

# Transparencia y Reproducibilidad en Investigación Social

Julio Iturra y Martín Venegas

2021-07-02



# Contents

<b>1</b>	<b>Introducción</b>	<b>5</b>
<b>2</b>	<b>Transparencia</b>	<b>9</b>
2.1	¿Qué es la transparencia? Un concepto multidimensional . . . . .	9
2.2	¿Crisis en las ciencias? . . . . .	12
2.3	¿Que podemos hacer? . . . . .	19
<b>3</b>	<b>Reproducibilidad</b>	<b>21</b>
<b>4</b>	<b>Herramientas para la transparencia y irreproducibilidad</b>	<b>23</b>
4.1	Pre-registros . . . . .	23
4.2	Flujos de trabajos reproducibles. . . . .	23
<b>5</b>	<b>Palabras finales</b>	<b>27</b>



# Chapter 1

## Introducción

Imagine que usted es un chef, y como tal, disfruta enormemente del arte de cocinar. Como amante de la cocina, cada vez que alguna celebración especial se avecina, es tradición preparar sus mejores recetas para sus invitados. Sin embargo, en esta ocasión, la celebración se llevará a cabo en un restaurante, lo cual implica que la preparación de la comida no depende de usted. Con bastantes dudas, usted accede, teniendo en mente la siguiente pregunta: ¿Cómo puedo tener certeza de que se seguirán los procedimientos adecuados para que la preparación sea de calidad? Pues, estimado lector, es esta la pregunta que subyace a todo el proceso que engloba al conocimiento científico.

La ciencia, al igual que la cocina, no se trata solamente de los productos. Una preparación no aparece por generación espontánea, sino que requiere el seguimiento riguroso de una receta, en la cual usamos los ingredientes adecuados, la cantidad apropiada y los tiempos de cocción. Esta receta no solo contribuye a la preparación de la comida, sino que hace posible que dicho plato pueda ser **reproducido** cada vez que se lo desee, garantizando el mismo resultado. Entonces, como buen chef, la respuesta a la pregunta no es tan compleja: necesitamos ser capaces de evaluar el proceso de elaboración de la comida para asegurarnos de que es el adecuado. Dicho de otro modo, el proceso de preparación debe ser **transparente** y estar abierto al escrutinio de otros cocineros expertos, como también para quienes estén adentrándose en la cocina. En el caso de la comida, que el restaurante cuente con una cocina abierta o construida en torno a ventanales bastaría para lograr este objetivo. Sin embargo, dentro del campo de la ciencia ¿cómo logramos que el proceso de producción del conocimiento científico esté abierto al escrutinio público?

Esta pregunta puede ser algo engañosa ¿acaso los procesos de investigación en la ciencia no están ya abiertos al escrutinio público? La narrativa actual pareciese sugerir que no. En el último tiempo ha primado el diagnóstico de que la ciencia está viviendo una crisis, donde polémicas situaciones de malas

prácticas académicas han salido a la luz. Un ejemplo de estas malas prácticas es el falseamiento de datos. Uno de los casos más emblemáticos es el de Diderik Stapel, una figura académica con alto prestigio en el campo de la psicología social a quién se le acusó y confirmó de falseamiento de datos. Más de 10 años de investigación y 150 artículos -algunos de ellos en las revistas más prestigiosas- fueron puestos en duda a raíz de las malas prácticas de Stapel. Así también el trabajo de muchos colegas y estudiantes fue desacreditado. El caso del ex doctor Stapel acabó en la revocación de su doctorado, la retracción de 58 artículos de investigación y, prácticamente, el fin de su carrera académica.

Casos como el de Diderik Stapel existen muchos (ver Abril Ruiz, 2019), sin embargo, no son las prácticas más recurrentes. Dentro de la investigación científica existen una serie de prácticas que caen en un terreno gris cuando se trata de su evaluación ética, estas son las llamadas *prácticas cuestionables de investigación*. La preocupación dentro de la comunidad científica es que la acumulación y la poca fiscalización de estas prácticas lleven a una ciencia poco transparente, con dificultad en torno a la reproducibilidad de los análisis y de sus resultados, y que a la larga se pierda la confianza en el quehacer científico.

La pérdida de la confianza en el quehacer científico afecta el objetivo de las ciencias sociales. Generalmente se tiene la concepción de que es parte de las tareas de los científicos sociales el contribuir al bienestar de la sociedad por la vía de las herramientas de investigación. Esta es la noción de las ciencias sociales como un bien público (Thibodeaux, 2016). Bajo esa idea, los efectos de una crisis de credibilidad tienen dos posibles efectos concretos en las ciencias sociales. Primero, la falta de credibilidad podría afectar la confianza en los hallazgos y en las disciplinas que componen las ciencias sociales. Ya existe evidencia sobre un incremento de desconfianza hacia los científicos en países como Estados Unidos (Motta, 2018) y sumarle una crisis de la ciencia a esa situación solo generaría un ambiente mucho más complejo. Segundo, y estrechamente relacionado a lo anterior, una pérdida de credibilidad en la ciencia podría impactar en la elaboración de políticas sociales, mal orientando las prioridades y los recursos del país. Si se quieren evitar estas situaciones, es necesario tomar cartas en el asunto y orientar los esfuerzos a devolverle la credibilidad a las ciencias sociales.

Este manual busca contribuir a una ciencia social más creíble, basándose en el marco de la ciencia abierta. La ciencia abierta se entiende como:

“...la práctica de la ciencia de manera que otros puedan colaborar y contribuir, donde los datos de la investigación, las notas de laboratorio y otros procesos de investigación estén disponibles libremente, en términos que permitan la reutilización, redistribución y reproducción de la investigación y sus datos y métodos subyacentes”. (FORSTER, Open Science Teaching Resource)

Dentro de la ciencia abierta, nos centraremos en dos conceptos: transparencia y reproducibilidad. Si seguimos la metáfora de la cocina que planteamos en un principio, la transparencia y la reproducibilidad son dos conceptos similares,

pero no idénticos. La transparencia implicaría la posibilidad de evaluar y poner en discusión la receta y la ejecución de la misma. En cambio, la reproducibilidad apuntaría a que la receta sea lo suficientemente clara y precisa para que el mismo plato, con el mismo sabor, pueda ser preparado por cualquier persona que contase con los ingredientes y recursos necesarios. En el campo de las ciencias, esta distinción la hacemos entre: 1) la transparencia de los procesos de producción científica (e.g. transparentar el plan de análisis) y 2) la reproducibilidad de los análisis en los artículos (e.g. código de procesamiento y análisis de datos permite reproducir el artículo). Este manual estará estructurado en torno a estas dos conceptos.

Relacionado al primer concepto (transparencia), haremos un barrido un tanto más detallado sobre la crisis de la reproducibilidad que aquí hemos mencionado brevemente. Ahondaremos en los factores que contribuyen a su reproducción, a las razones éticas por el cuales es necesario hacer un cambio y en las recomendaciones que se pueden seguir para adoptar ciertos principios de transparencia. En el caso del segundo concepto (reproducibilidad), nos centraremos en los análisis reproducibles, teniendo un carácter mucho más práctico. Presentaremos las distintas consideraciones que hay que tener para que un análisis sea fácilmente reproducible, desde tipos de flujos de trabajo hasta herramientas específicas que faciliten la reproducibilidad de los análisis.

Estimado investigador o investigadora de las ciencias sociales, este documento va dirigido a usted. Independiente de su disciplina, de su trayectoria académica o de su conocimiento previo con respecto a estas temáticas, en este manual tenemos dos simples objetivos. El primero es que usted pueda convencerse de que, efectivamente, es necesario dar un giro en la forma que hacemos ciencia actualmente y que la adopción de prácticas relacionadas a la ciencia abierta son el primer paso en ese giro. El segundo es poder instruirlo en esa adopción de prácticas, particularmente en lo que respecta a la transparencia y reproducibilidad. Al final de este manual, usted será capaz tanto de argumentar por qué la transparencia y la reproducibilidad son un paso importante en el avance de las ciencias sociales, así como también contará con una serie de herramientas para llevar esto al quehacer académico del día a día.





## Chapter 2

# Transparencia

Esta sección tratará sobre la transparencia en la investigación científica, haciendo énfasis en las ciencias sociales. Nuestro objetivo es poder comunicar de forma clara y concisa tres puntos: a) qué es la transparencia, b) por qué la necesitamos y c) cómo podemos adoptarla. En esta sección veremos los dos primeros puntos, luego, en la sección de herramientas presentaremos algunas formas de adoptar la transparencia. La tónica de este escrito es la práctica, es decir; todo lo que presentemos acá tiene la finalidad de servir de camino para poder aprender y aprehender herramientas que promueven la transparencia. Dicho esto, comencemos con los dos primeros puntos.

### 2.1 ¿Qué es la transparencia? Un concepto multidimensional

Nuestro punto de partida es que la transparencia es un concepto amplio y multidimensional. Por eso, antes de adentrarnos en su complejidad, comencemos con una definición de diccionario para comprender lo más básico en la idea de transparencia. Según la Real Academia Española la transparencia es la cualidad de un cuerpo que permite ver a través de él. Un gran ejemplo para llevar esta definición a la práctica es el vidrio de una ventana. La transparencia del vidrio nos permite ver con claridad lo que está el otro lado, como por ejemplo un paisaje. Sin embargo, ¿qué ocurre cuándo la transparencia del vidrio se va perdiendo? La respuesta es simple, pero potente: la claridad con la que veíamos el paisaje se va difuminando. Esta pérdida de claridad puede dar como resultado que nuestra observación del paisaje se torne ambigua y errónea, o dicho de otro modo, cada vez será más difícil analizar el paisaje. En esta metáfora, el paisaje sería equivalente al proceso de producción científica, y el vidrio representa la claridad con la que podemos analizar estos procesos. De esta manera, la base de la idea de transparencia es que permite analizar con claridad un fenómeno,

una situación, o en este caso, un proceso.

¿Qué implica un proceso científico transparente? Ya existen algunas respuestas a esta pregunta. Por ejemplo, Breznau (2021) entiende la transparencia como una forma en que los investigadores pueden revelar el proceso, ideas y materiales que sustentan un argumento o una teoría, con tal de contribuir a una comunidad científica más ética. Otra perspectiva es la de Aczel et al. (2020), quienes proponen a la transparencia como un principio que permite evaluar y reproducir los hallazgos científicos, así como también a sintetizar investigaciones y contribuir a la ejecución de metanálisis. Estas perspectivas son una primera aproximación a las implicancias de un proceso científico transparente, no obstante, no son una presentación exhaustiva del concepto.

Para abarcar la complejidad del concepto de transparencia, y las múltiples formas que puede adoptar, nos basamos en la taxonomía de Elliott (2020). La taxonomía se basa en cuatro grandes preguntas: ¿por qué?, ¿quienes? ¿qué? y ¿cómo? Cada una de estas preguntas tiene al menos una dimensión asociada. La primera pregunta se refiere a los propósitos por los cuales adoptar la transparencia; la segunda pregunta apunta a la audiencia que está recibiendo la información; la tercera pregunta hace alusión al contenido que es transparentado y la cuarta pregunta consiste en cuatro dimensiones distintas sobre cómo adoptar la transparencia: actores, mecanismos, tiempo y lugar. También, la taxonomía propone una dimensión sobre los peligros que podrían afectar a las iniciativas que busquen promover la transparencia. Aunque esta última dimensión no es una forma de transparencia -en el mismo sentido que las otras dimensiones-, si tiene relevancia en tanto permite mejorar las iniciativas de transparencia. Una representación gráfica puede verse en la Figura N° 2.1.

Para comprender mejor la taxonomía la Figura N° 2.2 presenta una versión detallada de cada dimensión en conjunto a una lista no exhaustiva de variaciones. Por ejemplo, la dimensión del propósito sintetiza varios de los puntos que ya se han señalado en la literatura sobre ciencia abierta: hacer la ciencia más replicable, facilitar la interacción crítica, facilitar el reanálisis de resultados, entre otros. La mayoría de estos propósitos están estrechamente relacionados con el horizonte de una ciencia con mayor credibilidad. Dentro de las dimensiones restantes, existe una en particular que queremos recalcar: los contenidos. Aquí, los distintos contenidos que pueden ser transparentados van desde cosas complejas como los juicios de valor, hasta lo más concreto como datos, métodos y materiales, o dicho de otra forma, el diseño de investigación. El diseño de investigación es uno de los que más se ha promovido la transparencia últimamente, especialmente a raíz de los pre-registros. Profundizaremos en esto más adelante en el escrito.

En síntesis, un proceso de investigación transparente es uno que se puede evaluar con claridad y facilidad. La taxonomía de Elliott (2020) nos ayuda a comprender la diversidad de dimensiones en las que se puede pensar la idea de transparencia. Sin embargo, aun existe la pregunta del *por qué*. ¿Por qué es relevante la idea de transparencia? La dimensión de propósitos de la taxonomía ya nos ha dado

2.1. ¿QUÉ ES LA TRANSPARENCIA? UN CONCEPTO MULTIDIMENSIONAL11

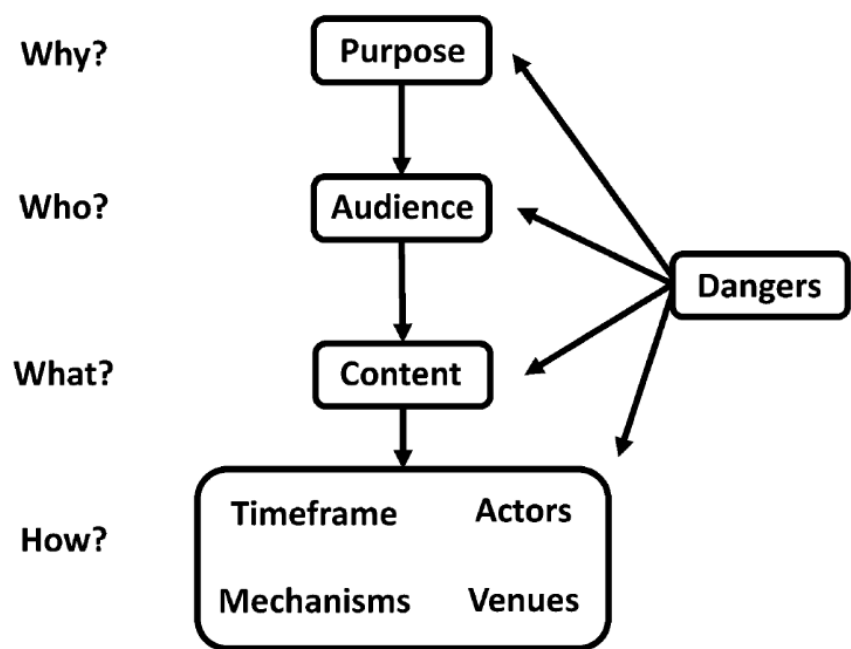


Figure 2.1: Taxonomía de Transparencia

algunas pistas: la ciencia puede ser mejor. Y no solo eso, sino que inclusive podría estar en crisis.

<b>Purpose</b> <ul style="list-style-type: none"> <li>Facilitating the reanalysis of results</li> <li>Making science more replicable</li> <li>Promoting innovation</li> <li>Maintaining accountability of experts</li> <li>Facilitating critical interaction</li> <li>Promoting high-quality policy making</li> <li>Enabling members of the public to make decisions in accord with their values</li> <li>Promoting trustworthiness</li> </ul>	<b>Audience</b> <ul style="list-style-type: none"> <li>Scientists doing the research</li> <li>Other scientists</li> <li>Other academics</li> <li>Policy makers</li> <li>Politicians</li> <li>Journalists</li> <li>Specific stakeholder groups</li> <li>General public</li> </ul>
<b>Content</b> <ul style="list-style-type: none"> <li>Data, methods, code, materials</li> <li>Interpretations of the data, methods, and code for nonspecialists</li> <li>Value judgments of many sorts</li> <li>Values or factors that influence the judgments</li> <li>Deliberations underlying reports</li> <li>Implications of value judgments</li> </ul>	<b>Timeframe</b> <ul style="list-style-type: none"> <li>Before research begins</li> <li>Throughout the research process</li> <li>Immediately after data are collected</li> <li>After publication</li> <li>During or after subsequent reviews or analyses of the research</li> </ul>
<b>Actors</b> <ul style="list-style-type: none"> <li>Scientists who performed the research</li> <li>Other scientists</li> <li>Scholars working in other fields (e.g., HPS or STS)</li> <li>Journalists</li> <li>Scientific societies</li> <li>Government agencies</li> <li>Nongovernmental organizations and civil society organizations</li> </ul>	<b>Mechanisms</b> <ul style="list-style-type: none"> <li>Discussions among scientists (oral or written)</li> <li>Interdisciplinary collaborations</li> <li>Collaborations with community members</li> <li>Government advisory bodies and other initiatives</li> <li>Adversarial proceedings (e.g., science courts)</li> </ul>
<b>Venues</b> <ul style="list-style-type: none"> <li>Communication by scientists (oral or written, including social media)</li> <li>Registries and repositories</li> <li>Science journalism</li> <li>Reports from government agencies</li> <li>Reports from nongovernmental organizations or community groups</li> </ul>	<b>Dangers</b> <ul style="list-style-type: none"> <li>Wasting scarce resources</li> <li>Slowing down science</li> <li>Harming companies</li> <li>Violating privacy</li> <li>Generating inappropriate skepticism</li> <li>Creating a false sense of trust</li> <li>Causing confusion</li> <li>Facilitating efforts to harass or mislead</li> </ul>

Figure 2.2: Variaciones por dimensión de transparencia

## 2.2 ¿Crisis en las ciencias?

En los últimos años, ha venido tomando fuerza la idea de que existe una crisis en la ciencia. Esta idea se basa en el diagnóstico de que gran parte de los artículos científicos en distintas disciplinas no son posibles de reproducir ni replicar. Tanto la reproducibilidad (emplear el mismo diseño y datos para reproducir los hallazgos de un artículo) o la replicabilidad (emplear el mismo diseño y distintos datos para obtener los mismos resultados) son componentes centrales de la ciencia. Es más, algunos de los mecanismos que permiten el avance del conocimiento científico son el escrutinio y la verificabilidad de sus hallazgos, por lo que un fallo en el cumplimiento de estos es un golpe directo a la credibilidad de la ciencia. La pregunta es ¿existe realmente una crisis en las ciencias?

Unas de las fuentes más comunmente citadas para introducir la idea de crisis es una encuesta realizada en la revista *Nature*. En esta encuesta, Baker (2016) logró obtener las opiniones de poco más de 1,500 investigadores de disciplinas como la química, ingenierías y la medicina, sobre tópicos relacionados a la reproducibilidad en las ciencias. El resultado principal muestra que un **90% de los encuestados está de acuerdo en la existencia una crisis**, donde un 52% piensa que es una gran crisis y un 38% la percibe como una ligera crisis. En este mismo estudio, se les pregunta por los factores que contribuyen a esta crisis, donde la cultura del *pública o perece* y el *reporte selectivo de resultados* aparecen como los protagonistas. Si bien la encuesta no es una muestra representativa de toda la comunidad científica, presenta una panorámica que lleva a, por lo menos, considerar la crisis de la ciencia como tema que merece atención.

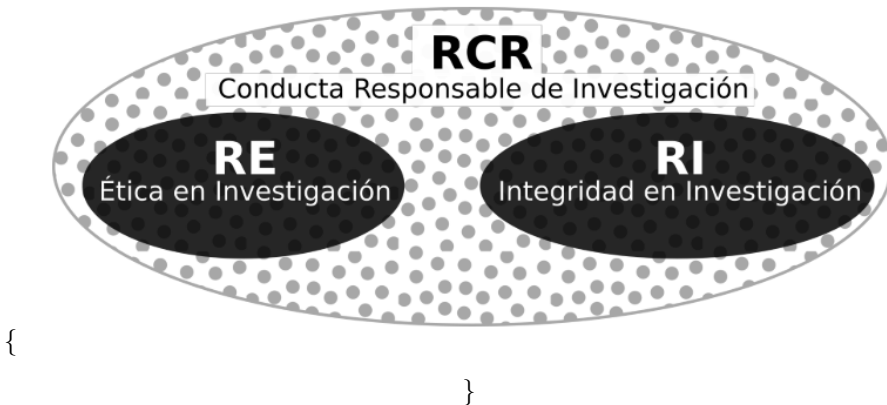
Actualmente, existe un cuerpo de literatura que se ha dedicado diagnosticar y proponer alternativas de solución ante la idea de una ciencia en crisis. Dentro del diagnóstico, variados estudios han orientado sus esfuerzos a esclarecer cuáles son los factores que podrían estar influenciando esta crisis. En esta sección, presentaremos estos factores en dos grandes dimensiones. La primera dimensión se relaciona a las prácticas de investigación, específicamente aquellas prácticas que ya sea en menor o mayor grado afectan la credibilidad de los hallazgos científicos en su conjunto. La segunda dimensión se centrará más en factores institucionales y que no dependen estrictamente de los investigadores. Una vez explicados estos factores, analizaremos como ciertas herramientas que promueven la transparencia en la investigación podrían c

Comenzemos con las prácticas de investigación. Para esquematizar de mejor manera qué es lo problemático de ciertas prácticas, es que utilizaremos el esquema conceptual de Steneck (2006). El esquema parte de una distinción básica entre la *ética en investigación* y la *integridad en investigación*, englobando ambas bajo el gran concepto de *Conducta Responsable de Investigación (RCR)* (ver Figura N° 2.2. A grandes rasgos, la RCR se puede entender como el “llevar a cabo la investigación de forma que se cumplan las responsabilidades profesionales de los investigadores, tal y como las definen sus organizaciones profesionales, las instituciones para las que trabajan y, en su caso, el gobierno y el público” (Steneck, 2006)<sup>1</sup>.

\begin{figure}

---

<sup>1</sup>A raíz de esta definición, y también en base al código de conducta detallado por XX es que Abril Ruiz (2019) proponen a la siguiente definición: “la RCR es el conjunto de normas, habitualmente plasmadas en códigos de conducta, que pretenden ser una guía para que nuestras acciones se mantengan dentro de la integridad investigadora” (p.144).



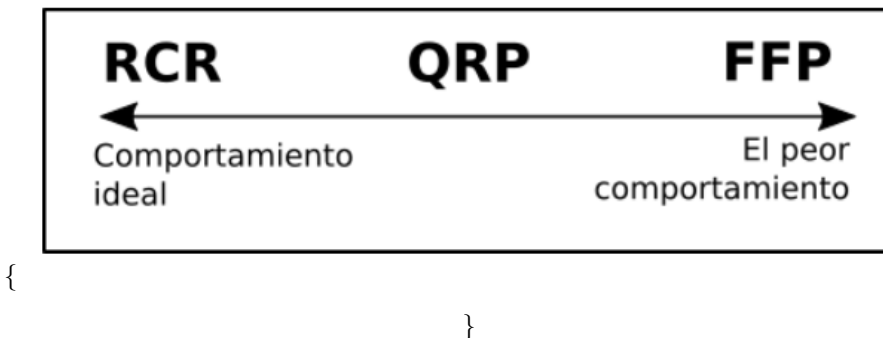
{

}

\caption{Conducta Responsable de Investigación. Imagen de Abril Ruiz (2019) basada en Steneck (2006)} \end{figure}

Dentro de este concepto, la ética de investigación está relacionada al comportamiento académico visto desde la óptica de los principios morales (Steneck, 2006), lo que se expresa en tópicos como el uso de datos, los consentimientos informados, el trato con pacientes -en el caso de las ciencias biomedicas-, por dar algunos ejemplos. La definición que ofrece Steneck (2006) señala que la ética de investigación se define como “el estudio crítico de los problemas morales asociados o que surgen en el curso de la investigación” (p.56) En cambio, la integridad en investigación se entiende como “poseer y adherirse firmemente a las normas profesionales, tal y como las señalan las organizaciones profesionales, las instituciones de investigación y, en su caso, el gobierno y el público” (Steneck, 2006, p.56). A diferencia de la ética de investigación, el concepto de integridad está regido por los estándares profesionales más que por los principios morales, su función es plantear una guía clara para la conducta investigativa, de ahí que sea el concepto utilizado por distintos códigos de conducta[].

\begin{figure}



\caption{Gradación del comportamiento integro en investigación. Imagen de

Abril Ruiz (2019) basada en Steneck (2006)} \end{figure} Habiendonos situando dentro del concepto de integridad en la investigación, podemos pasar a delinear las principales prácticas que atentan contra él y que se han propuesto como factores que contribuyen a la crisis en la ciencia. Tanto Steneck (2006) como distintos códigos de conducta de universidades e instituciones de financiamiento [] evalúan las prácticas de investigación en un continuo, que representa cuánto adhieren los investigadores a los principios de integridad científica. La Figura N° 2.2 esquematiza esta idea mostrando dos extremos, donde a la izquierda está el mejor comportamiento (RCR), y a la derecha el peor comportamiento (FFP). Las FFP son un abreviación en lengua inglesa para referirse a *Fabrication, Falsification, Plagiarism* (Invención, Falsificación y Plagio), también conocidas como *mala conducta académica*. En el medio del continuo están las *prácticas cuestionables de investigación* (QRP, por sus siglas en inglés) las cuáles refieren a “acciones que violan los valores tradicionales de la empresa de investigación y que pueden ser perjudiciales para el proceso de investigación” (*National Academies of Science* 1992 en Steneck, 2006, p.58). A diferencia de las FFP, las QRP son prácticas que tienen el potencial de dañar la ciencia y no que la dañan directamente (como las FFP).

Comprendidos ambos conceptos, veamos las situaciones que podrían categorizarse como FFP y las que podrían concebirse como QRP.

### 2.2.1 Mala conducta académica (FFP)

La mala conducta académica suelen ser situaciones polémicas y que muchas veces alcanzan gran cobertura mediática. El libro de Abril Ruiz (2019) presenta una serie de situaciones, en distintas disciplinas y años, en las que investigadores han sido descubiertos cometiendo prácticas que atentan directamente a la ciencia. Las situaciones son variadas, existen casos de manipulación de imágenes [], exageración de los registros de laboratorio, o de plano la invención de conjuntos de datos enteros. En esta sección veremos el caso de Diderik Stapel como ejemplo de malas prácticas de investigación y plantearemos la relación que tiene con la transparencia en la investigación.

#### 2.2.1.1 Diderik Stapel

Probablemente, el caso de Diderik Stapel sea uno de los más emblemáticos y representativos de este problema. Diderik Stapel era un investigador de la *Tilburg University* que se dedicaba al campo de la psicología social. Su carrera se caracterizó por una trayectoria ejemplar, obtuvo su M.A en psicología y comunicaciones el año 1991, se doctoró en psicología social el año 1997 y trabajó como profesor asociado primero en la *University of Groningen* (2000-2006) y desde el 2006 en la *Tilburg University*. Fue fundador del *Tilburg Institute for Behavioral Economics Research*, del cual el 2010 se convirtió en decano. Así también, fue galardonado con el premio a la trayectoria académica por la *Society of Experimental Social Psychology*. En breve, Stapel era una figura de alto estatus en el mundo académico, entre 1995 y 2015 publicó

aproximadamente 150 artículos en revistas científicas, algunas de las más prestigiosas (e.g. *Science*). Sin embargo, el año 2011 se confirmó que gran parte de su trayectoria académica era una farsa.

Después de la verificación de su culpa ante acusaciones de malas prácticas, la carrera de Diderik Stapel acabó. Fue desvinculado de la Tilburg University, se le rebocó su título de doctorado y toda su trayectoria académica fue investigada acusiocientemente. El informe final de la investigación encontró que más de 50 de sus artículos eran fraudulentos. De hecho, Stapel ocupa el tercer lugar en *Retraction Watch*, con 58 de sus artículos retractados. A efectos de este escrito, lo que más destacable es que Stapel cometió conductas de mala conducta académica durante más de 15 años. La pregunta es ¿cómo fue esto posible? ¿cómo ninguno de sus colegas, alumnos o co-autores se dio cuenta de sus malas prácticas? La respuesta breve es por la falta de transparencia durante el proceso de investigación.

Los artículos periodísticos que han profundizado en el caso (e.g. Carey, 2011) han relatado parte del proceso investigativo de Stapel. Dentro de los procesos de producción y análisis de datos, Stapel se caracterizaba por hacer todo el trabajo solo y “a puertas cerradas”. Es decir, nadie más que él tenía acceso a los datos brutos, ni tampoco a la ejecución de las pruebas estadísticas. Generalmente, Stapel compartía con sus colegas y alumnos de doctorado la base de datos lista, con las pruebas estadísticas ya hechas y, claro está, con resultados significativos. Estas prácticas no causaron sospechas durante muchos años, es más, a muchos de sus estudiantes les parecía una práctica normal y eficiente. Además, con el estatus de Stapel ¿qué podría estar mal? Sin embargo, era a raíz de esta práctica que Stapel tenía la oportunidad de inventar y falsificar datos a su conveniencia. Esto explica en gran parte de su trayectoria académica llena de grandes hallazgos.

El caso de Stapel deja un punto base sobre la mesa: la falta de transparencia en el proceso investigativo dio cabida a la mala conducta académica. Cómo nadie más colaboraba con el procesamiento de datos, ni tampoco parecía extraño que así fuera, las oportunidades para la falsificación de los datos estaba abierta. Ahora, esta no es necesariamente una relación de causalidad, la falta de transparencia no tiene porque terminar en conductas como fabricación o falsificación de datos. Sin embargo, tal y como lo argumentan (O’Boyle et al., 2017) si son una oportunidad para violaciones a la integridad científica más sutiles, tales como las QRP.

### 2.2.2 Prácticas cuestionables de investigación (QRP)

Recordemos, las QRP son prácticas que en si mismas no dañan directamente la empresa científica, pero si tienen el potencial de hacerlo. En la literatura sobre el tema, existen una multiplicidad de términos y listas de prácticas específicas que pueden categorizarse como cuestionables. Abril Ruiz (2019) hace una recopilación y traducción de distintos códigos de conducta de distintas



universidades y organismos, el cuál nosotros sistematizamos en la Figura N° X. La figura muestra un conjunto de prácticas divididos en cuatro categorías, de acuerdo a si las prácticas tienen que ver con: la redacción y el reporte de los resultados, el procesamiento de los datos, temas de citación y uso de ideas ajenas y, por último, sobre relaciones con otros actores en el campo de la ciencia. El objetivo de esta recopilación es meramente esquemático, en el sentido de poder gatillar la reflexión respecto a la experiencia del lector a partir de algunas situaciones concretas que podrán afectar las investigaciones.

Han habido algunos estudios que han intentado medir directamente la existencia de estas prácticas a través de encuestas cuantitativas. Fanelli (2009) hizo un metanálisis que tenía por objetivo sistematizar los resultados de estudios que hasta esa fecha habían abordado las prácticas de investigación desde encuestas cuantitativas. Los resultados mostraron que un 1.97% de investigadores había inventado datos al menos una vez y que un 33.7% había realizado alguna vez una QRP como “borrar puntos de los datos basados en un sentimiento visceral”. Un estudio más reciente, también basado en encuestas cuantitativas sobre prácticas, es el de John et al. (2012). En este estudio, los resultados mostraron que un 36.6% de quienes participaron alguna vez habían practicado alguna QRP. En detalle, analizando los porcentajes práctica a práctica se halló que el 50% de los psicólogos encuestados alguna vez reportaron selectivamente estudios que apoyaran su hipótesis; un 35% alguna vez reportaron resultados inesperados como esperados; y un 2% alguna vez reportó datos falsos. Estos estudios son una primera aproximación a la existencia de las QRP en la ciencia.

Existen ciertas prácticas que han sido tratadas con más énfasis en la literatura, y que son las que queremos destacar acá. Por un lado, está el *sesgo de publicación*. A grandes rasgos, los sesgos de publicación se refieren a la publicación exclusiva de resultados significativos, en desmedro de los no significativos. Dicho de otra forma, cuándo el criterio de selección para la publicación de resultados es la significancia, el cuerpo de literatura de esa temática particular comienza a sesgarse. El estudio de Franco et al. (2014) tiene por objetivo cuantificar esta situación, específicamente en ciencias sociales. En su estudio encontraron un patrón interesante: “estudios donde la hipótesis principales arrojan resultados nulos son 40% menos probables de ser publicados en una revista científica, en contraste a estudios que arrojen resultados significativos” (Christensen et al., 2019, p.41). Los descriptivos se pueden ver en la Figura N° 2.2.2.

Otra práctica es la del *p-hacking*...

Por último, está el *HARKing*, el cual...

Un estudio que se aleja de la encuesta directa sobre prácticas, pero que contribuye estimando cuál es el potencial de que una QRP afecte los resultados de un artículo es el de Simmons et al. (2011). En detalle, los autores buscan calcular el *likelihood* de obtener un falso positivo (error Tipo I)

de acuerdo a al nivel de manipulación intencionada de los datos. El procedimiento consistió en calcular 15.000 muestras de 20 observaciones por condición de tratamiento para un experimento hipotetico y determinaron cuatro posibles ajustes a los datos: 1) plantear dos variables dependientes correlacionadas, 2) incrementar la muestra entre 20 y 30 casos según condición de tratamiento, 3) controlar por una variable sociodemográfica (género) o plantear una interacción genero y condición de tratamiento y 4) flexibilidad en eliminar alguna condición de tratamiento. Como se puede ver en la Figura N° ??, el resultado principal es que a medida que aumenta la cantidad de manipulación en los datos, el likelihood de obtener un falso positivo aumenta progresivamente. Una situación en donde se realicen los cuatro ajustes planteados tiene un 60.7% de likelihood de encontrar un falso positivo.

Una herramienta que se ha utilizado para estimar la existencia de p-hacking en los cuerpos de literatura es la *p-curve*. La *p-curve* “describe la densidad de los *p-values* reportados en una literatura, aprovechando el hecho de que si la hipótesis nula fuera verdadera (es decir, sin efecto), los *p-values* deben distribuirse uniformemente entre 0 y 1” (Christensen et al., 2019, p.67.). De esta manera, en cuerpos de literatura que no sufran de p-hacking, la distribución de *p-values* debería ser asimetrica a la derecha, en cambio, si existe sesgo por p-hacking la distribución de *p-values* estaría distribuida de forma asimetrica a la izquierda. Simonsohn et al. (2014) proponen esta herramienta y la prueban en dos muestras de artículos de la *Journal of Personality and Social Psychology (JPSP)*. Las pruebas estadísticas consistieron en confirmar que la primera muestra de artículos (que presentaban signos de p-hacking) estaba sesgada, en cambio la segunda muestra (sin indicios de p-hacking), no lo estaba. Los resultados corroboraron las hipótesis, en detalle, los artículos que presentaban solamente resultados con covariables, resultaron tener una *p-curve* asimetrica a la izquierda.

\begin{figure}

TABLE 4.1 LIKELIHOOD OF OBTAINING A FALSE POSITIVE

Researcher degrees of freedom	Samples $p < .05$
(1) Two correlated dependent variables	9.5%
(2) Adding 10 more observations per cell	7.7%
(3) Controlling for or interacting with gender	11.7%
(4) Dropping (or not) one of three treatments	12.6%
Combine 1 and 2	14.4%
Combine 1, 2, and 3	30.9%
Combine 1, 2, 3, and 4	60.7%

NOTES: Table shows the likelihood of obtaining a  $p$ -value  $< .05$  with simulated data with no true effect. Given no true effect, we should expect 5 percent of the samples to have  $p$ -values  $< .05$ , but taking advantage of researcher degrees of freedom gives us many more low  $p$ -values.

{

}

\caption{Porcentajes de publicación de acuerdo a significancia de resultados.  
Imagen de Christensen et al. (2019).} \end{figure}

### 2.3 ¿Que podemos hacer?

Tanto en las ciencias sociales, como en otras disciplinas han ido emergiendo una variedad de recomendaciones y consejos que contribuyen a la adopción de la transparencia. Por ejemplo, Crüwell et al. (2018) propone formar investigadores y estudiantes a partir de la promoción de siete principales tópicos: entender la ciencia abierta; acceso abierto; la importancia de los datos, material y código abierto, los análisis reproducibles; los pre-registros; la replicación y, por último, la enseñanza de la ciencia abierta. De forma similar,...



## Chapter 3

# Reproducibilidad

En el marco de la denominada “crisis de la ciencia”, existe una serie de iniciativas que han buscado hacerse cargo de un problema frecuente en las investigaciones, principalmente de carácter cuantitativo. En este sentido, ¿Cuántas veces nos hemos enfrentado a un trabajo publicaco que no comparte sus materiales, y que por tanto, es imposible acceder a los procedimientos que dieron luces de sus resultados?



## Chapter 4

# Herramientas para la transparencia y irreproducibilidad

### 4.1 Pre-registros

### 4.2 Flujos de trabajos reproducibles.

Otro tratamiento extensivo de la idea de transparencia se da en las TOP Guidelines. Estos son principios que buscan alcanzar un formato de investigación reproducible a través del aumento de la transparencia en el proceso y los productos de investigación (Nosek et al., 2014). Estos principios sirven de guía a las revistas académicas para poder adherir progresivamente al ideal de transparencia en la ciencia. Son ocho principios:

1. Citación
2. Transparencia de datos
3. Transparencia de métodos analíticos (código)
4. Transparencia de los materiales
5. Transparencia del diseño y el análisis
6. Preregistro de estudios
7. Pre registro de planes de análisis
8. Replicación

A grandes rasgos, el principio de citación propone que las normas de citado deben ampliarse también a los datos y códigos, permitiendo reconocer su autoría intelectual (Nosek et al., 2015). Los principios de transparencia de datos, métodos analíticos, materiales y diseño y análisis (2 a 5) refieren a la transparencia en su forma más concreta: la apertura del proceso de

investigación para su evaluación. El detalle puesto a los principios responde a la generalizabilidad que se le busca dar a los principios. Por ejemplo, un estudio observacional cuantitativo no tiene material que transparentar, pero sí datos y métodos analíticos. Así también, un estudio cualitativo quizás no tenga código que transparentar, pero sí un diseño y una bitacora detalla el proceso de análisis. En el caso de los principios relacionados al pre-registro,

Nosek et al. (2015) argumenta que registrar los estudios los hace más descubribles, incluso si no son publicados. Así también, los preregistros del plan de análisis contribuyen a distinguir entre los análisis confirmatorios y exploratorios (ver Nosek et al. (2018) para un manejo detallado del tema). Por último, el principio de replicación fomenta las oportunidades para la corrección de artículos y redirecciona la investigación en vías más prometedoras (Nosek et al., 2015).

Cada uno de estos principios cuenta con tres niveles, que sirven para medir el grado de inclusión de la transparencia por parte de una revista científica. La Figura N° 4.1 muestra cada nivel en relación el principio correspondiente. Se

añade un nivel 0 que no cumple los estándares de transparencia con la finalidad de tener una comparación. Por ejemplo, para los estándares de transparencia del método de análisis (código), el nivel 1 dicta que las revistas deben solicitar la existencia del código de análisis, en cambio, el nivel 3 es más estricto en plantear que el código de análisis debe estar almacenado en un repositorio confiable y que el análisis será reproducido durante el proceso de revisión. El mismo método se puede aplicar para el preregistro del plan de análisis. En el nivel 1 las revistas promueven el uso de preregistros, en cambio, en el nivel 3 los preregistros son obligatorios y también reconocidos. En suma, las TOP Guideliness son una iniciativa que contribuye a la apertura en la ciencia cambiando los requerimientos de las revistas.

Los dos ejemplos entregados (la taxonomía y las TOP Guideliness) permiten comprender con un poco más de detalle el concepto de transparencia. Sin embargo, aun queda la interrogante del por qué. ¿Cuál es la problemática que lleva a considerar necesaria la adopción de la transparencia? En breve, la respuesta es la crisis en las ciencias, y específicamente en las ciencias sociales.



Summary of the eight standards and three levels of the TOP guidelines				
Levels 1 to 3 are increasingly stringent for each standard. Level 0 offers a comparison that does not meet the standard.				
	LEVEL 0	LEVEL 1	LEVEL 2	LEVEL 3
Citation standards	Journal encourages citation of data, code, and materials—or says nothing.	Journal describes citation of data in guidelines to authors with clear rules and examples.	Article provides appropriate citation for data and materials used, consistent with journal's author guidelines.	Article is not published until appropriate citation for data and materials is provided that follows journal's author guidelines.
Data transparency	Journal encourages data sharing—or says nothing.	Article states whether data are available and, if so, where to access them.	Data must be posted to a trusted repository. Exceptions must be identified at article submission.	Data must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Analytic methods (code) transparency	Journal encourages code sharing—or says nothing.	Article states whether code is available and, if so, where to access them.	Code must be posted to a trusted repository. Exceptions must be identified at article submission.	Code must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Research materials transparency	Journal encourages materials sharing—or says nothing.	Article states whether materials are available and, if so, where to access them.	Materials must be posted to a trusted repository. Exceptions must be identified at article submission.	Materials must be posted to a trusted repository, and reported analyses will be reproduced independently before publication.
Design and analysis transparency	Journal encourages design and analysis transparency or says nothing.	Journal articulates design transparency standards.	Journal requires adherence to design transparency standards for review and publication.	Journal requires and enforces adherence to design transparency standards for review and publication.
Preregistration of studies	Journal says nothing.	Journal encourages preregistration of studies and provides link in article to preregistration if it exists.	Journal encourages preregistration of studies and provides link in article and certification of meeting preregistration badge requirements.	Journal requires preregistration of studies and provides link and badge in article to meeting requirements.
Preregistration of analysis plans	Journal says nothing.	Journal encourages preanalysis plans and provides link in article to registered analysis plan if it exists.	Journal encourages preanalysis plans and provides link in article and certification of meeting registered analysis plan badge requirements.	Journal requires preregistration of studies with analysis plans and provides link and badge in article to meeting requirements.
Replication	Journal discourages submission of replication studies—or says nothing.	Journal encourages submission of replication studies.	Journal encourages submission of replication studies and conducts blind review of results.	Journal uses Registered Reports as a submission option for replication studies with peer review before observing the study outcomes.

Figure 4.1: Variaciones por dimensión de transparencia



## Chapter 5

# Palabras finales

Hemos aprendido sobre transparencia y reproducibilidad.



# Bibliography

Abril Ruiz, A. (2019). *Manzanas podridas: Malas prácticas de investigación y ciencia descuidada*.

Aczel, B., Szaszi, B., Sarafoglou, A., Kekecs, Z., Kucharský, Š., Benjamin, D., Chambers, C. D., Fisher, A., Gelman, A., Gernsbacher, M. A., Ioannidis, J. P., Johnson, E., Jonas, K., Kousta, S., Lilienfeld, S. O., Lindsay, D. S., Morey, C. C., Munafò, M., Newell, B. R., Pashler, H., Shanks, D. R., Simons, D. J., Wicherts, J. M., Albarracin, D., Anderson, N. D., Antonakis, J., Arkes, H. R., Back, M. D., Banks, G. C., Beevers, C., Bennett, A. A., Bleidorn, W., Boyer, T. W., Cacciari, C., Carter, A. S., Cesario, J., Clifton, C., Conroy, R. M., Cortese, M., Cosci, F., Cowan, N., Crawford, J., Crone, E. A., Curtin, J., Engle, R., Farrell, S., Fearon, P., Fichman, M., Frankenhuis, W., Freund, A. M., Gaskell, M. G., Giner-Sorolla, R., Green, D. P., Greene, R. L., Harlow, L. L., de la Guardia, F. H., Isaacowitz, D., Kolodner, J., Lieberman, D., Logan, G. D., Mendes, W. B., Moersdorf, L., Nyhan, B., Pollack, J., Sullivan, C., Vazire, S., and Wagenmakers, E.-J. (2020). A consensus-based transparency checklist. *Nature Human Behaviour*, 4(1):4–6.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, 533(7604):452–454.

Breznau, N. (2021). Does Sociology Need Open Science? *Societies*, 11(1):9.

Carey, B. (2011). Fraud Case Seen as a Red Flag for Psychology Research. *The New York Times*.

Christensen, G. S., Freese, J., and Miguel, E. (2019). *Transparent and Reproducible Social Science Research: How to Do Open Science*. University of California Press, Oakland, California.

Crüwell, S., van Doorn, J., Etz, A., Makel, M. C., Moshontz, H., Niebaum, J., Orben, A., Parsons, S., and Schulte-Mecklenbeck, M. (2018). 7 Easy Steps to Open Science: An Annotated Reading List.

Elliott, K. C. (2020). A Taxonomy of Transparency in Science. *Canadian Journal of Philosophy*, pages 1–14.

- Fanelli, D. (2009). How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. *PLOS ONE*, 4(5):e5738.
- Franco, A., Malhotra, N., and Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, 345(6203):1502–1505.
- John, L. K., Loewenstein, G., and Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science*, 23(5):524–532.
- Motta, M. (2018). The Dynamics and Political Implications of Anti-Intellectualism in the United States. *American Politics Research*, 46(3):465–498.
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S., Breckler, S., Buck, S., Chambers, C., Chin, G., and Christensen, G. (2014). Transparency and Openness Promotion (TOP) Guidelines.
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., Ishiyama, J., Karlan, D., Kraut, A., Lupia, A., Mabry, P., Madon, T., Malhotra, N., Mayo-Wilson, E., McNutt, M., Miguel, E., Paluck, E. L., Simonsohn, U., Soderberg, C., Spellman, B. A., Turitto, J., VandenBos, G., Vazire, S., Wagenmakers, E. J., Wilson, R., and Yarkoni, T. (2015). Promoting an open research culture. *Science*, 348(6242):1422–1425.
- Nosek, B. A., Ebersole, C. R., DeHaven, A. C., and Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, 115(11):2600–2606.
- O’Boyle, E. H., Banks, G. C., and Gonzalez-Mulé, E. (2017). The Chrysalis Effect: How Ugly Initial Results Metamorphosize Into Beautiful Articles. *Journal of Management*, 43(2):376–399.
- Simmons, J. P., Nelson, L. D., and Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, 22(11):1359–1366.
- Simonsohn, U., Nelson, L. D., and Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, 143(2):534–547.
- Steneck, N. H. (2006). Fostering integrity in research: Definitions, current knowledge, and future directions. *Science and Engineering Ethics*, 12(1):53–74.
- Thibodeaux, J. (2016). Production as social change: Policy sociology as a public good. *Sociological Spectrum*, 36(3):183–190.