

Encuestas de opinión en rankings universitarios: ¿Caja negra o medición confiable?

Adrián Quintero López

Universidad Nacional Autónoma de México

adrianql@unam.mx

Marion Lloyd

Universidad Nacional Autónoma de México

mlloyd@unam.mx

Área temática: Evaluación de docentes e investigadores

Resumen

Esta investigación aborda críticamente las metodologías de los rankings universitarios internacionales, enfocándose en la idoneidad de las encuestas de opinión como herramientas evaluativas. Los rankings universitarios han aumentado su influencia en políticas educativas y percepciones públicas, aunque su validez enfrenta críticas por la falta de transparencia y posibles sesgos en la recolección de datos. Por lo tanto, este estudio revisa sistemáticamente las metodologías utilizadas en estos rankings, poniendo énfasis en la rigurosidad y transparencia del diseño de encuestas. Se identifican problemas como la falta de rigor en la formulación de preguntas y sesgos en la selección de muestras. Estas deficiencias comprometen la fiabilidad de los rankings, un limitante particularmente preocupante dado que las encuestas pueden representar hasta el 50% del valor total de los mismos.

Palabras clave: Rankings universitarios; Metodologías de evaluación; Encuestas de opinión; Transparencia metodológica; Sesgos en evaluación.

Introducción

En las últimas dos décadas, los rankings universitarios han ganado presencia como una herramienta para la toma de decisiones, impactando a las políticas educativas y la opinión pública. Las clasificaciones pretenden hacer una evaluación comparativa de universidades a través de diversos indicadores, incluyendo encuestas de opinión. No obstante, las metodologías han sido objeto de múltiples críticas. Cuestionamientos sobre la validez, la transparencia y la

equidad de estos métodos sugieren que los resultados de los rankings pueden estar influenciados por sesgos y prácticas poco rigurosas.

Esta investigación examinará críticamente las metodologías de los rankings universitarios internacionales, con un enfoque particular en la idoneidad de las encuestas de opinión como herramienta de evaluación en los rankings. Este análisis aborda tanto los desafíos metodológicos inherentes en la recopilación y análisis de los datos de opinión, como las posibles limitaciones y sesgos que podrían influir en los resultados finales.

Críticas Metodológicas a los Rankings Internacionales

Los rankings han evolucionado rápidamente y ahora evalúan tanto la calidad general de las instituciones de educación superior (IES) a nivel global como regiones y áreas específicas de investigación. Las clasificaciones universitarias surgieron entre finales del siglo XIX y principios del XX (Hazelkorn y Mihut, 2011; Wilbers y Brankovic, 2021). Sin embargo, adquirieron una forma más estructurada en la década de 1980, con el surgimiento de las primeras clasificaciones nacionales de universidades en 1983.

La comparación internacional de universidades inició en 2003 con el Academic Ranking of World Universities (ARWU). En 2004, el suplemento del periódico *Times Higher Education* (THE), en colaboración con la consultora Quacquarelli Symonds (QS), lanzó el ranking THE-QS World University Rankings. A diferencia del ARWU, este ranking combinaba encuestas de reputación con información proporcionada directamente por las universidades.

La ausencia de indicadores comunes a nivel global propició el uso de encuestas de reputación, para comparar la calidad educativa entre instituciones (Usher y Savino, 2006). Un argumento favorable para la implementación de las encuestas fue la independencia de fuentes gubernamentales o institucionales (Usher y Savino, 2006). Esto, se asumía, contribuía a una evaluación más objetiva y libre de influencias externas.

Los rankings universitarios, creados por editoriales o entidades independientes, clasifican y jerarquizan a las IES basándose en puntuaciones derivadas de indicadores específicos previamente seleccionados (Usher y Savino, 2006; Lloyd et al, 2011). Pretenden medir el desempeño académico, especialmente en docencia e investigación. Estas clasificaciones de universidades ganaron popularidad al ser vistas como indicativos de calidad (Ordorika y Rodríguez, 2008; Lloyd et al, 2011).

Estas clasificaciones a menudo incluyen como parte de sus indicadores la utilización de encuestas de opinión, para medir de manera indirecta la calidad. Se parte de la premisa de que

los empleadores, académicos y administradores encuestados emiten opiniones informadas e imparciales sobre la calidad institucional (Usher y Savino, 2006). No obstante, la validez de este método se ha puesto en duda, ya que muchos encuestados no están familiarizados con cada institución evaluada, lo que podría sesgar los resultados del ranking (Usher y Savino, 2006).

Entre 2003 y 2023, se desarrollaron alrededor de 63 rankings internacionales y 71 rankings nacionales en 35 países distintos (Lloyd y Quintero, 2023). Cabe señalar que más de la mitad de los rankings internacionales incorporan encuestas de opinión como parte central de sus metodologías (Lloyd y Quintero, 2023). Además, el puntaje asignado a las encuestas representa cerca del 50% del valor total de los rankings (Fidler y Parsons, 2008; Huang, 2012; Jacqmin, 2021).

Los resultados de las encuestas de opinión pueden ser consideradas científicas porque se asume que registran observaciones de manera precisa (Usher y Savino, 2006). Sin embargo, desde el surgimiento de los rankings, han surgido críticas en torno a las metodologías y resultados de los rankings. Estudios cuantitativos han puesto en evidencia discrepancias entre los resultados de los indicadores y el posicionamiento de las universidades, la selección de indicadores, la arbitrariedad en la asignación de valor (ponderación) de los indicadores, así como la presencia de sesgos en los datos (Florian, 2007; Billaut et al, 2009; Huang, 2012; Stergiou y Lessenich, 2014; Selten et al., 2020; Chirikov, 2022). Simultáneamente, otros tipos de análisis han cuestionado la opacidad metodológica, la generalización excesiva de los resultados, la tendencia a favorecer un tipo ideal de universidad (anglosajona), y la fluctuación en el posicionamiento de las universidades (Usher y Savino, 2006; Ordorika y Rodríguez, 2008; Marginson, 2014).

A pesar de esto, los rankings continúan siendo utilizados como herramientas de control sobre la producción académica de docentes, investigadores y universidades en general. De manera paradójica, la rendición de cuentas en las universidades está en cierta medida condicionada por los resultados de estos rankings, los cuales no están sujetos a pruebas de puntajes, validez ni confiabilidad. Esto plantea un dilema sobre la equidad y la justicia de utilizar tales rankings como base para decisiones académicas y administrativas importantes.

En respuesta al surgimiento de los rankings, se conformó el Grupo Internacional de Expertos en Rankings (IREG), el cual, junto con la Asociación de Cooperación Académica (ACA) y la Asociación de Universidades Europeas (UEA), se dedicaron a analizar metodologías y hacer recomendaciones. En 2006, un grupo de líderes académicos creó los “Principios de Berlín”, que contenía “buenas prácticas” sobre el diseño de rankings (IHEP, 2006). Desde 2011, en América Latina se han realizado varios encuentros de académicos y rectores donde se han expresado

preocupaciones sobre los sesgos y usos indebidos de los rankings. Entre ellos se encuentran los distintos Encuentros de RedES Universitarias y Consejos de Rectores de América Latina y el Caribe, coordinados por la UNESCO, y el encuentro “Las universidades latinoamericanas ante los rankings internacionales: impactos, alcances y límites”, organizado por la UNAM, entre otros (Lloyd y Quintero, 2023).

No obstante, muchos de los esfuerzos de los organismos internacionales han contribuido a la legitimación del uso de rankings, a pesar de las críticas teóricas y metodológicas (Barron, 2017). En el mejor de los casos, han emitido recomendaciones que son obligatorias y no cuentan con respaldo legal o normativo. Como resultado, las evaluaciones de los rankings universitarios no están supervisadas ni reguladas por ninguna autoridad internacional, regional y/o local.

Método

Este estudio comprende una revisión sistemática de las metodologías de rankings universitarios, internacionales y nacionales. Se centrará en la rigurosidad y transparencia con que se aplican las encuestas de opinión, destacando la importancia del diseño de las encuestas. Por tanto, el objetivo principal es evaluar la idoneidad de las encuestas empleadas en los rankings.

Para garantizar la validez de los datos recolectados, es esencial considerar aspectos como el tamaño de la muestra, la población objetivo y las técnicas de muestreo. Independientemente de los objetivos de las encuestas, se subraya la necesidad de un buen diseño, que incluya una muestra representativa y libre de sesgos (Herrera Viedma et al., 2024). Estos factores son cruciales para asegurar la precisión de los resultados obtenidos.

El estudio abarca un periodo de dos décadas, desde 2003 hasta 2023. Durante este tiempo, se han desarrollado y diversificado las metodologías utilizadas en los rankings universitarios (Lloyd y Quintero, 2023). Este enfoque permite enriquecer la comprensión de las diversas perspectivas adoptadas por las agencias de rankings.

La selección de rankings para el análisis se basó en la inclusión de instituciones mexicanas, la implementación de encuestas de opinión y su vigencia hasta el 2023. Este enfoque permitió analizar 18 metodologías de rankings internacionales y 3 nacionales. Estas clasificaciones fueron creadas por 10 agencias, mayormente centradas en la publicación de revistas y periódicos.

Análisis de rankings basados en encuestas de opinión

Las encuestas de opinión están diseñadas para evaluar el prestigio percibido por académicos de otras instituciones (*peer review*) respecto a la calidad de enseñanza e investigación de una universidad específica (Selten et al., 2020). Estos métodos se orientan principalmente a los académicos y empleadores, y en menor medida a los estudiantes. Los académicos reciben una ponderación media de 0.28 en rankings internacionales y 0.25 en los nacionales, mientras que los empleadores tienen una media de 0.21 en los internacionales y 0.73 de los nacionales. Los estudiantes únicamente son consultados en rankings que evalúan ciudades y sistemas de educación superior.

Una de las problemáticas de las encuestas de opinión es la poca transparencia en las preguntas, generales y específicas. Todas las metodologías incluyen una breve descripción sobre la escala de evaluación, que varía de 1 a 10 o de 1 a 100. La falta de detalles sobre las preguntas limita la capacidad de realizar análisis estadísticos, como el alfa de Cronbach o el coeficiente de Spearman-Brown, que permitan determinar la consistencia interna y la estabilidad de las respuestas (AERA et al., 2018; Mendiola et al., 2022).

Las encuestas utilizadas en los rankings universitarios generalmente se formulan de dos maneras. Por un lado, se emplean preguntas estructuradas sobre la calidad de una institución en la docencia e investigación (Fidler y Parsons, 2008; Huang, 2012). Por el otro, se solicita a los encuestados que enlisten a las mejores universidades, según su opinión, ordenándolas de mejor a peor. En ambos casos, no se especifica cómo se asignan los puntajes a cada institución.

Además, en los rankings de QS, los resultados de estas encuestas se presentan como sub-puntajes de las dimensiones de docencia e investigación. Esto implica que el valor de las encuestas sea recalculado y normalizado para la dimensión correspondiente, complicando la determinación exacta del puntaje de las universidades. Adicionalmente, a pesar de que se afirma que los resultados de las encuestas están equilibrados para garantizar una representación geográfica y/o por área de conocimiento, no se proporcionan los criterios utilizados para realizar dichos ajustes (Fidler y Parsons, 2008).

La selección de participantes a menudo se realiza mediante la nominación de universidades externas. Es decir, a una institución determinada se le solicita una base de datos de académicos para ser encuestados. Esto influye significativamente en los resultados de los rankings, favoreciendo la inclusión de académicos cuyas respuestas son favorables a quien los recomendó (Rindova et al., 2018).

Existen algunos errores y sesgos relacionados con la selección de participantes. En primer lugar, la falta de información sobre la tasa de respuesta (también llamada como *tasa de retorno*).

Los datos disponibles señalan en 2006, QS mostró una tasa de *retorno* menor al 1% (Huang, 2012). Lo anterior es un reflejo de la poca participación de la comunidad y representatividad en este tipo evaluaciones. También, existen problemas asociados con la distribución desigual de cuestionarios, lo cual provoca sesgos regionales (Huang, 2012). Aunque algunas metodologías señalan que aplican “componentes exponenciales” y factores para asegurar la distribución por país (e incluso por área de conocimiento), en la realidad no se explica como hacen este proceso, que valores asignan y cuál es el resultado de dicha distribución.

En segundo lugar, otra consecuencia de los problemas en las tasas de retorno es que han provocado la implementación de períodos acumulativos en los cuestionarios, frecuentemente de 5 años. Esta estrategia, que consiste en utilizar la suma de opiniones para evaluar instituciones en un año específico, es empleada por la mayoría de los rankings. Es decir, las encuestas de opinión no miden el rendimiento de las universidades en el año anterior, si no en los últimos 5 años.

Esta forma de analizar los datos puede provocar un “efecto de anclaje”, donde los resultados de un año afectan las evaluaciones de reputación futuras (Bowman y Bastedo, 2010). Asimismo, puede producirse el “efecto Mateo”, por el cual instituciones mejor posicionadas tienden a recibir puntuaciones más altas (Fernández-Cano et al, 2018; Rindova et al, 2018; Herrera-Viedma et al, 2024).

En tercer lugar, algunas instituciones recurren a prácticas deshonestas, utilizando estrategias para modificar o inducir datos favorables (*gaming*), con tal de influir en los resultados de los rankings (Lydgate, 2018; Chirikov, 2022; Rindova et al., 2018). Jacqmin (2021) argumenta que “en la práctica, es posible manipular la encuesta en la etapa de diseño de esta, por ejemplo, encuestando a académicos más inclinados a ser favorables hacia un subconjunto de instituciones, o en la etapa de análisis, cuando se incorporan juntos para construir el subindicador” (p. 13).

La credibilidad y la fiabilidad de las encuestas de opinión dependen directamente de la robustez de las metodologías utilizadas, lo cual influye significativamente en la validez de los resultados de dichos rankings. Existen diversos métodos para evaluar la fiabilidad de las encuestas de opinión, como la prueba de test-retest, la estimación del error de medición y el análisis de la consistencia interna (AERA et al., 2018; Mendiola et al, 2022). La mayoría de los rankings limitan sus aparados metodológicos con poca o nula información detallada sobre la confiabilidad y validez de las encuestas. Llama la atención que de las metodologías analizadas,

únicamente un par de clasificaciones mexicanas muestren algún comentario sobre el intervalo de confianza y el margen de error estadístico.

Los principios de Berlín expresan que en el diseño y ponderación de datos, los rankings deben “ser transparente respecto a la metodología utilizada en la creación del ranking”, además de “elegir indicadores de acuerdo a su relevancia y validez” (IHEP, 2006)—cuestión que casi nunca es acatada en las encuestas de opinión empleadas por los rankings—. Una causa de la falta de transparencia en el proceso de evaluación por medio de encuestas es que estas son frecuentemente realizadas por las mismas agencias que producen el ranking o por empresas contratadas.

Además, bajo el precepto de la “secrecía de información”, ocultan los detalles metodológicos. Esta práctica afecta la validez y la confiabilidad del método empleado, así como la percepción de su rigor y transparencia (Rindova et al., 2018). Una de las consecuencias de la poca transparencia es el cuestionamiento sobre la “parcialidad” de los resultados, ya que numerosas universidades muestran puntuaciones similares en estos indicadores, lo que pone en duda la objetividad de la evaluación (Huang, 2012).

Conclusiones

En suma, las encuestas de opinión constituyen una metodología científica ampliamente empleada en diversos campos de la educación, como la evaluación de estudiantes y docentes. No obstante, la credibilidad de las encuestas de opinión depende en gran medida de la eficacia de la metodología empleada y de su implementación. En el contexto de los rankings universitarios, los estudios y críticas metodológicas revisados destacan varias deficiencias tanto en la metodología como en la implementación.

Entre estas críticas se incluyen la falta de rigor y transparencia en el proceso de elaboración de las encuestas, así como la ausencia de información detallada sobre la confiabilidad y validez de estas (Fidler y Parsons, 2008; Rindova et al., 2018; Jacquemin, 2021). Estas limitaciones, sumadas a los sesgos en la selección de muestras y la formulación de preguntas, comprometen la fiabilidad de los rankings basados en este tipo de evaluaciones, lo cual es particularmente relevante dado que los resultados de las encuestas representan cerca del 50% del valor de los rankings.

Es esencial reconocer la influencia de los rankings no solo en el ámbito académico sino también en la sociedad en general. Los rankings son instrumentos que, a menudo, no consideran el contexto nacional ni la heterogeneidad del sistema de educación superior. Por lo tanto, es

fundamental promover el desarrollo y uso de sistemas de información y estudios comparativos, objetivos y confiables, que no solo faciliten una toma de decisiones más ajustada a la realidad de nuestra sociedad, sino que también propicien una evaluación más justa y equitativa de las instituciones de educación superior.

Referencias

- American Educational Research Association, American Psychological Association, y National Council on Measurement in Education. (2018). *Estándares para pruebas educativas y psicológicas* (M. Lieve, Trans.). Washington, DC: American Educational Research Association.
- Barron, G. S. (2017). The Berlin Principles on Ranking Higher Education Institutions: limitations, legitimacy, and value conflict. *Higher Education*, 73, 317–333.
- Billaut, J.-C., Bouyssou, D., y Vincke, P. (2009). Should you believe in the Shanghai ranking? An MCDM view. *Scientometrics*, 84(1), 237–263.
- Bowman, A.B., y Bastedo, M.N. (2010). Anchoring effects in world university rankings: exploring biases in reputation scores. *Higher Education*, 61, 431–444. <https://doi.org/10.1007/s10734-010-9339-1>
- Chirikov, I. (2022). Does conflict of interest distort global university rankings? *Higher Education*, 86, 791-808.
- Fernández-Cano, A., Curiel-Marin, E., Torralbo-Rodríguez, M., y Vallejo-Ruiz, M. (2018). Questioning the Shanghai Ranking methodology as a tool for the evaluation of universities: an integrative review, *Scientometrics*, 116, 2069-2083. <https://doi.org/10.1007/s11192-018-2814-7>
- Fidler, B., y Parsons, C. (2008). World university ranking methodologies: stability and variability. *Higher Education Review*, 40(3), 15-34.
- Florian, R.V. (2007). Irreproducibility of the results of the Shanghai academic ranking of world universities. *Scientometrics*, 72(1), 25–32
- Hazelkorn, E., y Mihut, G. (Eds.). (2021). *Research Handbook on University Rankings: Theory, Methodology*. Edward Elgar Publishing.
- Herrera-Viedma, E., Arroyo-Machado, W., y Torres-Salinas, D. (2024). Losing objectivity: The questionable use of surveys in the Global Ranking of Academic Subjects. *Quantitative Science Studies*. https://doi.org/10.1162/qss_c_00289
-

- Huang, M.H. (2012). Opening the black box of QS World University Rankings, *Research Evaluation*, 21(1), 71-78. <https://doi.org/10.1093/reseval/rvr003>
- Institute for Higher Education Policy (IHEP) (2006). *Berlin Principles on Ranking of Higher Education Institutions*. https://www.ihep.org/wpcontent/uploads/2014/05/uploads_docs_pubs_berlinprinciplesranking.pdf
- Jacqmin, J. (2021). Do ads influence rankings? Evidence from the higher education sector. *Education Economics*, 29(5), 509–526. <https://doi.org/10.1080/09645292.2021.1918642>
- Lloyd, M., Ordorika, I. y Rodríguez Gómez, R. (2011). Los Rankings Internacionales de Universidades, su impacto, metodología y evolución. *Cuadernos de Trabajo de la Dirección General de Evaluación Institucional*, 2(7).
- Lloyd, M., y Quintero López, A. (2023). *Los rankings universitarios en América Latina (2013-2022)*. UNAM.
- Lydgate, C. (12 de septiembre, 2018). *Reed and the rankings game*. Reed College. <https://www.reed.edu/apply/college-rankings.html>
- Marginson, S. (2014). University rankings and social science. *European Journal of Education*, 49(1), 45–59.
- Mendiola, M. S., y González, A. M. (2022). *Evaluación y aprendizaje en educación universitaria: estrategias e instrumentos*. CUAIEED, Coordinación de Universidad Abierta, Innovación Educativa y Educación a Distancia, Universidad Nacional Autónoma de México.
- Ordorika, I, y Rodríguez Gómez, R. (2008). Comentarios al Academic Ranking of World Universities 2008. *Cuadernos de Trabajo de la Dirección General de Evaluación Institucional*, 1(1).
- Rindova, V. P., Martins, L. L., Srinivas, S. B., y Chandler, D. (2018). The Good, the Bad, and the Ugly of Organizational Rankings: A Multidisciplinary Review of the Literature and Directions for Future Research. *Journal of Management*, 44(6), 2175-2208. <https://doi.org/10.1177/0149206317741962>
- Selten, F., Neylon, C., Huang, C.K., y Groth, P. (2020). A longitudinal analysis of university rankings. *Quantitative Science Studies*, 1(3), 1109–1135. https://doi.org/10.1162/qss_a_00052
- Stergiou, K.I., y Lessenich, S. (2014). On impact factors and university rankings: From birth to boycott. *Ethics in Science and Environmental Politics*, 13(2), 101–111
-



CONGRESO INTERNACIONAL DE EDUCACIÓN | EVALUACIÓN 2024



Usher, A., y Savino, M. (2006). *A world of difference: A global survey of university league tables*. Toronto. Educational Policy Institute.

Wilbers, S., y Brankovic, J. (2021). The Emergence of University Rankings: A Historical-Sociological Account. *Higher Education*. <https://doi.org/10.1007/s10734-021-00776-7>
