estudios sostienen que los equipos de investigación con diversidad de género producen mejor ciencia (Fortunato et al., 2018). Sin entrar en el debate sobre si la diversidad es un valor en sí mismo, lo cierto es que, en esta área, la diversidad produce valor; de modo que la igualdad de oportunidades no solo es justa, también es rentable.

Este artículo recoge algunos avances analíticos del proyecto GRANteD *Grant allocation disparities from a gender perspective*² (Cruz-Castro y Sanz-Menéndez, 2020), un proyecto comparativo cuyos principales objetivos son: (i) clarificar el concepto de sesgo de género frente al de brecha de género en la concesión de proyectos de investigación, (ii) identificar la existencia de sesgos y de desigualdades en las distintas fases que van desde la decisión de presentarse a una convocatoria hasta la decisión final de concesión, (iii) investigar las causas de los distintos tipos de sesgo y (iv) determinar el impacto de la financiación en las carreras de investigadoras e investigadores. Se trata de un área donde faltan datos, pero sobre todo falta teoría y un modelo explicativo general que vaya más allá de los estudios de caso y descriptivos. La evidencia que tenemos sobre cómo funcionan los programas de financiación, las diferencias por género y la existencia de discriminación está muy fragmentada, y basada en estudios locales que en su mayoría no son extrapolables a otros contextos y que, en general, hace que no podamos hablar de resultados concluyentes sobre las explicaciones causales.

Es necesario mucho esfuerzo aún en el terreno de la clarificación de los conceptos, y el análisis crítico de algunos de los argumentos comúnmente aceptados, y que a veces sirven de base a las políticas públicas. Gran parte de los resultados contradictorios se deben a que en realidad se están analizando fenómenos distintos con las mismas medidas, o el mismo fenómeno con medidas muy dispares; por ejemplo, términos como sesgo, discriminación, prejuicio, parcialidad, favoritismo, injusticia, desigualdad, inequidad, diferencia etc. se utilizan a menudo como conceptos equivalentes e intercambiables, cuando en realidad no lo son y representan procesos distintos.

Lo cierto es que el sistema de ciencia se ha construido sobre el modelo normativo mertoniano (CUDOS) (Merton, 1942) que está asociado a la idea de equidad (justicia en relación con el mérito) más que a la igualdad.

En este artículo se revisa la literatura desde la perspectiva de los estudios de ciencia, con el foco puesto en la necesidad de un enfoque dinámico que preste atención a los mecanismos y a los procesos acumulativos. Este estudio no es un meta-análisis ni pretende ser una revisión exhaustiva. Responde a una selección y como tal, es parcial. Existe una literatura desde los estudios de género distinta de la perspectiva que aquí se adopta, al igual que existe un área de estudios políticos de la ciencia que tampoco se aborda.

1. UNA REFLEXIÓN CRÍTICA SOBRE ARGUMENTOS COMUNES

Con frecuencia, la confusión conceptual conduce a argumentos causales imprecisos; por ejemplo, es posible que toda discriminación produzca desigualdad, pero no toda desigualdad o diferencia es fruto de la

² Programa H2020 Swifts. Proyecto financiado por el Grant Agreement N.º 824574, al que se agradece financiación.

discriminación. Con todo, es un hecho que en general las científicas consiguen menos fondos para investigar y tienen, respecto a los solicitantes, menores tasas de éxito que los hombres, también en España (UMyC – Ministerio de Ciencia e Innovación, 2011). Por ejemplo, según datos recientes de la Agencia Española de Investigación referidos a 2018, el porcentaje de proyectos de investigación presentados por una mujer como IP fue del 32,4 % frente al 67,6 % de proyectos con IP masculino; de las solicitudes aprobadas, el 29,5 % estaban lideradas por mujeres frente al 70,5 % lideradas por hombres³. Sin embargo, el papel real que la discriminación y los sesgos de los evaluadores tienen en ese tipo de resultados es algo sobre lo cual la evidencia disponible es diversa y no siempre concluyente.

Demostrar que hay discriminación en la asignación de proyectos es una tarea muy compleja porque las personas a las que comparamos, en este caso los investigadores e investigadoras solicitantes, tendrían que ser iguales o similares en todo, menos en aquello sobre lo que pensamos produce el sesgo (su género); sabemos que en muchos de los casos eso no es así porque hombres y mujeres no llegan en las mismas o similares circunstancias al momento de la solicitud. Las desigualdades de partida, incluidas aquellas que tienen que ver con los méritos que se evalúan, son fundamentales en la explicación.

Se trata de partir de la necesidad de distinguir los resultados (las tasas de éxito) de los mecanismos (que tienen lugar en el proceso de evaluación en paneles de evaluadores) y de no atribuir las diferencias en cualquier tipo de resultado (las tasas de éxito) a un mecanismo en particular (por ejemplo, la discriminación por género por parte de los que evalúan), sin antes controlar las posibles diferencias de partida entre hombres y mujeres, y otras explicaciones alternativas. Esto, lejos de ser sencillo, plantea unas dificultades metodológicas considerables, entre ellas, la dificultad de observar los procesos sociales en entornos reales.

Un ejemplo de explicación simplista, pero no por ello ausente en el discurso sería el siguiente: dada una cierta distribución de solicitantes entre hombres y mujeres, observamos una tasa de éxito que no es proporcional a las mujeres que se han presentado y concluimos que la diferencia entre investigadoras solicitantes y financiadas se explica porque ha habido un sesgo de género en la evaluación en contra de las mujeres que ha favorecido a los hombres. Este tipo de argumentación hay que evitarla, al menos por dos razones: la primera es que es mala ciencia social; la segunda es que la crítica sin fundamento sólido, aunque durante un tiempo contribuya a visibilizar un problema, a la larga, produce entre las científicas un efecto de desánimo que hace que se presenten menos porque dan por descontado un trato discriminatorio.

Por tanto, una cuestión metodológica muy importante es que no podemos partir de un resultado (las tasas de éxito), e inferir de él directamente un proceso (la discriminación). De hacerlo, estaríamos cayendo en lo que Cole (1979) denominó «residualismo ingenuo» que tantas veces está en la base de las correlaciones simples. En otras palabras, la premisa debe ser que la discriminación es sólo una de las causas posibles de las diferencias observadas en los resultados.

Por añadidura, tasas de éxito iguales para hombres y mujeres no garantizan que el proceso de evaluación y asignación no haya sido discriminatorio. Si ese fuese el caso, los sesgos de género tendrían fácil solución desde las políticas públicas y desde las agencias de financiación: solo habría que definir una línea base, por ejemplo, la proporción de solicitantes por género, y repartir la financiación disponible proporcionalmente. ¿Sería esto aceptable? ¿Lo sería para la política? ¿Para los financiadores? ¿Para la comunidad científica? Para algunos, sería aceptable si asumimos que, por definición, ambos grupos tienen la misma calidad, o que, la igualdad de resultados como valor es socialmente preferible a otros valores, como la equidad. En el ámbito de la ciencia, ambas son ideas controvertidas.

Un paso en la dirección correcta es establecer si hay una relación entre el éxito relativo en la obtención de financiación para investigar y los méritos de los candidatos en el momento de la solicitud. Si, desde la evidencia empírica, somos capaces de establecer esa relación, al menos estaremos controlando una de las explicaciones posibles que rivalizan con la existencia de sesgos: si, a igualdad de méritos, las tasas de éxito son distintas a lo esperable en función de las solicitudes, el residual se lo atribuimos al sesgo, la parcialidad o la discriminación; este residualismo ya no es tan ingenuo y se conoce como «residualismo sofisticado» (Cole y Fiorentine, 1991). La revisión de la literatura revela que es aquí donde se quedan la mayoría de investigaciones en esta área. Es necesario plantearse un modelo más ambicioso, de carácter secuencial, que tenga

³ Datos referidos a la convocatoria 2018 del Programa de Generación de Conocimiento (PGC). En el programa de Retos de Investigación (RTI), el porcentaje de solicitudes lideradas por mujeres fue del 38,9 % y el porcentaje solicitudes aprobadas con una mujer como Investigadora Principal fue el 35,9 % (Agencia Estatal de Investigación, 2020). En la convocatoria de contratos Ramón y Cajal de 2019, el porcentaje de solicitudes presentadas por hombres fue 60 % frente al 40 % de solicitudes de mujeres; las tasas de éxito, sin embargo, fueron similares: 11 % y 11,5 % respectivamente (Agencia Estatal de Investigación, 2021).

en cuenta todas las fases del proceso desde las desigualdades de partida, pasando por la solicitud hasta los efectos posteriores en las carreras (Cruz-Castro y Sanz-Menéndez, 2020).

A continuación, se enumeran algunos avances analíticos que merece la pena resaltar y que ponen de manifiesto la necesidad de estudiar los procesos y su secuencia, para luego profundizar en las secciones que siguen en lo que la literatura revisada nos revela de cada uno de ellos.

En primer lugar, es importante reconocer que existen desigualdades de partida, que no tienen que ver con las publicaciones y los índices de impacto, desigualdades que en algunos casos son estructurales. La productividad no es solo una característica de las personas, también es un reflejo de las posiciones que ocupan estas en la estructura social de la academia y del acceso a los recursos que esas posiciones permiten (Xie y Shauman, 1998). Por ejemplo, hay más investigadoras en universidades menos prestigiosas, con contratos temporales o a tiempo parcial, o en rangos académicos inferiores, o en ciertos campos de la ciencia; es lo que conocemos como «segregación». La segregación social es un concepto fundamental en las ciencias sociales e ignorarlo lleva con frecuencia a errar el diseño de las políticas.

En segundo lugar, hay desigualdades en la fase de solicitud, y por diversas razones, con frecuencia las mujeres se presentan menos a las convocatorias de financiación de la investigación, y esas razones hay que investigarlas. Es lo que conocemos como «autoselección».

En tercer lugar, puede haber parcialidad/discriminación dentro de los paneles de evaluación, y esta falta de imparcialidad también puede tener orígenes diversos (género, institución, proximidad cognitiva); la fuente de la discriminación es importante para diseñar las políticas destinadas a reducirla. Son lo que conocemos como «sesgos».

Por último, la financiación de proyectos tiene efectos en las carreras de hombres y mujeres, y que esos efectos varían en función de las prácticas de las organizaciones que emplean a los investigadores. Es lo que conocemos como el «ciclo de la credibilidad científica» (Latour y Woolgar, 1979; Rip, 1994).

2. EL EFECTO DE LAS DIFERENCIAS PREVIAS EN LA ESTRUCTURA DE OPORTUNIDADES

Las diferencias en las tasas de solicitud y éxito por género en las convocatorias de proyectos y becas de investigación pueden verse afectadas por desigualdades previas. La literatura en torno a este tema es ingente y se centra mayoritariamente en la menor participación de niñas y jóvenes en las ramas educativas de algunas ciencias y en las ingenierías desde etapas tempranas hasta la universidad. Esta explicación, conocida como el modelo del *pipeline* pone el énfasis fundamentalmente en las fugas (*leaks*) que llevan a una reducción de las tasas de participación femenina en cada etapa, y las políticas basadas en este tipo de evidencia se orientan a aumentar dichas tasas.

Aunque ampliamente presente en el discurso de las políticas públicas, algunos autores han formulado críticas a ese modelo; por ejemplo, Xie y Shauman (2003) consideran que es demasiado simple y no captura la complejidad de los resultados educativos y de carrera; además, el enfoque se basa en exceso en la idea del desarrollo, de modo que cualquier trayectoria que no supone un progreso de una etapa a otra se considera un fracaso, y presta poca atención a las interdependencias entre las decisiones educativas y otros acontecimientos a la largo de la vida. Sin embargo, a mi modo de ver, la crítica más interesante se sitúa en el terreno metodológico, y se resume en la imposibilidad de analizar adecuadamente las dinámicas que se producen en el tiempo con datos transversales, y la necesidad de llevar a cabo análisis de cohortes y estudios longitudinales.

Con todo, parece innegable que una parte importante de las mujeres que muestran interés por las carreras científicas se pierde en cada transición educativa y profesional en muchos sistemas. Hace décadas, era común atribuir las diferencias en las tasas de participación femenina en carreras de ciencias a una participación previa menor en secundaria en dichas ramas, junto con unos peores resultados en las calificaciones en estas materias. Estas explicaciones basadas en la capacidad han sido bastante criticadas, y confrontadas por perspectivas que ponen el énfasis en las expectativas, y en cómo los estereotipos socioculturales estructuran la desigualdad, y generan sesgos implícitos en las preferencias de niños y jóvenes de ambos sexos que conducen posteriormente a distribuciones segregadas por género (Ridgeway, 2014).

Desde el punto de vista del diseño de las políticas, reconocer que las mujeres no están representadas por igual en los distintos campos científicos (en algunos son mayoría) es esencial para entender que, en programas de financiación de la investigación que se organizan por áreas (la mayoría, por otra parte), las disparidades de género en la distribución de los resultados tienen en parte su origen en dinámicas de segregación previa.

Algo similar ocurre con la segmentación ocupacional en etapas posteriores de la carrera y la menor representación de las mujeres en posiciones altas de las escalas académicas; el conocido como «techo de cristal» (Rosser, 2004) es un tema recurrente en el que también falta evidencia de carácter longitudinal. En ese sentido, algunos estudios han demostrado, para Estados Unidos, que, con el tiempo, no hay diferencias significativas en la tasa de éxito a la hora de acceder a la titularidad entre hombres y mujeres (National Research Council, 2010); y efectivamente, la clave está en el factor tiempo, porque, otros estudios muestran que, además de tardar más en acceder a la permanencia (*tenure*), las mujeres son más proclives a abandonar la ciencia y la academia antes de que esto ocurra (Goulden et al., 2011).

Además de estas diferencias previas de carácter estructural, la causa más frecuente en el análisis de las diferencias por género en la consecución de proyectos tiene que ver con diferencias en los méritos, y dentro de éstos, en la productividad científica. Esta literatura es muy extensa. La afirmación empírica tradicional ha sido que las mujeres publican menos que los hombres y que sus trabajos tienen un impacto menor (Cole y Zuckerman, 1984; Long, 1992; Sandström y van den Besselaar, 2016, entre otros). Otras investigaciones apuntan a un cierre progresivo de dicha brecha (Xie y Shauman, 1998; Thelwall, 2018) y al hecho de que, en algunas instituciones, la productividad aumenta con el rango tanto para hombres como para mujeres (Mauleón et al., 2008). El problema principal de la mayoría de estudios que analizan la relación entre género y productividad es que no precisan los mecanismos causales.

Las explicaciones de la menor productividad de las mujeres pueden clasificarse de modo muy amplio en dos categorías relacionadas con lo que Sonnert y Holton (1995) denominaron el modelo de la diferencia y el modelo del déficit. El primero sostiene que las mujeres se comportan de modo diferente porque tienen motivaciones distintas y distinto nivel de compromiso con la carrera científica; por el contrario, el modelo del déficit sostiene que las mujeres tienen las mismas aspiraciones y objetivos que los hombres pero que reciben un trato diferente; así, su menor productividad sería el resultado de menos oportunidades, obstáculos en la carrera, y de problemas para la obtención de recursos y para colaborar con otros investigadores. Ambos tipos de factores son compatibles y pueden reforzarse entre sí (Mairesse y Pezonni, 2015). La productividad científica individual y la financiación tienen una relación circular y acumulativa.

Por último, es importante señalar un problema metodológico que afecta a mucha de la evidencia empírica sobre la menor productividad científica de las mujeres, me refiero al abandono de la carrera. Muchos de los trabajos en este campo utilizan muestras que únicamente incluyen individuos activos en la carrera científica y académica, incurriendo en lo que estadísticamente se conoce como censura de datos por la derecha. Preston (2004) ya señaló importancia de estudiar cohortes completas y no solo a las científicas que permanecían en la academia.

3. PARTICIPACIÓN EN LAS CONVOCATORIAS: PREFERENCIAS FRENTE A EXPECTATIVAS

Como ya se ha mencionado, con frecuencia, el debate sobre las diferencias de financiación por género se centra en las disparidades en las tasas de éxito en las convocatorias. Sin embargo, para poder analizar correctamente la cuestión del éxito, hay que entender previamente la decisión de presentarse o no. Que las mujeres se presentan menos es un hecho, y lo es incluso en campos científicos donde hay paridad de género (Ley y Hamilton, 2008); es decir, las tasas de solicitud no se corresponden proporcionalmente con el conjunto potencial de solicitantes (Comisión Europea, 2009). Muchos son los factores que influyen; algunos son estructurales, y tienen que ver con la segregación ocupacional previa; en otras palabras, las mujeres están menos representadas en aquellas posiciones o categorías desde las que es más probable (o incluso posible) presentarse; pero las variables familiares también importan (Blake y La Valle, 2000).

Para estudiar las diferencias en participación, Stephan y El-Ganainy (2007) proponen distinguir entre explicaciones basadas en el contexto, la mayoría estructurales, de aquellas basadas en factores que afectan por una parte a la oferta (actitudes hacia la competición y el riesgo, o las preferencias sobre conciliación de vida laboral y familiar), y por otra a la demanda (papel de las redes de colaboración, preferencias de evaluadores o empleadores, entre otros).

Entre los factores estructurales más estrechamente relacionados con el diseño de los programas de financiación de la investigación se encuentran los criterios de elegibilidad. Por ejemplo, requisitos asociados al tipo de contrato o a la categoría profesional pueden producir desventajas para las mujeres a la hora de presentarse dado que, en muchos países, éstas se encuentran sobre-representadas en la contratación temporal, a tiempo parcial o en categorías de menor rango. En ese sentido, un hallazgo interesante de una

encuesta llevada a cabo en el Reino Unido (Blake y La Valle, 2000) era que las investigadoras tenían más probabilidad que los investigadores de incluir en el presupuesto de los proyectos su propio salario, lo que sugería una mayor representación de las primeras en categorías no permanentes. En suma, una parte de las diferencias de género en las tasas de solicitud está arraigada en las prácticas de empleo de las instituciones (Comisión Europea, 2009).

Entre los factores relacionados con la oferta existe un debate, en especial en economía, acerca de si las mujeres tienen más aversión al riesgo y actitudes más desfavorables a la competición que los hombres. La evidencia, sin embargo, dista de ser concluyente, y algunas revisiones críticas de la literatura han identificado importantes sesgos metodológicos en la mayoría de los trabajos (Nelson, 2014), y ponen en cuestión la existencia de diferencias intrínsecas, señalando que conductas que pueden parecer aversión al riesgo son en realidad respuestas a incentivos y recompensas diferentes (Fernández-Mateo y Kaplan, 2018). Estos mismos autores, además, advierten sobre la confusión entre efectos de oferta y demanda, y sobre la conveniencia de preguntarse si son las mujeres las que eligen no aspirar a las posiciones más altas de la jerarquía o a trabajos en los segmentos mejor pagados, o si, por el contrario, ciertas posiciones o trabajos no resultan atractivos para las mujeres debido a experiencias pasadas de fracaso en la contratación o en la promoción.

4. LOS PROCESOS DE EVALUACIÓN

4.1. Qué y cómo se evalúa

Cuando hablamos de evaluación de la investigación y de los investigadores, una distinción interesante es aquella entre mérito y valor (Antal et al., 2015). Lincoln y Guba (1980) argumentan que mientras que el mérito es intrínseco al individuo y puede evaluarse con relación a una serie de criterios previamente acordados, el valor es extrínseco y para determinarlo hay que confrontar la aportación del individuo con una serie de requerimientos externos. En las organizaciones de investigación, el ejemplo paradigmático de evaluación en la que se estiman no solo la calidad y el mérito sino también el valor del candidato para el centro o el departamento, es aquella que regula el acceso a puestos permanentes. Las agencias de financiación, por lo general, ponen el énfasis en los méritos, pero cuando evalúan el ajuste entre un proyecto o un individuo y la institución de acogida, están considerando aspectos de valor.

Como se dijo en la introducción, cualquier análisis serio sobre la existencia de sesgos de género en la evaluación debe tener en cuenta los méritos. En un estudio ya clásico de los años 90, Wennerås y Wold (1997) analizaron el proceso de evaluación por pares del Consejo de Investigación de Medicina sueco para la concesión de becas postdoctorales y mostraron que la meritocracia no era el estándar; por el contrario, las relaciones sociales entre candidatos y miembros de los paneles jugaba un papel significativo, y a la vez, las mujeres necesitaban más del doble de publicaciones que los hombres para obtener puntuaciones similares. Una reproducción del estudio diez años después confirmó el papel de la proximidad institucional pero no el sesgo de género (Sandström y Hällsten, 2008). Este ejemplo ilustra el problema principal de la investigación en este campo: mientras que unos estudios muestran hallazgos en favor de la existencia de sesgos, otros no; además, la magnitud de los efectos varía sustancialmente de un estudio a otro.

Otro tema relevante es si los propios instrumentos de medida de los méritos pueden introducir sesgos en la evaluación. Los que abogan por el uso de criterios bibliométricos alegan que estas métricas son neutrales con respecto al género. Otros, más críticos, sostienen que la bibliometría, cuando se usa para evaluar individuos y no instituciones, puede contribuir a perpetuar desigualdades previas, y que, en cualquier caso, los indicadores bibliométricos son *proxies* del mérito, a veces descontextualizados (Nielsen, 2018). Un ejemplo muy citado es del índice h, que está sesgado hacia los investigadores que publican en sub-campos con altas tasas de citación, que se correlaciona con el número de publicaciones y por lo tanto con la edad académica, y además privilegia la publicación con múltiples coautores, algo que coloca en desventaja a las mujeres (Abramo et al., 2013). Aquí de nuevo hay que reconocer que la evidencia no es concluyente y hay resultados en un sentido y en otro; en un estudio sobre la agencia italiana de evaluación de universidades e institutos de investigación, Jappelli et al. (2017), si bien hallaron diferencias significativas por género, no encontraron que el análisis bibliométrico como método de evaluación perjudicara a las mujeres comparado con la evaluación por pares, sino al contrario.

Un área muy prometedora en el campo de la evaluación de la investigación es el análisis lingüístico de los informes de los evaluadores. En un trabajo reciente sobre informes de evaluación de solicitudes de

renovación de proyectos del NIH (*National Institute of Health*), Magua et al. (2017) mostraban que mientras los investigadores eran descritos como «líderes» o «pioneros» en su campo, con una investigación «altamente innovadora» y «altamente significativa», las investigadoras eran caracterizadas como «expertas» en su campo y parte de «entornos de investigación excelentes». Aunque es sabido que la formalización de los criterios de evaluación y de los procedimientos afectan a la calidad y a la transparencia de los procesos de revisión, algunos autores han alertado sobre la necesidad de que las categorías de los dictámenes eviten cierto tipo de carga semántica y limiten el margen interpretativo de los revisores; este suele ser el caso de categorías del tipo «aceptable» (en inglés *good enough*).

No solo las palabras, también los números pueden tener una carga simbólica. Las calificaciones cuantitativas son ubicuas en las organizaciones modernas y el tipo de escalas de calificación puede tener efectos. Un experimento reciente llevado a cabo por Rivera y Tilcsik (2019) en una universidad norteamericana, revela que el uso de escalas de 1 a 6 para evaluar la calidad docente, favorece a las profesoras más que el uso de escalas de 1 a 10; los autores sugieren que el número de puntos en una escala importa en la medida en que los estereotipos de brillantez están asociados al significado cultural de ciertos números.

Las condiciones organizativas también importan. En una revisión de hace unos años, Heilman (2001) identificó cuatro tipos de factores organizativos que perjudicaban la evaluación del desempeño de las mujeres. En primer lugar, la ambigüedad en los criterios de evaluación: cuanto más vagos y genéricos, mayor facilidad para que la información pueda presentarse de modo que encaje con ideas preconcebidas; en segundo lugar, la falta de estructura en los procesos de evaluación y toma de decisiones; en tercer lugar, la ambigüedad con respecto al origen de los resultados: hay factores organizativos que diluyen la contribución de los individuos a trabajos colectivos; por ejemplo, poner demasiado énfasis en los equipos puede tener consecuencias negativas si el trabajo en grupo favorece atribuciones colectivas del mérito que limitan la visibilidad de las contribuciones individuales; en cuarto lugar, la ambigüedad acerca de las causas de los logros anteriores; en este sentido, la discriminación positiva o el establecimiento de cuotas pueden tener efectos no deseados relacionados con la percepción de haber tenido un trato favorable a lo largo de la carrera (Kravitz y Platania, 1993). Van den Brink y Stobbe (2014) afirman que conviene re-conceptualizar este dilema organizativo como una paradoja en la que la «ayuda» (su percepción) se convierte en una dificultad.

4.2. Quién evalúa y quién decide

El modo predominante de evaluación de la investigación para su financiación es la evaluación por pares, *peer review*, que puede tener lugar o no en el contexto de paneles o comités. A pesar de haber recibido críticas relativas a su fiabilidad, equidad y valor predictivo (Bornman, 2011), este método sigue siendo el estándar general en las agencias de financiación (Chubin y Hackett, 1990).

Los paneles de evaluación son un entorno social y como tal, sujeto a dinámicas complejas. Desde el punto de vista empírico, la importancia de estudiar las características de los evaluadores radica en los efectos de interacción. No solo los atributos de los evaluados sino también los de los evaluadores pueden dar lugar a sesgos y por ello es relevante incluirlos en los análisis. De hecho, las características de los evaluadores y las dinámicas dentro de los paneles (género tanto del evaluador como del evaluado, conflicto de interés, presión para la conformidad y el acuerdo, etc.) importan en tal grado que pueden cambiar la calificación de un proyecto de financiable a no financiable (Tamblyn et al., 2018).

En el sistema científico, la revisión por pares cumple la función de «guardián» o de filtro de entrada. Se ha discutido mucho sobre el impacto de la presencia de mujeres (o más bien de más mujeres) en los paneles de evaluación y selección. Una composición más equilibrada podría implicar una mayor conciencia de género y más atención a los sesgos implícitos (Van den Brink et al., 2006), pero la evidencia sobre las prácticas y dinámicas de los paneles de evaluación de agencias de financiación es muy limitada. En comparación, hay muchos más trabajos sobre comités de selección y promoción académica, que, en general, muestran unas tasas de éxito menores para las mujeres (Van den Brink, 2010); de ellos se pueden extraer algunas conclusiones de interés, entre ellas, la más importante es que a veces se confunde el efecto del género de los evaluadores con el efecto de las redes o conexiones sociales o institucionales, que suelen operar en favor de los hombres, por la menor densidad de las redes de las mujeres (Checchi et al., 2019). Este tipo de interacciones puede ser la causa que subyace a la falta de acuerdo respecto a si realmente la presencia de mujeres en los comités de evaluación tiene un impacto positivo en las tasas de éxito femeninas (Bagues et al., 2017).

Una de las fuentes principales de sesgo de género en las evaluaciones son los estereotipos. Los estereotipos pueden considerarse un «atajo» cognitivo, por el cual es más probable que los individuos con-

sideren más fiable y veraz la información que se ajusta a sus expectativas estereotipadas que las que no (Ellemers, 2018). La distinción entre estereotipos descriptivos y normativos es interesante. Según Heilman (2012) los primeros consisten en unas expectativas menores con respecto al desempeño de las mujeres, que derivan de la percepción de que hay un desajuste entre las características de las mujeres y los atributos que se suponen necesarios para el éxito. Desde el punto de vista de las políticas, lo importante es que este tipo de percepciones pueden cambiar con la información adecuada. No ocurre lo mismo con los estereotipos normativos o prescriptivos, que son ideas preconcebidas sobre cómo «deberían» ser las mujeres.

La mayor parte de la evidencia disponible en el tema de los estereotipos es de tipo experimental, y lo cierto es que, en su mayoría, no encuentra diferencias entre hombres y mujeres: tener estereotipos de género parece algo transversal (Moss-Racusin et al., 2012). No obstante, hay que tener cierta cautela con esta metodología porque los experimentos en esta área suelen ser muy específicos en cuanto al contexto, y, por tanto, poco generalizables. Entre las explicaciones teóricas que intentan dar cuenta de la activación de estereotipos en los procesos de evaluación está la denominada «teoría de la congruencia de roles» (Knobloch-Westernwick y Glynn, 2013), de acuerdo a la cual, las creencias sobre los científicos y los hombres se parecen, mientras que las creencias sobre los científicos y las mujeres no; esta incongruencia afectaría negativamente a la evaluación de las mujeres en entornos científicos.

Una comparación incluso general de la literatura que afirma que existen sesgos de género en la evaluación de la investigación, frente a la que sostiene que dicho fenómeno es limitado y/o es cada vez menor, muestra claramente que la primera es mayoritaria. Esta constatación no es óbice para que desde la investigación haya que estar muy atentos a las explicaciones rivales que compiten con el sesgo a la hora de dar cuenta de las diferencias de resultados por género, y no inferir procesos desde los resultados solo a partir de datos referidos a las distribuciones.

En 2007, Ceci y Williams (2007) publicaron una revisión controvertida de la evidencia empírica disponible hasta entonces acerca de la existencia de sesgos de género en la evaluación de proyectos, de artículos y en la contratación académica, que servía de base a la mayoría de las políticas al respecto. Se trata de una revisión bastante amplia de más de dos décadas de datos (fundamentalmente de Estados Unidos), y a partir de la cual concluían que gran parte de la evidencia ya no era válida. A pesar de la controversia, una conclusión interesante era que las disparidades de género en los resultados resultan de una desigualdad de recursos que derivan de elecciones, libres o constreñidas, y que es a esas elecciones, y a que estén mejor informadas a través de la educación, donde hay que dirigir las intervenciones políticas y los recursos.

La idea de que las desigualdades de género en las carreras científicas se explican por las diferencias en los méritos y por las elecciones educativas y profesionales previas tiene su atractivo analítico, pero resiste mal la confrontación con el hecho de que cuanto más subimos en la jerarquía académica, más difícil es distinguir el mérito y las decisiones de carrera de otros factores que impiden el avance de las mujeres. En un artículo reciente, Treviño et al. (2018) analizan un caso interesante, el de las cátedras distinguidas (*named professorships*), un tipo de nombramiento que implica una distinción honorífica y un complemento salarial, en el área de economía de la empresa de las universidades americanas. Sus hallazgos revelan que, tras controlar por el mérito y otros factores relevantes, el nombramiento de mujeres para estas cátedras era menos probable que el de los hombres; las mujeres, cuando accedían, lo hacían con mayores méritos relativos y, además, el efecto adverso del género se acentuaba si el candidato finalmente elegido era interno a la propia universidad.

La causa más probable de la heterogeneidad en la literatura sobre la existencia de sesgos y sus causas radica en la falta de definiciones comunes, muestras, métodos, etc. Y existe un riesgo cierto de que la falta de validez y/o fiabilidad lleve a algunos a poner en duda o incluso a negar todo el fenómeno en su conjunto, cuando la realidad es que diferentes estudios se centran en diferentes situaciones y contextos. No obstante, es importante no equiparar las críticas a estudios poco rigurosos o convincentes sobre la existencia de sesgos de género, con la negación de la existencia de desigualdades. Como ya he insistido, se trata de procesos distintos.

5. EL IMPACTO DE LA FINANCIACIÓN EN LA PRODUCTIVIDAD Y EN LAS CARRERAS

No es posible hablar del impacto de la financiación y los proyectos en las carreras de modo general. El avance en las carreras científicas está fuertemente afectado por el contexto institucional y organizativo, y no todas las organizaciones otorgan el mismo valor a la capacidad de obtener financiación para la investigación.

El modelo de financiación de la investigación predominante a nivel nacional también importa. Con todo, parece evidente que la financiación competitiva de proyectos es cada vez más importante para las carreras, y se usa cada vez más como indicador de calidad tanto a nivel individual como institucional. Algunos tipos de becas, proyectos y subvenciones han adquirido incluso un valor reputacional que va más allá del económico.

Hay muchos más trabajos sobre el impacto de la financiación en la productividad científica que en las carreras. La razón más obvia es que existen métricas al uso que facilitan lo primero, mientras que la investigación sobre trayectorias requiere de estudios mucho más complejos y costosos. En general, los estudios bibliométricos revelan efectos positivos, aunque con frecuencia presentan limitaciones metodológicas derivadas de la falta de grupos de control equivalentes (aquellos que no obtuvieron financiación o que se quedaron justo en el umbral), o de la falta de estudios longitudinales que puedan servir de contrafactual. Algunos estudios muestran que, en términos de productividad, las diferencias entre los financiados y los que no lo fueron no son significativas (Bornmann et al., 2008; Van den Beesselaar y Leydesdorff, 2009) aunque hay que señalar que estos estudios utilizan muestras de individuos que llegaron hasta las fases finales del proceso de selección, es decir, investigadores de alta calidad, en cualquier caso. Cuando las tasas de éxito son muy bajas (programas altamente competitivos) puede que este tipo de evaluación de impacto no sea muy pertinente.

Estudiar la productividad posterior de los investigadores que obtienen financiación sirve como una suerte de control de calidad del proceso de evaluación. Es lo que se conoce como validez predictiva de la selección. No obstante, inferir una evaluación de calidad a partir de la productividad posterior de los financiados adolece de problemas de circularidad, por la imposibilidad de descartar que el aumento de las publicaciones de los que tuvieron éxito en la convocatoria se deba precisamente a la financiación, y no tanto a su calidad previa. Hay que ser por tanto cauteloso con el tema de la atribución causal y la endogeneidad, y, cuando se hace evaluación de impacto, contar con los grupos de control adecuados.

Por otra parte, existe el problema añadido de las muchas variables que concurren en la explicación de la producción científica: edad académica, género, rango, redes de colaboración, institución, localización geográfica, entre otras. En un trabajo reciente sobre una cohorte de investigadores financiados por el ERC (*European Research Council*) a través de *starting* y *advanced grants* en el área de ciencias de la vida, Pina et al. (2019) muestran que el género no influye en las diferencias de productividad antes y después de la financiación, aunque los hombres con *starting grants* tienen más publicaciones tanto antes como después; por otra parte, parece que el momento de la carrera en la que se obtiene la financiación también influye, ya que el aumento de la productividad es mayor en los beneficiarios de *starting grants* que en los *advanced*; un último hallazgo interesante es que los beneficiarios de *starting grants* que trabajan en países con menos desarrollo científico tienen menos publicaciones y redes de colaboración menos densas, y que esta variable de localización influye más que el género. Como ya se ha señalado, un problema metodológico en este tipo de estudios es que, al analizar sólo individuos que sí obtuvieron la financiación con un tipo de diseño de «antes y después», no se cuenta con grupos de control.

La evidencia del impacto de los proyectos y becas de investigación en las carreras es limitada, aunque en general, muestra efectos positivos. Un estudio de la Comisión Europea sobre los becarios del programa Marie Curie, que utilizaba un grupo de control, revelaba que los beneficiarios de estas becas habían obtenido un puesto académico permanente y lideraban equipos con más frecuencia, y que estas ayudas tenían un impacto positivo en la reducción de la brecha de género en cuanto a las probabilidades de acceder a la titularidad o a la cátedra, el número de publicaciones y el acceso posterior a financiación internacional (Comisión Europea, 2014). En este caso, el problema metodológico principal es la dificultad de analizar a los individuos que abandonan la carrera, y entre los cuales hay razones para pensar hay más mujeres que hombres. Esto es especialmente importante en tiempos de crisis (Cruz-Castro y Sanz-Menéndez, 2016), cuando la presión académica competitiva se hace mayor y puede provocar efectos negativos en las oportunidades de las mujeres.

6. ALGUNOS EJEMPLOS DE POLÍTICAS DE LAS AGENCIAS DE INVESTIGACIÓN

En primer lugar, hay que distinguir entre políticas que inciden en el lado de la oferta, es decir las que intentan aumentar por ejemplo el número de mujeres que se presentan a las convocatorias (programas de mentores, formación), o en el de la demanda (las que intentan incidir en las organizaciones, su transparencia, la formalización de criterios que reduzcan los sesgos y la discriminación, la formación de evaluadores

etc.). Es importante identificar con precisión si el problema que se quiere abordar desde la política es de un tipo u otro, ya que a veces el diagnóstico los confunde.

Algunos trabajos llaman a cierta cautela a la hora de confiar demasiado en el impacto de la formación de evaluadores, intervenciones destinadas a que los evaluadores tomen conciencia de las propias actitudes discriminatorias, sobre el cambio de esas actitudes; es posible, incluso, que mostrar evidencia sobre la existencia de sesgos genere tanto resultados positivos como reacciones justificativas (Moss-Racussin et al., 2015).

Hay trabajos muy interesantes desde la economía que señalan la necesidad de identificar la fuente de la discriminación, y distinguir entre aquella que se basa en creencias y la que se basa en preferencias. Con diseños experimentales, hay autores que han mostrado como, cuando las creencias iniciales de los evaluadores favorecen a los hombres, las mujeres necesitan demostrar más para ser evaluadas igual; este tipo de discriminación puede mitigarse con información adicional sobre la calidad de las candidatas que desmonte esas creencias o prejuicios; por el contrario, si la causa de la discriminación es una preferencia negativa, hacer más visible la calidad de las mujeres no tiene efectos (Bohren et al., 2019). Siendo más optimistas y a la vez más realistas, centrarse en crear estructuras institucionales y normas puede ser mejor que intentar cambiar a los individuos.

Los motivos de las personas son difíciles de investigar y de demostrar; las políticas pueden ser más eficaces si se centran en otro tipo de mecanismos, de carácter normativo y organizativo, de procedimiento, etc. La literatura nos señala algunas formas de reducir el sesgo de género derivado de la activación de estereotipos. La primera es evitar la formación de expectativas negativas a través de factores de contexto: la representación proporcional de las mujeres es uno de esos factores contextuales; el género se visibiliza si hay un grupo minoritario, y existe evidencia que muestra que, si se aumenta la proporción de mujeres en el conjunto de solicitantes, la evaluación de las mismas también mejora (Sackett et al., 1991). La segunda es reducir la ambigüedad en el proceso de evaluación proporcionando a los evaluadores la mayor y mejor información posible de los candidatos de manera que se reduzca la tendencia a «rellenar» huecos de información con ideas preconcebidas; para ello, es imprescindible que los criterios de evaluación sean claros y concretos, y puedan aplicarse de manera idéntica con independencia de a quien se esté evaluando (Heilman, 2012); esto incluye reglas claras para la atribución del mérito individual en el trabajo de grupo. Por último, las organizaciones pueden aumentar los incentivos de los evaluadores para ser justos a través de mecanismos de rendición de cuentas (Reskin, 2003).

También es común distinguir entre políticas que se orientan a los procesos de solicitud o evaluación, y políticas que se orientan directamente al resultado de las convocatorias. Este es un tema complejo porque, aunque la mayoría de las políticas públicas son del primer tipo, a veces, el discurso político y académico que suele acompañar los diagnósticos y las soluciones contienen referencias a lo segundo. Cuando las políticas tienen como objetivo directamente los resultados, existe un riesgo de discriminación inversa y de señalización adversa.

Otro tipo de políticas son las que se centran en el diseño y contenido de las convocatorias. Aquí, hay algunos ejemplos de diseños interesantes centrados en que el lenguaje con el que se describen los perfiles de los candidatos que se buscan sea lo más neutro posible. Rasgos tradicionalmente asociados a lo masculino que pueden cambiarse por otros más neutrales son: «potencial para el liderazgo» por «potencial para contribuir significativamente», «importancia» por «influencia», «innovación» por «originalidad», y «creatividad» por «inventiva» (Escobar Álvarez et al., 2019). Otro ejemplo interesante es incluir en las guías para la redacción de cartas de referencia, la recomendación de que se dichas cartas se centren en los méritos de los candidatos y se eviten referencias a circunstancias personales como el género, la edad, o a capacidades y roles ajenos al mundo profesional.

La evaluación ciega, en el sentido de no conocer la identidad y otros atributos del sujeto evaluado, ha sido el estándar tradicional en la evaluación de artículos de investigación para su publicación. Solo recientemente se ha empezado a experimentar con este tipo de enfoque en otros tipos de evaluaciones y los resultados son variados. Mientras que algunos estudios muestran que el anonimización sí reduce sesgos relacionados con características personales del evaluado como el género, la proximidad social, o la conformidad intelectual con el evaluador (Goldin y Rouse, 2000), otros, como el estudio de Kolev et al. (2019) sobre la evaluación de proyectos innovadores de la Fundación Gates, señalan que, a pesar de la evaluación ciega, los proyectos presentados por mujeres reciben menores puntuaciones. Curiosamente, en este último estudio, cuando se controla por una serie de medidas relativas al lenguaje de los títulos y las descripciones de los proyectos (palabras genéricas *versus* específicas), las diferencias desaparecen, lo que sugiere que, en ese caso, la brecha en las puntuaciones tiene que ver con diferencias en el estilo de comunicación.

Los dos tipos de políticas más comunes relativas al diseño de las convocatorias y al proceso de evaluación están sujetas a cierta controversia. Las primeras son las denominadas en inglés *clock stopping policies*, son aquellas que amplían los plazos a considerar para los solicitantes que han tenido hijos. La evidencia

ha demostrado que cuando estas políticas se aplican tanto a hombres como a mujeres, las brechas de solicitudes no se reducen, y que, cuando se aplican a plazos para solicitar la *tenure* en sistemas como el estadounidense, las diferencias por género pueden incluso aumentar (Antecol et al., 2018). Estos resultados ponen de manifiesto que este tipo de políticas neutrales no tienen en cuenta adecuadamente la realidad de las pérdidas de productividad asociadas la maternidad y la paternidad por separado.

Las segundas son las cuotas de género en la composición de los comités de evaluación y selección. Aunque en algunos países estas cuotas son obligatorias, la evidencia sobre sus efectos es diversa y en general no es concluyente. La Comisión Europea se comprometió a alcanzar un 40% de mujeres en las estructuras de asesoramiento de Horizonte 2020; sin embargo, algunas estadísticas ponen en duda su eficacia; por ejemplo, con datos agregados para 2008-2012, el *European Research Council* (ERC) no encontró correlación entre el número de mujeres en los paneles de evaluación y las tasas de éxito de las solicitantes (Vernos, 2013). Tampoco hay que subestimar la sobrecarga de trabajo de evaluación que la implementación de las cuotas puede suponer para las investigadoras, en especial en áreas donde están poco representadas, y que puede ir en detrimento del propio trabajo de investigación.

7. REFLEXIÓN FINAL SOBRE EL CAMBIO INSTITUCIONAL

Uno de mis objetivos en este artículo ha sido subrayar el hecho de algunas de las dinámicas que llevan a que las mujeres se presenten menos y obtengan menos proyectos de investigación son estructurales. Muchas de las diferencias entre hombres y mujeres a la hora de participar en las convocatorias tienen su origen en las prácticas de empleo de universidades y centros de investigación. Aunque tenga un componente derivado de la secuencia de desarrollo, la baja representación de las mujeres en las categorías académicas superiores es un problema que persiste en muchos países; la opacidad y discrecionalidad de los procesos de toma de decisiones a nivel local es otra traba para la implementación de políticas. Para algunos, la introducción de sistemas de habilitación y acreditación basados en indicadores bibliométricos ayuda a la transparencia (Marini y Meschitti, 2018).

Para otros, la evaluación basada en la revisión por pares tiene tantas limitaciones que sería mejor utilizar un sistema de lotería como mecanismo de asignación de la financiación para la investigación, una vez cumplidos unos niveles estándares de calidad (Roumbanis, 2019) y a la vez incrementaría la diversidad epistemológica, la equidad y la imparcialidad. No existe por el momento evidencia a favor ni en contra; lo que sin duda genera esta idea es una especie de derrota moral vinculada a dar por perdida la posibilidad de cambiar las organizaciones y las instituciones.

Desde una perspectiva más positiva, Steward y Valiant (2018) han esbozado algunos principios generales para crear estructuras favorecedoras del cambio. En primer lugar, la necesidad de centrarse en las políticas y en las prácticas; para estos autores, puede que la formación de evaluadores y el cambio de actitudes estén sobrevaloradas en comparación con el diseño de reglas y procedimientos que hagan difíciles conductas discriminatorias. En segundo lugar, la conveniencia de basar las políticas en el conocimiento experto y en la evidencia de lo que funciona y de lo que no. En tercer lugar, aunque sin duda hay casos de buenas prácticas de los que aprender, las políticas no pueden importarse sin más de otros contextos esperando que tendrán el mismo impacto; saber que una política ha funcionado en alguna parte es tan importante como preguntarse si funcionará aquí (Cartwright y Hardie, 2012). Por último, el principio de «cambia lo que puedas donde puedas» es una llamada al pragmatismo y un reconocimiento de que las organizaciones rara vez cambian de manera radical sino incremental y que, un poco de cambio, en la medida en que ponga en marcha un proceso, es mejor que la ausencia de cambio.

REFERENCIAS BIBLIOGRÁFICAS

- Abramo, G., D'Angelo, C. A. y Murgia, G. (2013). Gender differences in research collaboration. *Journal of Informetrics*, 7(4), 811-822. <https://doi.org/10.1016/j.joi.2013.07.002>
- Agencia Estatal de Investigación – AEI (2020). Proyectos de I+D de «Generación de Conocimiento». Proyectos de I+D+i «Retos Investigación». Convocatorias 2018. [informe de análisis]. https://www.ciencia.gob.es/portal/site/MI-CINN/menuitem.791459a43fdf738d70fd325001432ea0/?vgnextoid=1c85ed0682835610VgnVCM1000001d04140aRCRD&vgnextchannel=ede586fd6c544610VgnVCM1000001d04140aRCRD&vgnextfmt=formato2&id3=b257eacaf341f610VgnVCM1000001d04140a_____

- Agencia Estatal de Investigación (2021). Programa Ramón y Cajal para la contratación de investigadores de trayectoria destacada 2018/2019. Análisis [informe]. https://www.ciencia.gob.es/stfls/MICINN/AEI/ficheros/Informe_Programa_Ramon_y_Cajal_convocatorias_2018_19.pdf
- Antal, A. B., Hutter, M. y Stark, D. (2015). *Moments of Valuation: Exploring Sites of Dissonance*. Oxford University Press.
- Antecol, H., Bedard, K. y Stearns, J. (2018). Equal but Inequitable: Who Benefits from Gender-Neutral Tenure Clock Stopping Policies? *American Economic Review*, 108(9), 2.420-2.441. <https://doi.org/10.1257/aer.20160613>
- Bagues, M., Sylos-Labini, M. y Zinovyeva, N. (2017). Does the Gender Composition of Scientific Committees Matter? *American Economic Review*, 107(4), 1.207-1.238. <https://doi.org/10.1257/aer.20151211>
- Blake, M. y La Valle, I. (2000). *Who applies for research funding? Key factors shaping funding application behaviour among women and men in British higher education institutions* [informe]. National Centre for Social Research. <https://wellcomecollection.org/works/qctj6ypv>
- Bohren, J. A., Imas, A. y Rosenberg, M. (2019). The Dynamics of Discrimination: Theory and Evidence. *American Economic Review*, 109(10), 3.395-3.436. <https://doi.org/10.1257/aer.20171829>
- Bornmann, L. (2011). Scientific peer review. *Annual Review of Information Science and Technology*, 45(1), 199-246. <https://doi.org/10.1002/aris.2011.1440450112>
- Bornmann, L., Mutz, R. y Daniel, H.-D. (2008). How to detect indications of potential sources of bias in peer review: A generalized latent variable modeling approach exemplified by a gender study. *Journal of Informetrics*, 2(4), 280-287. <https://doi.org/10.1016/j.joi.2008.09.003>
- Cartwright, N. y Hardie, J. (2012). *Evidence-Based Policy: A Practical Guide to Doing It Better*. Oxford University Press.
- Ceci, S. J. y Williams, W. M. (eds.) (2007). *Why Aren't More Women in Science? Top Researchers Debate the Evidence* (1.ª ed.). American Psychological Association. <https://doi.org/10.1037/11546-000>
- Checchi, D., Cicognani, S. y Kulic, N. (2019). Gender Quotas or Girls' Networks? Evidence from an Italian Research Selection. *Work, Employment and Society*, 33(3), 462-482. <https://doi.org/10.1177/0950017018813071>
- Chubin, D. E. y Hackett, E. J. (1990). *Peerless Science: Peer Review and U. S. Science Policy*. SUNY series in Science, Technology, and Society. State University of New York Press.
- Cole, J. R. (1979). *Fair Science: Women in the Scientific Community*. The Free Press.
- Cole, J. y Zuckerman, H. (1984). The productivity puzzle: persistence and change in patterns of publication of men and women scientists. *Advances in Motivation and Achievement*, 2, 217-258.
- Cole, S. y Fiorentine, R. (1991). Discrimination against women in science: The confusion of outcome with process. En H. Zuckerman, J. R. Cole y J. T. Bruer (eds.), *The Outer Circle: Women in the Scientific Community* (pp. 205-226). Yale University Press.
- Cruz Castro, L. y Sanz Menéndez, L. (2020). *Grant Allocation Disparities from a Gender Perspective: Literature Review. Synthesis Report. GRANteD Project D.1.1*. Digital.CSIC. <https://doi.org/10.20350/digitalCSIC/10548>
- Cruz Castro, L. y Sanz Menéndez, L. (2016). The effects of the economic crisis on public research: Spanish budgetary policies and research organizations. *Technological Forecasting and Social Change*, 113, Part B, 157-167. <https://doi.org/10.1016/j.techfore.2015.08.001>
- Comisión Europea (ed.) (2009). *The gender challenge in research funding: Assessing the European national scenes* [informe]. Office for Official Publications of the European Communities. <https://www.genderportal.eu/resources/gender-challenge-research-funding-assessing-european-national-scenes>
- Comisión Europea (2014). *Marie Curie researchers and their long-term career development: A comparative study: final report* [informe]. Publications Office of the European Union. <https://publications.europa.eu/en/publication-detail/-/publication/9c2cb7d8-3773-430c-b1c0-db94ec421b01>
- Ellemers, N. (2018). Gender Stereotypes. *Annual Review of Psychology*, 69, 275-298. <https://doi.org/10.1146/annurev-psych-122216-011719>
- Escobar Alvarez, S. N., Jaggi, R., Abbuhl, S. B., Lee, C. J. y Myers, E. R. (2019). Promoting gender equity in grant making: what can a funder do? *The Lancet*, 393(10.171), e9-e11. [https://doi.org/10.1016/S0140-6736\(19\)30211-9](https://doi.org/10.1016/S0140-6736(19)30211-9)
- Fernandez-Mateo, I. y Kaplan, S. (2018). Gender and Organization Science: Introduction to a Virtual Special Issue. *Organization Science*, 29(6), 1.229-1.236. <https://doi.org/10.1287/orsc.2018.1249>
- Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B., Vespignani, A., Waltman, L., Wang, D. y Barabási, A.-L. (2018). Science of science. *Science*, 359(6.379), artículo eaao0185. <https://doi.org/10.1126/science.aao0185>
- Goldin, C. y Rouse, C. (2000). Orchestrating Impartiality: The Impact of «Blind» Auditions on Female Musicians. *American Economic Review*, 90(4), 715-741. <https://doi.org/10.1257/aer.90.4.715>
- Goulden, M., Mason, M. A. y Frasc, K. (2011). Keeping Women in the Science Pipeline. *The ANNALS of the American Academy of Political and Social Science*, 638(1), 141-162. <https://doi.org/10.1177/0002716211416925>
- Heilman, M. E. (2001). Description and Prescription: How Gender Stereotypes Prevent Women's Ascent Up the Organizational Ladder. *Journal of Social Issues*, 57(4), 657-674. <https://doi.org/10.1111/0022-4537.00234>
- Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, 32, 113-135. <https://doi.org/10.1016/j.riob.2012.11.003>
- Jappelli, T., Nappi, C. A. y Torrini, R. (2017). Gender effects in research evaluation. *Research Policy*, 46(5), 911-924. <https://doi.org/10.1016/j.respol.2017.03.002>

- Knobloch-Westerwick, S., Glynn, C. J. y Huge, M. (2013). The Matilda Effect in Science Communication: An Experiment on Gender Bias in Publication Quality Perceptions and Collaboration Interest. *Science Communication*, 35(5), 603-625. <https://doi.org/10.1177/1075547012472684>
- Kolev, J., Fuentes-Medel, Y. y Murray, F. (2019). *Is Blinded Review Enough? How Gendered Outcomes Arise Even Under Anonymous Evaluation* [NBER Working Paper Series, 25.759]. National Bureau of Economic Research. <https://doi.org/10.3386/w25759>
- Kravitz, D. A. y Platania, J. (1993). Attitudes and beliefs about affirmative action: Effects of target and of respondent sex and ethnicity. *Journal of Applied Psychology*, 78(6), 928-938. <https://doi.org/10.1037/0021-9010.78.6.928>
- Latour, B. y Woolgar, S. (1979). *Laboratory Life: The Construction of Scientific Facts* (2.ª ed.). Princeton University Press. <https://doi.org/10.2307/j.ctt32bbxc>
- Ley, T. J. y Hamilton, B. H. (2008). The Gender Gap in NIH Grant Applications. *Science*, 322(5.907), 1.472-1.474. <https://doi.org/10.1126/science.1165878>
- Lincoln, Y. S. y Guba, E. G. (1980). The Distinction Between Merit and Worth in Evaluation. *Educational Evaluation and Policy Analysis*, 2(4), 61-71. <https://doi.org/10.2307/1163674>
- Long, J. S. (1992). Measures of Sex Differences in Scientific Productivity. *Social Forces*, 71(1), 159-178. <https://doi.org/10.2307/2579971>
- Magua, W., Zhu, X., Bhattacharya, A., Filut, A., Potvien, A., Leatherberry, R., Lee, Y.-G., Jens, M., Malikireddy, D., Carnes, M. y Kaatz, A. (2017). Are Female Applicants Disadvantaged in National Institutes of Health Peer Review? Combining Algorithmic Text Mining and Qualitative Methods to Detect Evaluative Differences in R01 Reviewers' Critiques. *Journal of Women's Health*, 26(5), 560-570. <https://doi.org/10.1089/jwh.2016.6021>
- Mairesse, J. y Pezzoni, M. (2015). Does Gender Affect Scientific Productivity? A Critical Review of the Empirical Evidence and a Panel Data Econometric Analysis for French Physicists. *Revue Économique*, 66(1), 65-113. <https://doi.org/10.3917/reco.661.0065>
- Marini, G. y Meschitti, V. (2018). The trench warfare of gender discrimination: evidence from academic promotions to full professor in Italy. *Scientometrics*, 115(2), 989-1.006. <https://doi.org/10.1007/s11192-018-2696-8>
- Mauleón, E., Bordons, M. y Oppenheim, C. (2008). The effect of gender on research staff success in life sciences in the Spanish National Research Council. *Research Evaluation*, 17(3), 213-225. <https://doi.org/10.3152/095820208X331676>
- Merton, R. K. (1942). The normative structure of science. En *The Sociology of Science. Theoretical and Empirical Investigations* (pp. 267-278). The University of Chicago Press.
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J. y Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences – PNAS*, 109(41), 16.474-16.479. <https://doi.org/10.1073/pnas.1211286109>
- Moss-Racusin, C. A., Molenda, A. K. y Cramer, C. R. (2015). Can Evidence Impact Attitudes? Public Reactions to Evidence of Gender Bias in STEM Fields. *Psychology of Women Quarterly*, 39(2), 194-209. <https://doi.org/10.1177/0361684314565777>
- National Research Council (2010). *Gender Differences at Critical Transitions in the Careers of Science, Engineering, and Mathematics Faculty*. The National Academies Press. <https://doi.org/10.17226/12062>
- Nelson, J. A. (2014). The power of stereotyping and confirmation bias to overwhelm accurate assessment: the case of economics, gender, and risk aversion. *Journal of Economic Methodology*, 21(3), 211-231. <https://doi.org/10.1080/1350178X.2014.939691>
- Nielsen, M. W. (2018). Scientific Performance Assessments Through a Gender Lens: a Case Study on Evaluation and Selection Practices in Academia. *Science & Technology Studies*, 31(1), 2-30. <https://doi.org/10.23987/sts.60610>
- Pina, D. G., Barać, L., Buljan, I., Grimaldo, F. y Marušić, A. (2019). Effects of seniority, gender and geography on the bibliometric output and collaboration networks of European Research Council (ERC) grant recipients. *Plos One*, 14(2), artículo e0212286. <https://doi.org/10.1371/journal.pone.0212286>
- Preston, A. E. (2004). *Leaving Science: Occupational Exit from Scientific Careers* (1.ª ed.). Russell Sage Foundation. <https://www.russellsage.org/publications/leaving-science>
- Reskin, B. F. (2003). Including Mechanisms in Our Models of Ascriptive Inequality: 2002 Presidential Address. *American Sociological Review*, 68(1), 1-21. <https://doi.org/10.2307/3088900>
- Ridgeway, C. L. (2014). Why Status Matters for Inequality. *American Sociological Review*, 79(1), 1-16. <https://doi.org/10.1177/0003122413515997>
- Rip, A. (1994). The republic of science in the 1990s. *Higher Education: The International Journal of Higher Education Research*, 28(1), 3-23. <https://doi.org/10.1007/BF01383569>
- Rivera, L. A. y Tilcsik, A. (2019). Scaling Down Inequality: Rating Scales, Gender Bias, and the Architecture of Evaluation. *American Sociological Review*, 84(2), 248-274. <https://doi.org/10.1177/0003122419833601>
- Rosser, S. V. (2004). *The Science Glass Ceiling: Academic Women Scientist and the Struggle to Succeed* (1.ª ed.). Routledge.
- Roubanis, L. (2019). Peer Review or Lottery? A Critical Analysis of Two Different Forms of Decision-making Mechanisms for Allocation of Research Grants. *Science, Technology & Human Values*, 44(6), 994-1.019. <https://doi.org/10.1177/0162243918822744>

- Sackett, P. R., DuBois, C. L. y Noe, A. W. (1991). Tokenism in performance evaluation: The effects of work group representation on male-female and White-Black differences in performance ratings. *Journal of Applied Psychology*, 76(2), 263-267. <https://doi.org/10.1037/0021-9010.76.2.263>
- Sandström, U. y Besselaar, P. V. D. (2016). Quantity and/or Quality? The Importance of Publishing Many Papers. *Plos One*, 11(11), artículo e0166149. <https://doi.org/10.1371/journal.pone.0166149>
- Sandström, U. y Hällsten, M. (2008). Persistent nepotism in peer-review. *Scientometrics*, 74(2), 175-189. <https://doi.org/10.1007/s11192-008-0211-3>
- Sonnert, G. y Holton, G. (1995). *Who Succeeds in Science? The Gender Dimension*. Rutgers University Press.
- Stephan, P. E. y El-Ganainy, A. (2007). The entrepreneurial puzzle: explaining the gender gap. *The Journal of Technology Transfer*, 32(5), 475-487. <https://doi.org/10.1007/s10961-007-9033-3>
- Stewart, A. J. y Valian, V. (2018). *An Inclusive Academy: Achieving Diversity and Excellence*. The MIT Press.
- Tamblyn, R., Girard, N., Qian, C. J. y Hanley, J. (2018). Assessment of potential bias in research grant peer review in Canada. *Canadian Medical Association Journal – CMAJ*, 190(16), e489-e499. <https://doi.org/10.1503/cmaj.170901>
- Thelwall, M. (2018). Do females create higher impact research? Scopus citations and Mendeley readers for articles from five countries. *Journal of Informetrics*, 12(4), 1.031-1.041. <https://doi.org/10.1016/j.joi.2018.08.005>
- Treviño, L. J., Gomez-Mejia, L. R., Balkin, D. B. y Mixon, F. G. (2018). Meritocracies or Masculinities? The Differential Allocation of Named Professorships by Gender in the Academy. *Journal of Management*, 44(3), 972-1.000. <https://doi.org/10.1177/0149206315599216>
- UMyC – Ministerio de Ciencia e Innovación (2011). *Libro blanco: situación de las mujeres en la ciencia española* (Sánchez de Madariaga, I, De la Rica Goiricelaya, S. y Dolado Lobregad, J.J. (coords.)). Unidad de Mujeres y Ciencia – UMyC. <https://www.ciencia.gob.es/stfls/MICINN/Ministerio/FICHEROS/UMYC/LibroBlanco-Interactivo.pdf>
- Van den Besselaar, P. y Leydesdorff, L. (2009). Past performance, peer review and project selection: a case study in the social and behavioral sciences. *Research Evaluation*, 18(4), 273-288. <https://doi.org/10.3152/095820209X475360>
- Van den Brink, M., Brouns, M. y Waslander, S. (2006). Does excellence have a gender? A national research study on recruitment and selection procedures for professorial appointments in The Netherlands. *Employee Relations*, 28(6), 523-539. <https://doi.org/10.1108/01425450610704470>
- Van den Brink, M. (2010). *Behind the Scenes of Science: Gender Practices in the Recruitment and Selection of Professors in the Netherlands*. Amsterdam University Press.
- Van den Brink, M. y Stobbe, L. (2014). The support paradox: Overcoming dilemmas in gender equality programs. *Scandinavian Journal of Management*, 30(2), 163-174. <https://doi.org/10.1016/j.scaman.2013.07.001>
- Van den Brink, M., Brouns, M. y Waslander, S. (2006). Does excellence have a gender? A national research study on recruitment and selection procedures for professorial appointments in The Netherlands. *Employee Relations*, 28(6), 523-539. <https://doi.org/10.1108/01425450610704470>
- Vernos, I. (2013). Quotas are questionable. *Nature*, 495, 39. <https://doi.org/10.1038/495039a>
- Wennerås, C. y Wold, A. (1997). Nepotism and sexism in peer-review. *Nature*, 387, 341-343. <https://doi.org/10.1038/387341a0>
- Xie, Y. y Shauman, K. A. (1998). Sex Differences in Research Productivity: New Evidence about an Old Puzzle. *American Sociological Review*, 63(6), 847-870. <https://doi.org/10.2307/2657505>
- Xie, Y. y Shauman, K. A. (2003). *Women in science: Career processes and outcomes*. Harvard University Press.