



FLACSO
ARGENTINA

ÁREA DE ÉTICA, DERECHOS Y BIENES PÚBLICOS GLOBALES

PROGRAMA DE BIOÉTICA

MAESTRIA EN BIOÉTICA

Trabajo final: Conducta en Investigación desde la Ética de la
Transparencia: Una Discusión sobre las Prácticas Cuestionables en
Investigación.

Estudiante/Autor: PABLO EZEQUIEL FLORES KANTER

Director: MARIANO MOSQUERA

Trabajo final para optar por el grado académico de Magister en Bioética

Fecha: /0/2023

Conducta en Investigación desde la Ética de la Transparencia: Una Discusión sobre las Prácticas Cuestionables en Investigación.

Resumen

En un contexto de baja credibilidad científica y críticas al quehacer científico, la discusión se ha focalizado en analizar las conductas cuestionables en investigación y proponer taxonomías pertinentes para la evaluación del comportamiento científico. Sin embargo, la gran parte de estas definiciones y taxonomías carecen de un abordaje ético en su formulación. El objetivo principal de este trabajo es agregar la perspectiva ética en las discusiones en torno a este tópico. El presente trabajo culmina con la propuesta de una taxonomía que contempla el principio ético general de transparencia como aspecto central en su desarrollo, y que pretende ser superadora en relación con los conceptos sobre conductas cuestionables en investigación y taxonomías previamente propuestas. Con todo, se espera que el presente trabajo aporte mayor claridad a los fines de evaluar la propia conducta en investigación y la de nuestros pares; pero que, sobre todo, promueva un giro hacia una forma de hacer ciencia más íntegra, y más responsable.

Palabras Claves: Ética en Investigación; Conducta en Investigación Éticamente Responsable; Integridad Científica.

Aunque el t3pico no es enteramente nuevo (Barber, 1976), hoy m3s que en otras 3pocas parece haber una preocupaci3n creciente sobre el quehacer cient3fico y sus implicancias. Esta preocupaci3n lleva a cuestionamientos acerca de la forma en las que se hace (hacemos) investigaci3n (Gelman & Loken, 2016; Nelson et al., 2018; Nosek et al., 2022). En este contexto se han empezado a utilizar t3rminos como los de “grados de libertad del investigador” y “pr3cticas cuestionables en investigaci3n” (Manapat et al., 2022; Simmons et al., 2011). Si bien se han desarrollado muchos trabajos sobre estos temas, actualmente no se ha alcanzado un consenso respecto a: 1- el significado de estos conceptos y 2- qu3 conductas denominar como cuestionables y porqu3. En lo personal, veo que las taxonom3as disponibles que hacen referencia a la integridad cient3fica, y particularmente asociada a las pr3cticas cuestionables de investigaci3n, no discuten de manera directa las consideraciones 3ticas involucradas. Esta limitaci3n impide clarificar lo cuestionable de las conductas que se han descrito en la literatura, pero adem3s impiden entender el porqu3 de su denominaci3n como cuestionables. El objetivo principal de este trabajo es agregar la perspectiva 3tica en las discusiones en torno a este t3pico. Al incorporar conceptos 3ticos se espera que el presente trabajo permita, principalmente, dar mayor claridad a las actuales taxonom3as y conceptos que se manejan. As3 la intenci3n final es ayudar a las y los investigadoras/es de la regi3n a tener mayores recursos para juzgar la propia conducta y la de sus pares.

Me parece oportuno entonces comenzar por (a) describir las distintas maneras en que se ha definido el t3rmino de conducta cuestionable en investigaci3n, y (b) mostrar las taxonom3as que se han propuesto a la fecha a los fines de categorizar estas conductas en investigaci3n. A partir de esto incorporar3 algunos conceptos 3ticos que anteriormente

no se han considerado de manera explícita, e intentaré reordenar a partir de estas incorporaciones la información previa con el objetivo de alcanzar una taxonomía más clara sobre el quehacer científico.

Definiciones de Conductas Cuestionables en Investigación.

Como mencione, no existe una definición única del concepto de conductas o prácticas cuestionables en investigación (de aquí en adelante las mencionaré como QRP por sus siglas en inglés). Sin embargo, es posible al menos distinguir dos grandes grupos de definiciones: 1- aquellas que identifican a las QRP en función de si aumentan la probabilidad de reportar un “falso positivo”; y 2- aquellas que identifican las QRP principalmente con prácticas de reporte deficientes.

En las definiciones que identifican a las QRP en función de si aumentan la probabilidad de reportar “falsos positivos”, tenemos como principales ejemplos los trabajos descritos en Ioannidis (2005, 2012), John et al. (2012), Bouter et al. (2016), Munafò et al. (2017), y Sacco et al. (2019). Así, encontramos en estos autores el uso de términos como “amenazas para una ciencia reproducible” (Munafò et al., 2017), “ciencia poco rigurosa” (Bouter et al., 2016) o (baja) “credibilidad científica” (Ioannidis, 2012). Pero veamos a continuación más en detalle algunas de estas definiciones.

En John et al. (2012), no son explícitamente definidas como conductas cuestionables, pero se las describe como un área gris en las prácticas aceptadas (“gray area of acceptable practice”). Se ubican en un área gris porque estas prácticas incrementan las chances de falsos positivos en las pruebas de significancia estadística (incrementan las

chances de que la hipótesis nula sea incorrectamente rechazada: error tipo 1). Los autores identifican algunas de estas prácticas: 1. En un artículo, no reportar todas las variables dependientes medidas; 2. Decidir si recolectar más datos en función de si se alcanzó un resultado estadísticamente significativo; 3. En un artículo, no reportar todas las condiciones del estudio; 4. Frenar la recolección de datos más tempranamente que lo planeado dado que se encontró el resultado que se estaba buscando; 5. En un artículo hacer un mal redondeo de un valor p : por ejemplo, reportar que un valor p de .054 es menor a .05; 6. En un artículo, reportar selectivamente los estudios que han funcionado; 7. Decidir cuándo excluir datos luego de mirar el impacto que tienen en los resultados; 8. En un artículo, reportar resultados inesperados como habiendo sido predichos desde el principio; 9. En un artículo, afirmar que los resultados no están afectados por las variables demográficas cuando eso no es seguro o se sabe que sí; y 10. Falsificar datos.

Sacco et al. (2018) plantean que las QRP son cuestionables en parte porque ellas puedan potencialmente incrementar la probabilidad de generar falsos-positivos. Similarmente, Munafò et al. (2017) habla de las amenazas para una ciencia reproducible. Una de las amenazas está en los sesgos del investigador, en la tendencia a ver estructura en aquello que es azaroso, lo que aumenta las chances de llegar a conclusiones falsas. Se menciona que esta sobreinterpretación del “ruido” es facilitada por la rapidez, automaticidad y flexibilidad en el análisis de datos. Los autores afirman que, en una gran base de datos existen cientos o miles de alternativas de abordajes de análisis de datos razonables; llevando estas múltiples posibilidades analíticas a aumentar las chances de generar resultados “falsos-positivos”. Los autores mencionan que resulta difícil para los investigadores reconocer el P-hacking (i.e., la realización de análisis

múltiples a los fines de encontrar un resultado estadísticamente significativo) porque el sesgo de confirmación los lleva a aceptar aquellos resultados que se ajustan más a sus expectativas. Surgen luego hipótesis que se ajustan a dichos resultados y son luego reportadas sin indicar o reconocer su origen a posteriori (lo que se conoce como HARKing). Se afirma que el no control de este proceder incrementa de manera dramática el ratio de falsos descubrimientos.

Para Bouter et al. (2016), las QRP son “ofensas menores” que suelen denominarse también como “ciencia poco rigurosa”. Así, según estos autores las QRP refieren principalmente a la transgresión de principios metodológicos que amenazan contra la validez y relevancia de la investigación; pudiendo ser cometidas de manera intencional o inintencionalmente. Finalmente, Ioannidis (2005; 2012), al explicar porque la mayoría de los resultados son falsos (Ioannidis, 2005) o al hablar de la trayectoria de la credibilidad científica (Ioannidis, 2012), presenta datos sobre la proporción de resultados científicos que son correctos. Si bien el autor tampoco define explícitamente las QRP, menciona una serie de conductas que están en la base del déficit en la credibilidad y del impedimento del alcance de resultados más confiables y una ciencia autocorrectiva, entre las que destaco: el sesgo de publicación; la flexibilidad en la recolección y análisis de datos; no reportar todos los resultados; fabricar resultados; otras QRP (citando aquí al trabajo ya descrito de John et al., 2012); y estudios con baja potencia estadística.

Como dije, este primer grupo de definiciones se caracteriza por identificar las QRP en función de si aumentan la probabilidad de publicar resultados falsos. Sin embargo, no hay consistencia en las conductas que se identifican como QRP, presentando conductas que son muy diferentes entre sí. Algunas de estas QRP refieren a decisiones

metodológicas: potencia estadística, flexibilidad en la recolección y análisis de datos, el P-hacking, o decidir si recolectar más datos o excluir datos en función de los resultados. Mientras que otras de estas QRP refieren al reporte en el artículo: no reportar todos los resultados, HARKing, no reportar todas las variables medidas, o no reportar otros aspectos relevantes del procedimiento.

Tampoco se especifica, o no hay consistencia en el mejor de los casos, respecto de si todas estas conductas están en el mismo nivel de incumplimiento o gravedad. Por ejemplo, Bouter et al. (2016) se basa en la taxonomía de Steneck (2006) al afirmar que existe un espectro que permite distinguir entre malas prácticas en investigación “menores” y “mayores”. En el extremo de este espectro están las malas conductas en investigación “mayores”, claramente erradas e intencionales en el mayor de los casos: fabricación, falsificación y plagio. Estas se distinguen de las “ofensas menores” (ubicadas en un punto medio del continuo entre la conducta responsable en investigación y las malas conductas) que suelen denominarse QRP o “ciencia poco rigurosa”. Mientras que, en el caso de Ioannidis (2012) y John et al. (2012) no se hace de manera explícita esta diferenciación, y se ubican bajo el mismo paraguas de conductas las prácticas de falsificación y fabricación de datos. Más aún, otros autores parecen hacer depender del contexto la consideración de las QRP como más o menos buenas o malas conductas. Así, Sacco et al. (2018), entendiendo las QRP como flexibilidad en el análisis y recolección de datos o usando el término más general de “grados de libertad del investigador”, mencionan que existen muchas situaciones en donde la utilización de las QRP no solo puede estar justificada, sino aún ser necesaria.

Paso ahora al otro grupo de definiciones, aquellas que identifican las QRP principalmente con prácticas de reporte deficientes. En este caso tenemos los ejemplos de Simmons et al., (2011), Wigboldus y Dotsch (2016), Chin et al. (2021), y Stefan y Schönbrodt (2022). Stefan y Schönbrodt (2022) hablan de las QRP en términos de explotar los grados de libertad del investigador (en el diseño, análisis y reporte) con el objetivo de obtener un producto publicable. En Simmons et al. (2011), se afirma que el no reporte de la flexibilidad en la recolección y análisis de datos permite a los investigadores presentar casi cualquier resultado como estadísticamente significativo: “In fact, it is unacceptably easy to publish “statistically significant” evidence consistent with any hypothesis.” Es aquí donde surge el término grados de libertad del investigador (researcher degrees of freedom), y los autores lo utilizan para referirse a la gran cantidad de decisiones que el investigador debe tomar en el curso de la investigación (e.g., recolectar más datos; excluir observaciones; combinar y comparar condiciones; variables a ser consideradas; combinación y transformación de mediciones). El problema según estos autores, y aquello que puede identificarse como QRP, está en el reporte: “It is rare, and sometimes impractical, for researchers to make all these decisions beforehand. Rather, it is common (and accepted practice) for researchers to explore various analytic alternatives, to search for a combination that yields “statistical significance,” *and to then report only what “worked.”* (el resaltado es mío). Entonces no parece ser tanto los “grados de libertad del investigador”, sino más bien explotar esto en favor de publicar un resultado que cumple las expectativas del investigador/a o editor (en el caso del sesgo de publicación), lo que conlleva un aumento en el ratio de falsos positivos: “...flexibility in data collection, analysis, and reporting dramatically increases actual false-positive rates...” Finalmente

los autores afirman que estas QRP no son el producto de una intención maliciosa de engañar.

En el caso de Wigboldus y Dotsch (2016), los autores reanalizan las prácticas descritas en John et al. (2012), afirmando que algunas de estas son más que cuestionables. Por ejemplo, falsificar datos y presentarlos como reales es simplemente una mala conducta, difícilmente podremos catalogarlo como cuestionable según estos autores. Mencionan que distinto es el caso de otras conductas, en donde son más cuestionables en términos de reporte que en términos de prácticas. Los autores toman como ejemplo una de las prácticas cuestionables mencionadas en John et al. (2012), la práctica de “Decidir cuándo recolectar más datos luego de ver si los resultados son estadísticamente significativos”. Según Wigboldus y Dotsch (2016) esta no es una práctica cuestionable, más allá de que pueda inflar las chances de cometer un error tipo 1 (la ocurrencia de un falso positivo). Lo que hace cuestionable esta práctica según los autores es fallar en reportar en el paper que efectivamente estos pasos se siguieron. Lo cuestionable entonces está en el reporte¹. Al no reportar esta toma de decisión, los lectores son incapaces de evaluar correctamente el valor de la evidencia en los resultados reportados. Más aún, los autores afirman que si esta toma de decisiones que es en esencia exploratoria se presenta como si hubiera sido preespecificada de antemano (se reporta como confirmatoria), la práctica ya no es siquiera cuestionable sino enteramente mala. En este último caso, los autores parecen distinguir entre errores “honestos” en el reporte de la clara intención de engañar con el reporte. Interesantemente, estos autores

¹ En este sentido, los autores coinciden con la definición de QRP descrita por Chin et al. (2021): QRP refieren a una serie de prácticas que, *cuando no se reportan transparentemente*, distorsionan la precisión del reporte (e.g., exagerando los tamaños del efecto o produciendo resultados estadísticamente significativos).

distinguen entre conducta fraudulenta y conducta cuestionables de las conductas vinculadas a malas formas de hacer ciencia. Para esto describen un ejemplo en donde un investigador decide que su percepción es suficiente para dar cuenta de la realidad y es más precisa de lo que pueden hacerlo los potenciales participantes. Debido a esto, decide completar por sí solo miles de cuestionarios. Luego reporta sus resultados en el manuscrito. En el método, el investigador involucrado hace totalmente transparente su proceder indicando precisamente cómo obtuvo los datos. Wigboldus y Dotsch (2016) mencionan que, si bien esto es sin duda una mala forma de hacer ciencia, no puede catalogarse a la conducta (ni al reporte) del investigador como fraudulenta(o) ni como cuestionable; por el contrario, el investigador ha informado de manera transparente (abierta) y veraz cómo se obtuvieron y analizaron sus datos. Acá los autores proponen entonces dos aspectos distintos que hay que juzgar en el proceder del investigador: a- si es veraz, abierto-transparente en su reporte (área que compete a la identificación de conductas cuestionables o malas conductas); b- si realiza buena ciencia (área que compete a la identificación de la aplicación pertinente y correcta de los métodos). En lo que concierne al análisis de datos, Wigboldus y Dotsch (2016) plantean que no puede hablarse allí de prácticas cuestionables. La parte cuestionable resulta en no informar veraz y explícitamente el carácter-naturaleza-forma de estos análisis (exploratorio vs confirmatorio, por ejemplo): “It is not the research practice that is potentially questionable; it is the reporting practice that is potentially questionable”. Los autores también plantean usar el término “prácticas cuestionables” solo para referirse a las prácticas no éticas como falsificar datos, dar un trato no ético a los participantes o maltratar animales experimentales; y, en realidad, proponen que se denominen directamente como “mala

práctica”. La mayor parte de lo englobado como “área gris” en el análisis de datos (John et al., 2012) debe denominarse según estos autores como “práctica de reporte cuestionable”. La buena práctica de reporte consiste así en reportar de manera explícita y honesta la forma en que se procedió.

Los trabajos referenciados con anterioridad hacen entonces un recorte un tanto más específico de aquellas conductas que deberían catalogarse como QRP. Lo que termina definiendo esto último es la práctica de reporte: si un reporte es transparente-abierto no se define como QRP, mientras que si no es transparente-abierto se identifica como QRP. Algunos como Simmons et al. (2011) siguen poniendo el foco en las consecuencias que esto trae para la publicación de resultados falsos; mientras que Wigboldus y Dotsch (2016) no hacen referencia a esa consecuencia o la ponen en un plano diferente: independientemente de que puedan potenciar el error tipo 1 (falsos positivos), lo cuestionable está en si se reporta de manera transparente y abierta o no. Wigboldus y Dotsch (2016) parecen ir más allá, cuando hacen la distinción entre el reporte y la calidad metodológica de lo reportado. Puede decirse así que los autores definen a las QRP no en base a la calidad metodológica de las prácticas sino, exclusivamente, a la calidad del reporte. Entiendo que por esa razón sugieren cambiar el término QRP por “conductas de reporte cuestionable”. Sumado a todo esto, y sin tenerlo como propósito expreso, Wigboldus y Dotsch (2016) terminan ofreciendo una taxonomía de conducta científica (o al menos esta puede deducirse claramente de lo desarrollado por los autores). Volveré sobre esta taxonomía más adelante.

Concluyendo este apartado, de los antecedentes que he descrito existen algunas coincidencias que permiten agrupar las distintas definiciones dadas. Sin embargo, he

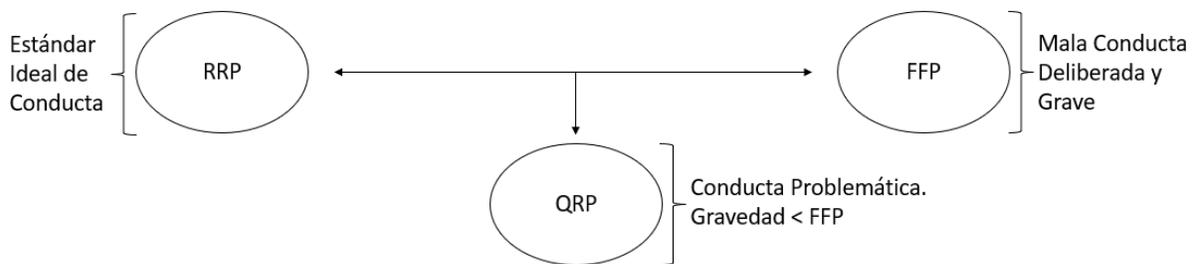
intentado mostrar que dentro de estas dos grandes definiciones de las QRP existen diferencias que resulta relevante atender. Es importante también notar que hasta aquí he citado antecedentes que se encargan principalmente de identificar las QRP, pero casi ninguno de estos se ha propuesto desarrollar o mencionar una taxonomía que no solo identifique estos comportamientos, sino que además los distinga de otros tipos de conductas más o menos graves, más o menos ideales. A continuación, desarrollaré las taxonomías que a la fecha se han propuesto en relación a las QRP.

Prácticas Cuestionables en Investigación (QRP): Taxonomías.

Una de estas taxonomías ha sido descrita por Steneck (2006). El autor propone que la conducta en investigación puede ser entendida como un continuo, desde el estándar institucional ideal, denominada Conducta Responsable en Investigación (RRP por sus siglas en inglés), a la peor conducta caracterizada por prácticas de Fabricación, Falsificación y Plagio (FFP) (véase figura 1). Respecto al polo FFP, la mayoría de los interesados acuerdan en que estos comportamientos deben ser evitados. En el polo opuesto, la RRP refiere a conducir la investigación en coherencia con las responsabilidades profesionales del investigador; responsabilidades que son definidas por las organizaciones profesionales, instituciones académicas involucradas, y el gobierno. En el medio de este continuo se ubican las Conductas Cuestionables en Investigación (QRP para seguir con la misma denominación de los párrafos anteriores). En otras palabras, si bien las QRP conllevan ciertos problemas para la investigación y sus objetivos, no pueden identificarse como buenas prácticas o coherentes con los estándares institucionales, pero tampoco se consideran tan graves como las FFP.

Contrariamente a las FFP, las QRP por su carácter de gravedad menor no ameritan, según este autor, acciones-sanciones gubernamentales. En este contexto las QRP son definidas como acciones que van en contra de los estándares ideales y que probablemente sean perjudiciales para el proceso de investigación.

Figura 1. Taxonomía de Steneck (2006).



Los ejemplos que Steneck (2006) pone como QRP son los siguientes: no presentar de manera precisa y honesta el aporte que el grupo de investigadores/as ha realizado en la investigación reportada; presentar errores y equivocaciones en el reporte, tales como error en notas y bibliografías, o resumir de manera inadecuada o errónea los resultados y conclusiones en el resumen; no proveer información suficiente sobre el método, que posibiliten a un investigador externo evaluar o replicar los resultados; y presentar sesgos, definidos como la toma de decisiones en la investigación basadas en razones ajenas a lo científico-académico. Otras QRP mencionadas en el trabajo de Steneck (2006) son: no declarar adecuadamente conflictos de interés vinculados a los resultados de la investigación y la agencia financiadora del proyecto; no presentar datos que contradigan las investigaciones previas del investigador/a; ocultar detalles de la metodología o los

resultados en un paper o propuesta; usar un diseño de investigación inadecuado o inapropiado; sobreinterpretación de resultados estadísticamente significativos en estudios pequeños; reporte selectivo del método, análisis y resultados; resultados negativos no publicados; análisis inadecuados de subgrupos; análisis post-hoc no admitidos; y análisis efectuados por el espónsor de la investigación.

Vemos que si bien Steneck (2006) se ocupa de ofrecer una taxonomía de conducta científica, disparidades que notábamos en las definiciones sobre las QRP en el apartado anterior siguen siendo notorias aquí. Entre estas destaco las siguientes. Si bien se distingue entre FFP y QRP, no está claro lo que termina diferenciando una de otra. Por ejemplo, en QRP se ubican comportamientos como ocultar detalles de la metodología o los resultados y no presentar de manera precisa y honesta el aporte del equipo de investigación involucrado, conductas que, al igual que las FFP, son deshonestas o tienen una clara intencionalidad de engañar. Por lo tanto, no termina de quedar claro el criterio para que algunas de estas conductas se ubiquen en el extremo de la “peor conducta” mientras otras queden en un lugar intermedio. Otro aspecto a resaltar es que Steneck (2006) no distingue tampoco entre las prácticas de reporte y las practicas metodológicas. Para el autor prácticas bien diferentes como cometer errores “honestos”, no proveer de información suficiente, la realización de análisis inadecuados, y la presentación deshonestas de información caen bajo el mismo rótulo de QRP.

Como una alternativa a la propuesta de que Steneck (2006), Manapat et al. (2022) proponen una nueva taxonomía. En ésta se definen las prácticas cuestionables en investigación como opciones metodológicas que tienden a alejarse de las buenas prácticas. Las QRP se definen aquí como prácticas en investigación de las que se tiene

evidencia suficiente de que son erradas o subóptimas. A estas QRP los autores contraponen el concepto de los “grados de libertad del investigador” (RDF por sus siglas en inglés). Manapat et al. (2022) hacen un uso diferencial de este término, distinto al uso dado por Simmons et al. (2011), definiéndola como una ambigüedad honesta en cómo mejor tomar una decisión metodológica, considerando metodologías alternativas aceptables y justificables. Sobre estas primeras definiciones los autores citados hacen las siguientes aclaraciones:

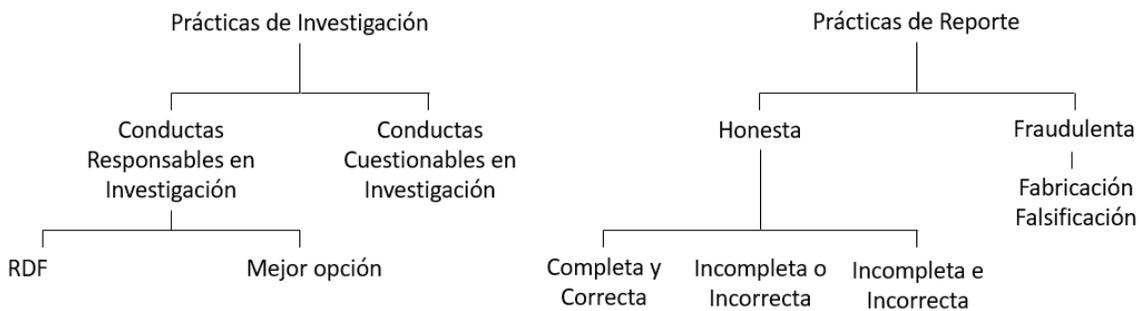
- 1- El definir si una conducta cae dentro de la categoría de QRP o RDF depende de la cantidad de evidencia que exista al momento a favor o en contra de la práctica en cuestión. La elección entre procedimientos que presentan aproximadamente la misma evidencia favorable se considera también dentro de las RDF.
- 2- En este sentido, la categorización de una práctica no es estática sino contexto dependiente. En otras palabras, si una nueva investigación provee evidencia clara en contra de una decisión metodológica, una práctica inicialmente considerada como RDF pasa a ser categorizada como QRP.
- 3- La distinción entre RDF y QRP también está basada en base a si se ha adquirido información suficiente para tomar una decisión defendible. La no capacidad de hacer una toma de decisión metodológica defendible, en función de haber adquirido información suficiente, es razón para categorizar a dicha conducta como QRP.

Un ejemplo de QRP que dan los autores es la no consideración de métodos de imputación acorde ante casos ausentes al aplicar una regresión lineal, la cual lleva a resultados sesgados y pruebas de significación estadísticas incorrectas. Podemos ver

que Manapat et al. (2022) hacen un recorte mucho más específico sobre las QRP en contraposición a Steneck (2006), considerando específicamente decisiones vinculadas al análisis de datos, que pueden, según la evidencia recopilada al momento, considerarse inadecuadas o subóptimas (y en este sentido, los autores coinciden con el primer grupo de definiciones sobre QRP: aquellas que identifican a las QRP en función de si aumentan la probabilidad de reportar un “falso positivo”).

Continuando, reproduzco a continuación el diagrama que presentan Manapat et al. (2022) para describir su taxonomía:

Figura 2. Taxonomía de Manapat et al. (2022).



Como puede visualizarse, en Manapat et al. (2022) se hace una distinción entre dos grandes categorías. Por un lado, están las prácticas de investigación, las acciones o tomas de decisiones que el investigador/a puede tomar. Aquí ubican los conceptos que han elaborado los autores sobre las QRP y las RDF. Por otro lado, ubican las prácticas de reporte, diferenciando entre un reporte honesto y fraudulento. El reporte honesto puede ir según los autores desde completo y correcto a incompleto e incorrecto; mientras

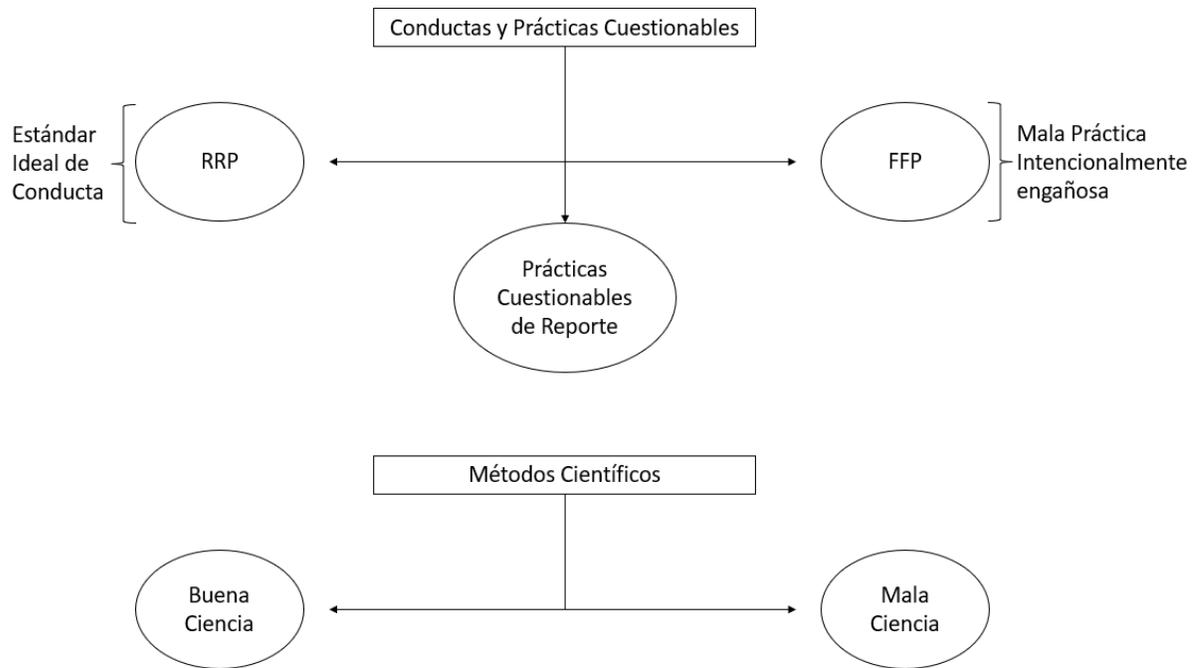
que el reporte fraudulento, esto es con intenciones de engañar y mentir, se caracteriza por conductas de fabricación y falsificación. De esta manera, la propuesta de Manapat et al. (2022) pasa de una noción gradual a una taxonomía dicotómica: o hay conducta responsable o hay QRP; o hay reporte honesto o hay reporte deshonesto.

Considero que la taxonomía de Manapat et al. (2022) es superadora en algunos puntos importantes respecto a la taxonomía previa de Steneck (2006). En primer lugar, considero que es oportuno distinguir entre prácticas de investigación y de reporte. Y, en segundo lugar, es relevante también diferenciar cuándo hay intencionalidad de engañar de cuando no la hay. Sin embargo, en la propuesta de Manapat et al. (2022), similar a lo que pasaba con la taxonomía de Steneck (2006), no se expresa un criterio que permita comprender la diferenciación entre conductas fraudulentas y las QRP, más allá de que una se asocie al reporte y otra a las prácticas de análisis de datos. Tampoco permite dar respuestas respecto a donde ubicar conductas que son deshonestas pero que no conllevan la fabricación o la falsificación, como ser el HARKing o el reporte (mal intencionadamente) selectivo. En otras palabras, el paso de una taxonomía gradual a una dicotómica conlleva dejar de lado (malas) conductas muy presentes en el contexto de investigación.

Finalmente, mencionaré la taxonomía que puede derivarse del trabajo de Wigboldus y Dotsch (2016), sin que haya sido objetivo propio de los autores presentar una taxonomía alternativa. De lo desarrollado por estos autores se deriva una taxonomía continua, similar en estos términos a la de Steneck (2006), con polos que diferencian extremos y puntos intermedios. En el caso de Wigboldus y Dotsch (2016) la conducta de falsificar datos y presentarlos como reales se considera de manera categórica una mala conducta.

En esta categoría de mala conducta ubican otros comportamientos vinculados con el reporte intencionalmente selectivo y engañoso, como ser el reportar que fue exploratorio como si hubiera seguido un abordaje confirmatorio. Otras conductas asemejadas habitualmente con lo que se denomina P-Hacking, tales como “Decidir cuándo recolectar más datos luego de ver si los resultados son estadísticamente significativos”, no son ni malas conductas ni cuestionables según estos autores. Lo cuestionable está en el reporte, y por ello ubicarían en el medio de este continuo lo que denominan Prácticas Cuestionables en el Reporte, que implica no brindar información completa sobre el procedimiento seguido, y en donde la intención no es engañar. A diferencia de las taxonomías anteriores, para Wigboldus y Dotsch (2016) los aspectos fraudulentos y cuestionables (del reporte) en el proceder del investigador deben ser analizados de manera separada a la calidad científica de lo reportado. Con todo, la taxonomía derivada de Wigboldus y Dotsch (2016) podría quedar diagramada de la siguiente manera:

Figura 3. Taxonomía derivada de Wigboldus y Dotsch (2016).



Se establece entonces una diferenciación clara con las anteriores taxonomías y definiciones sobre conducta cuestionable. Aquí los autores hacen explícito que lo que es cuestionable radica en la forma de presentar la información y siempre y cuando no haya una intencionalidad de engañar. Caso contrario, la conducta debe referirse como mala práctica. Además, para estos autores las prácticas referidas a la validez o pertinencia de los métodos empleados no debe ser analizado bajo el mismo eje que las conductas cuestionables: estas refieren a un eje de análisis distinto que evalúa la calidad científica de lo reportado. Por el contrario, el eje de las conductas cuestionable incluye indagar si hay intencionalidad de engañar y si se reporta toda la información de manera transparente.

Síntesis de los Aportes Previos y Propuesta de un Abordaje Diferente.

En suma, he intentado mostrar hasta aquí qué se ha propuesto en referencia al concepto de conductas cuestionables en investigación (QRP), y sobre las taxonomías que han buscado categorizar y ordenar este tipo de comportamientos. Al hacerlo, he querido hacer notar las similitudes y diferencias entre las distintas propuestas. En el caso de los conceptos de QRP, he notado que una gran parte de estas definiciones identifica como QRP a aquellas prácticas, principalmente vinculadas al análisis de datos, que aumentan la probabilidad de falsos descubrimientos; mientras que otros aportes se focalizan más en el reporte, haciendo notar que lo cuestionable no se ubica tanto en las prácticas (sea por caso el análisis de datos), sino en la manera en que esas prácticas son reportadas. Sin embargo, ninguna de estas definiciones hace alusión explícita a aspectos éticos ni principios involucrados que permitan analizar las conductas en consideración. Si bien la honestidad, más precisamente la intencionalidad de engañar, es mencionada en algunos casos para hacer referencia a las malas conductas como la falsificación y la fabricación, no hay un abordaje específico ni claro respecto al lugar que ocupa la ética y los principios éticos en la definición de estas conductas. Algo similar puede concluirse respecto a las taxonomías mencionadas. La única de ellas que hace alusión explícitamente a la ética es la de Steneck (2006), pero considero que no lo hace con fines de profundizar en los aportes de la ética, sino a los fines de justificar la manera de evaluar-medir los comportamientos. Para obtener medidas concretas de estas QRP, Steneck (2006) menciona que es preciso centrarse en lo que denomina integridad profesional dentro del estudio de la integridad científica: la adherencia a los estándares profesionales (guías claras respecto a qué debe hacer el investigador según las instituciones involucradas). En este sentido el autor considera preciso distinguir entre las preguntas o debates acerca

de lo que el investigador debe hacer (lo que según el autor corresponde a la ética en investigación) de las guías, más o menos precisas, de lo que el investigador debe hacer (lo que se corresponde con la integridad profesional en investigación). Para comprender esta distinción entre principios morales y estándares profesionales, y sus roles diferenciados en la investigación, el autor da el siguiente ejemplo: Una obligación moral de “ser sincero y veraz”, aunque esté respaldado por un código profesional, no funciona en la vida profesional de la misma manera que un requerimiento institucional acerca de “registrar y reportar los datos de manera precisa en un registro fechado y firmado”. De nuevo, es preciso resaltar que esta distinción la realiza el autor a los fines de obtener mediciones concretas acerca de las QRP y poder evaluarlas: según Steneck (2006) es más fácil evaluar-medir adherencia a normas que disposiciones morales (veracidad, honestidad, sinceridad, justicia). Como se aprecia, no es que Steneck (2006) afirme que su taxonomía no compete a la ética; en efecto, lo que termina definiendo como requerimientos institucionales puede ser entendido como la concretización, como él mismo dice, de principios éticos. Pero si es cierto que más allá de esta justificación que da el autor, no hay mayor profundización en la ética ni sus implicancias en la taxonomía propuesta.

Con todo, considero fundamental centrar la discusión en los aspectos éticos para alcanzar el objetivo de lograr mayor claridad y precisión respecto a la forma de definir y categorizar las conductas en investigación. A continuación, desarrollaré algunas nociones éticas que creo centrales para estos fines.

Ética en Investigación: Aportes Conceptuales para el Abordaje de la Conducta Científica.

Es posible identificar ciertos estándares (principios-reglas) éticos en el quehacer científico. Estos estándares se basan sobre dos fundamentos conceptuales, la ciencia y la moral (Resnik, 2005; Shamoo & Resnik, 2009; Tijdink et al., 2021). De esta forma, la conducta ética en ciencia implica a) no violar los comúnmente aceptados estándares morales y b) promover el avance de las metas científicas (como ser, la búsqueda del conocimiento y la resolución de problemas prácticos). Ambos aspectos se encuentran íntimamente ligados en la práctica y es difícil (o mejor dicho no es posible) separarlos (Bosma & Granger, 2022). Por ejemplo, el fraude es una conducta no ética porque implica mentir y engañar, lo que es moralmente incorrecto; pero, además, porque las conductas fraudulentas como la fabricación de datos promulga errores y promueve una falta de confianza en los resultados científicos (Shamoo & Resnik, 2009). Si bien las conductas fraudulentas pueden no provocar un daño directo en los sujetos humanos, pueden de igual manera generar un impacto negativo en la confianza social de la ciencia si en base a estas conductas se producen resultados que no sirven o que implican derrochar los recursos públicos (Resnik, 2018). Sin embargo, es importante tener presente que, dado que se presentan datos o resultados fraudulentos, errados o sesgados, estos pueden causar daños en la sociedad o incrementar el riesgo de producir daños en las personas.

Entre los principios éticos de la ciencia (entre los que se encuentran la libertad, el crédito, la responsabilidad social, la legalidad, el respeto mutuo, y el respeto por los sujetos) aquí me interesa centrarme en tres principios éticos fundamentales, la honestidad, la prudencia, y la apertura (Bosma & Granger, 2022; Resnik, 2005; Shamoo & Resnik, 2009;

Tijdink et al., 2021)². Considero que estos principios están en el centro de los debates actuales acerca de la falta de credibilidad de la ciencia y la baja reproducibilidad de los hallazgos científicos; y permiten llegar a una conceptualización y taxonomía más clara del quehacer científico. Además, estos principios están también asociados fuertemente con el hincapié que se hace actualmente en las prácticas de ciencia abierta para la superación de los problemas de confianza y reproducibilidad que enfrenta la ciencia.

Volviendo a los principios sobre los que me interesa focalizarme, la honestidad como principio indica que el científico no debe fabricar, falsificar, o describir engañosamente los datos o resultados. El investigador debe ser objetivo, no sesgado, y veraz en todos los aspectos del proceso de investigación (Resnik, 2005). Cuando se habla de deshonestidad en ciencia es relevante distinguir entre deshonestidad y errores “honestos”. Si bien ambos producen consecuencias similares obstaculizando el alcance de los objetivos científicos, la deshonestidad involucra el intento de engaño (mintiendo, ocultando información, o describiendo engañosamente) a una audiencia que espera que le digan la verdad. Nótese que lo que aquí distingue una buena o mala conducta (en términos de honestidad) es la motivación o intención (de engañar o mentir) del científico.

² En este sentido, el presente trabajo no se centra en el abordaje ético como disciplina académica: aquella centrada en la filosofía moral que busca responder a antiguas preguntas acerca del deber, el honor, la integridad, la virtud, la justicia, el buen vivir, entre otras. En otras palabras, aquí no interesa hacer el foco en las teorías éticas (e.g., la ética de la virtud; el utilitarismo; el abordaje kantiano de la ética); el foco se pone aquí en los principios éticos de la ciencia: la ética como estándares de conducta que permiten la distinción entre bien y mal, correcto e incorrecto, etc. (principios que por supuesto se derivan de los desarrollos teóricos sobre la ética). En este sentido, podría afirmarse que este trabajo se centra más en la ética normativa (McCullough et al., 2004). Tampoco me centraré aquí de manera directa en los principios éticos más generales de autonomía, no maleficencia, beneficencia y justicia. Aquellos interesados en ampliar sobre las distintas teorías éticas, así como los principios más generales u otros principios vinculados a la actividad científica, pueden consultar en Resnik (2005) y Shamoo y Resnik (2009).

Por su parte, el principio de prudencia señala que el investigador debe evitar los errores en la investigación: debe minimizar los errores humanos y metodológicos, evitando el autoengaño, los sesgos y los conflictos de interés (Resnik, 2005). En este contexto es importante distinguir entre diferentes tipos de errores en el proceso de investigación: por ejemplo, errores vinculados con los instrumentos de medición, errores vinculados con la interpretación y el análisis de los datos (e.g., métodos estadísticos), errores asociados con la utilización de supuestos teóricos y los sesgos en la inferencia; entre otros. Aquí la falta de prudencia se distingue de la deshonestidad en cuanto la falta de prudencia no implica la intencionalidad de engañar. Si bien, en comparación con la deshonestidad, la falta de prudencia no es una falencia u ofensa ética tan seria, es relevante evitar estos errores “honestos” dado que pueden también desperdiciar recursos, erosionar la confianza, o llevar consecuencias sociales negativas y aún desastrosas. En este sentido, este tipo de errores pueden ser catalogados como conductas negligentes.

Finalmente, el principio de apertura indica que los científicos deben compartir los datos, resultados, ideas, técnicas y materiales (World Medical Association, 2013). Deben permitir que otros investigadores revisen su trabajo (Resnik, 2005, 2018). El principio de apertura promueve el avance del conocimiento debido a que: a) permite a los científicos revisar y criticar sus respectivos trabajos; b) ayuda a construir una atmósfera de cooperación y confianza en la ciencia; y c) posibilita que los científicos usen los recursos de manera eficiente (Tijdink et al., 2021). Sumado a lo anterior, la apertura previene a la ciencia de volverse dogmática, no crítica, y sesgada. Pero este deber no es solo científico, si no también moral. Los científicos tienen una obligación-deber moral de evitar el ocultamiento a los fines de ayudar a las personas (Bosma & Granger, 2022).

Los tres principios éticos descritos pueden englobarse en el concepto o principio más general de transparencia. Al referir a la transparencia se indica que el investigador es responsable de la completud-integridad y precisión-exactitud del reporte, y que este debe adherir a las guías aceptadas para un reporte ético de la información (Aczel et al., 2019; Bosma & Granger, 2022; Needleman et al., 2008; Nicholls et al., 2016; World Medical Association, 2013). La responsabilidad se pone, por un lado, en la integridad del individuo, el cual debe ser honesto en su investigación, evidenciándose esto en la forma de un reporte que debe ser completo y exacto-preciso. Además, los principios más generales de justicia y reciprocidad requieren que el investigador contribuya a la acumulación del cúmulo de conocimiento científico, siendo para esto último necesario que el contenido de la investigación sea utilizable. Un paso fundamental para el alcance estos objetivos es la mejora de los reportes. Presentar reportes completos y precisos es además actuar acorde con las expectativas que la sociedad tiene acerca de la conducta en las actividades de investigación. Pero la transparencia no solo implica obligaciones y valores intrínsecos o vinculados a la integridad de los investigadores, sino que también refiere a obligaciones y valores externos: reflejan los aspectos vinculados al uso final de los datos y los resultados. Como vimos anteriormente, estas obligaciones incluyen los principios del uso justo de los recursos, así como la promoción de los beneficios y la evitación de los daños o perjuicios. Es claro que un reporte deficiente impide la apropiada implementación de los resultados de la investigación. Esto puede conllevar a exponer a las personas a un riesgo innecesario o imposibilitarles de acceder a un beneficio que se pierde por el uso inapropiado de los resultados (producto de un informe deficiente). Los reportes científicos deficientes también pueden precipitar un mal uso o desperdicio de

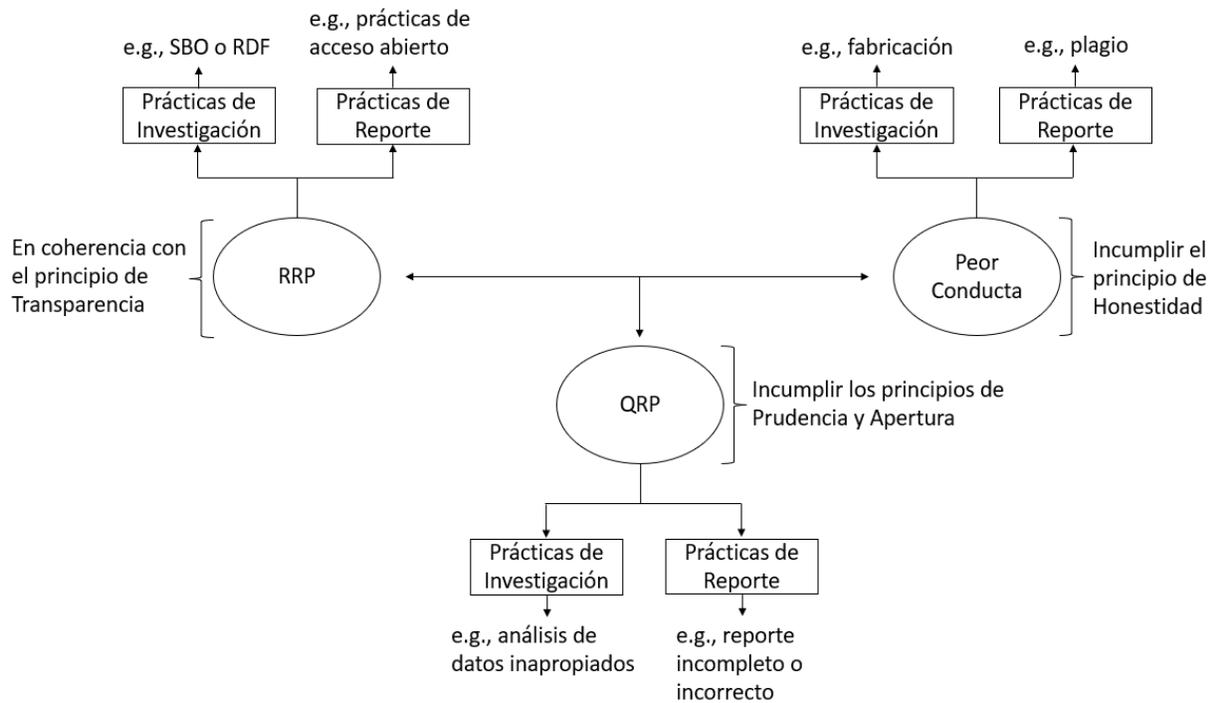
los recursos destinados con fines de reproducibilidad de los hallazgos. Es por todo esto que se afirma que las (buenas) prácticas de reporte son, principalmente, una obligación moral (Aczel et al., 2019; Bosma & Granger, 2022; Moher, 2007; Needleman et al., 2008).

Es posible concluir aquí indicando que la conducta en investigación debe analizarse bajo el principio más general de transparencia. Este principio general involucra tres principios fundamentales: la honestidad, la prudencia, y la apertura. A partir de este desarrollo ético es posible generar una taxonomía alternativa que distinga los comportamientos en investigación en base al incumplimiento de estos principios.

Conducta en Investigación: Propuesta de Taxonomía Basada en Principios Éticos.

El desarrollo ético antecedente posibilita plantear un continuo en donde se diferencian las distintas conductas en investigación (véase figura 4): 1) Fraudulentas, que incumplen la honestidad; y 2) Cuestionables, que no se alinean con los principios de Prudencia y Apertura. En (2) es posible distinguir entre a- acciones que refieren a no presentar de manera transparente y abierta los procesos de la investigación y b- emplear métodos de manera inadecuada.

Figura 4. Taxonomía basada en la Ética de la Transparencia.



A diferencia de los aportes precedentes, esta propuesta taxonómica es clara con respecto a los principios éticos en los que se fundamenta y permite distinguir entre lo que está mal (ser deshonesto), entre lo que es cuestionable o malo, pero no tan grave como engañar (no ser transparente en el reporte), y entre lo que no es ni malo ni cuestionable (probar múltiples análisis, aplicar abordajes exploratorios)³. También permite entender qué se espera del investigador en términos de conducta íntegra o responsable. Lo que los/as investigadores/as deben entender también es que si bien uno puede ubicar en distintas partes del continuo la “gravedad moral” de estas conductas (en términos morales, faltar a la verdad es más grave que no mostrar una información o equivocarse debido a un “error honesto”), en términos de consecuencias e implicancias éticas estas

³ Esta nueva propuesta está muy en línea con el reciente trabajo de Bosma y Granger (2022), quienes proponen abordar las prácticas vinculadas a la transparencia desde una perspectiva ética.

conductas tienden a converger: suelen llevar a las mismas consecuencias ya descritas de disminución de la confianza social, desperdicio de recursos y producción de daños.

La taxonomía que propongo tiene algunas coincidencias y diferenciaciones con las taxonomías previas. En coincidencia con las taxonomías de Steneck (2006) y Wigboldus y Dotsch (2016), esta taxonomía ubica las conductas en un continuo que permite diferenciar en función de la gravedad de ciertas conductas. A diferencia de las propuestas previas, aquí se hacen explícitos los principios éticos involucrados e incumplidos en cada caso. Asimismo, se toma los aportes de Manapat et al. (2022) para distinguir entre reporte y práctica, pero en la taxonomía aquí propuesta las conductas quedan agrupadas de distinta manera. Mientras en Manapat et al. (2022) las conductas vinculadas a un reporte honesto pero incompleto e incorrecto no se considera QRP, aquí sí debido al incumplimiento de los principios de Prudencia y Apertura, y en función de las consecuencias que traen aparejadas estas conductas. Sin embargo, sí coincidimos con Manapat et al. (2022) en que la utilización de métodos no pertinentes o injustificados debe ser considerado una QRP. Como estos autores, también considero que las RRP conllevan en este sentido o elegir una opción única bien establecida (SBO en el diagrama, y para usar el mismo término que en Manapat et al., 2022) o basar la decisión en una serie de procedimientos que son razonablemente justificables (RDF en el diagrama, y para usar el mismo término que en Manapat et al., 2022). De esta manera la taxonomía también coincide con la propuesta de Wigboldus y Dotsch (2016), en tanto considera QRP aquellos reportes cuestionables o deficientes; aunque a diferencia de estos autores, no consideramos que la calidad de la ciencia deba analizarse en un eje

independiente. Por el contrario, la ciencia de calidad involucra el respeto del principio ético de Prudencia.

En síntesis, la taxonomía que aquí se propone postula diferenciar las conductas de investigación considerando el cumplimiento/incumplimiento de los principios específicos derivados del principio general de Transparencia. La conducta responsable en investigación (RRP) se define así por prácticas que son coherentes con los principios de honestidad, prudencia y apertura. Esto implica el hacer uso de procedimientos acordes y justificables metodológicamente, pero además reportar de manera honesta, completa y abierta el proceder llevado a cabo y los materiales involucrados en el estudio. El incumplimiento del principio general de Transparencia conlleva consecuencias que van desde los daños y perjuicios a las personas y la sociedad, al deterioro de la confianza social en la ciencia y el desperdicio de los recursos. Finalmente, la diferenciación entre cuestionable y estrictamente malo en la conducta de investigación queda establecida por si se incumple o no el principio de honestidad.

Consideraciones Finales

Si bien actualmente se ha hecho un correcto hincapié en las (malas) formas de hacer ciencia y en la necesidad de analizar críticamente nuestro proceder para mejorar el avance y la aplicación práctica de los conocimientos científicos, la falta de consenso en las definiciones y taxonomías propuestas en esta área impiden un correcto análisis de nuestras prácticas como investigadores/as y la de nuestros pares. Esto puede generar miedos innecesarios en los investigadores noveles, así como el abandono de (buenas)

prácticas por ser en algunos casos (mal) catalogadas como cuestionables (Wigboldus & Dotsch, 2016). Aquí he intentado mejorar las conceptualizaciones y propuestas previas, tomando como aporte central a la ética y sus principios vinculados al quehacer científico.

Quisiera cerrar este aporte respondiendo al siguiente interrogante: ¿cómo puedo hacer para que mi quehacer científico se ubique, según la taxonomía propuesta, en el extremo de la conducta responsable en investigación? Para esto considero oportuno presentar el modelo de transparencia propuesto por Tuval-Mashiach (2017). El modelo propone tres preguntas básicas que deben poder responderse para alcanzar un nivel de transparencia adecuado en los reportes de investigación⁴ en cualquier área del conocimiento. De esta manera el/la investigador/a debe ser capaz de expresar de manera clara el “qué”, “cómo” y “porqué” de las decisiones tomadas en la investigación.

El “qué” es lo primero que debemos poder responder, y refiere a poder clarificar con términos precisos y sin ambigüedad lo que hemos realizado (ejemplo, tipo de método y análisis de datos-información empleado). Esta primera pregunta puede parecer simple y trivial en un primer momento, pero no lo es. El investigador/a debe hacer un gran esfuerzo aquí por presentar lo que hizo sin ambigüedad, acudiendo a ejemplos que permitan llegar a un nivel de precisión adecuado en el uso de los términos. En el caso del “cómo”, aquí hay que hacer referencia al procedimiento seguido, o, en otras palabras, detallar cuales fueron las tomas de decisiones que se tomaron a lo largo de toda la investigación. En este punto, se alcanza un grado de transparencia necesaria cuando un investigador externo o independiente puede repetir -y entender con claridad- los pasos y

⁴ Refiero al reporte en un sentido más amplio, incluyendo aquí no solo el paper o versión final del manuscrito en sí mismo, sino todo aquel material que pueda considerarse como anexo (e.g., material suplementario en un repositorio externo de acceso abierto) que desarrolle con el detalle adecuado cada una de estas preguntas.

procedimientos que se describen en un reporte. Por su parte, en el “porqué” hay que justificar las razones que nos llevaron a optar por seguir un camino -tomar una decisión dada- en el contexto de múltiples alternativas razonables posibles. Implica así que el/la investigador/a de cuenta, por ejemplo, de las razones de su elección de un método dado, y sea capaz de justificar su elección comparando métodos alternativos. Como se ve, para que el modelo de transparencia se aplique correctamente es fundamental el pensamiento crítico y el desarrollo lento y a consciencia de la investigación (Antonakis, 2023; Chin et al., 2021). Además, el modelo se basa en el supuesto del respeto del principio de honestidad. En este sentido, es relevante aclarar que, si la intención de la persona es engañar, no existe modelo ni tecnología actual disponible que impida que lo haga (Wigboldus & Dotsch, 2016).

Contamos actualmente con excelentes recursos de acceso abierto que facilitan enormemente la correcta respuesta a los tres interrogantes fundamentales descritos: el “qué”, el “cómo” y el “porqué”. Entre estos se destaca el Open Science Framework (OSF; <https://osf.io>), que integra una gran variedad de recursos que nos posibilita plasmar de manera transparente, clara y completa el flujo de trabajo de la investigación (la honestidad-veracidad de lo reportado corre por cuenta del investigador). Otro recurso interesante en este sentido es la reciente lista de verificación de prácticas transparentes propuesta por Aczel et al. (2019; <http://www.shinyapps.org/apps/TransparencyChecklist/>)

Finalmente me parece importante que el lector y lectora de este trabajo tengan presente que existen circunstancias o factores que hacen más probable la inclinación de nuestra parte a, en el mejor de los casos, conductas cuestionables en investigación. Algunos de estos factores son más macro, en el sentido que involucran tendencias más generales

de hacer ciencia y la cultura científica predominante al momento (e.g., cultura que obliga a publicar en cantidad y rápidamente para la obtención de subsidios, becas o el mantenimiento de puestos de trabajo); mientras que otros factores son más micro, involucrando principalmente sesgos del investigador/a (Bosma & Granger, 2022; Tijdink et al., 2021). Entre estos sesgos se encuentran la tendencia a visualizar estructura en donde solo hay ruido, el reporte selectivo en favor de los resultados que cumplen las expectativas del investigador, la tendencia a enfatizar los resultados que alcanzan la significación estadística, entre muchos otros. No es mi objetivo profundizar aquí sobre estos sesgos puntuales, y el lector/a interesado/a puede consultar excelentes trabajos en donde estos tópicos ya han sido bien desarrollados (véase por ejemplo Antonakis, 2017). Por último, decir que para el alcance de una conducta en investigación responsable (i.e., integridad científica) es necesaria también la implementación de acciones en los distintos niveles en donde puede ubicarse la conducta científica (Tijdink et al., 2021): no solo involucrando el proceso científico en sí mismo, sino también abarcando el sistema científico más general (e.g., políticas; prácticas; sistema de recompensa; evaluación; y formación).

Con todo, espero que el presente trabajo haya cumplido con el objetivo de hacer más claro para los/as investigadores/as de la región la manera en que nuestra conducta y la de nuestros pares puede ser evaluada. Idealmente, espero que este aporte promueva un giro hacia una forma de hacer ciencia más íntegra, y más responsable.

Referencias

- 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., ... Wagenmakers, E.-J. (2019). A consensus-based transparency checklist. *Nature Human Behaviour*, 4(1), 4–6.
<https://doi.org/10.1038/s41562-019-0772-6>
- Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly*, 28(1), 5–21.
<https://doi.org/10.1016/j.leaqua.2017.01.006>
- Antonakis, J. (2023). In support of slow science: Robust, open, and multidisciplinary. *The Leadership Quarterly*, 34(1), 101676.
<https://doi.org/10.1016/j.leaqua.2023.101676>
- Barber, T. X. (1976). *Pitfalls in human research: Ten pivotal points*. Pergamon Press.
- Bosma, C. M., & Granger, A. M. (2022). Sharing is caring: Ethical implications of transparent research in psychology. *American Psychologist*, 77(4), 565–575.
<https://doi.org/10.1037/amp0001002>
- Bouter, L. M., Tjeldink, J., Axelsen, N., Martinson, B. C., & Ter Riet, G. (2016). Ranking major and minor research misbehaviors: Results from a survey among participants of four World Conferences on Research Integrity. *Research Integrity and Peer Review*, 1(1), 17. <https://doi.org/10.1186/s41073-016-0024-5>

- Chin, J. M., Pickett, J. T., Vazire, S., & Holcombe, A. O. (2021). Questionable Research Practices and Open Science in Quantitative Criminology. *Journal of Quantitative Criminology*. <https://doi.org/10.1007/s10940-021-09525-6>
- Gelman, A., & Loken, E. (2016). The Statistical Crisis in Science. In M. Pitici (Ed.), *The Best Writing on Mathematics 2015* (pp. 305–318). Princeton University Press. <https://doi.org/10.1515/9781400873371-028>
- Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLoS Medicine*, 2(8), e124. <https://doi.org/10.1371/journal.pmed.0020124>
- Ioannidis, J. P. A. (2012). Why Science Is Not Necessarily Self-Correcting. *Perspectives on Psychological Science*, 7(6), 645–654. <https://doi.org/10.1177/1745691612464056>
- John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, 23(5), Article 5.
- Manapat, P. D., Anderson, S. F., & Edwards, M. C. (2022). A revised and expanded taxonomy for understanding heterogeneity in research and reporting practices. *Psychological Methods*. <https://doi.org/10.1037/met0000488>
- McCullough, L. B., Coverdale, J. H., & Chervenak, F. A. (2004). Argument-based medical ethics: A formal tool for critically appraising the normative medical ethics literature. *American Journal of Obstetrics and Gynecology*, 191(4), 1097–1102. <https://doi.org/10.1016/j.ajog.2004.06.060>

Moher, D. (2007). Reporting research results: A moral obligation for all researchers.

Canadian Journal of Anesthesia/Journal Canadien d'anesthésie, 54(5), 331–335.

<https://doi.org/10.1007/BF03022653>

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du

Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A.

(2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(1),

0021. <https://doi.org/10.1038/s41562-016-0021>

Needleman, I., Moher, D., Altman, D. G., Schulz, K. F., Moles, D. R., & Worthington, H.

(2008). Improving the Clarity and Transparency of Reporting Health Research: A

Shared Obligation and Responsibility. *Journal of Dental Research*, 87(10), 894–

895. <https://doi.org/10.1177/154405910808701013>

Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's Renaissance.

Annual Review of Psychology, 69(1), 511–534. <https://doi.org/10.1146/annurev->

[psych-122216-011836](https://doi.org/10.1146/annurev-psych-122216-011836)

Nicholls, S. G., Langan, S. M., Benchimol, E. I., & Moher, D. (2016). Reporting

transparency: Making the ethical mandate explicit. *BMC Medicine*, 14(1), 44,

[s12916-016-0587-5](https://doi.org/10.1186/s12916-016-0587-5). <https://doi.org/10.1186/s12916-016-0587-5>

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler,

F., Hilgard, J., Kline Struhl, M., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel,

A. M., Scherer, L. D., Schönbrodt, F. D., & Vazire, S. (2022). Replicability,

Robustness, and Reproducibility in Psychological Science. *Annual Review of*

Psychology, 73(1), 719–748. <https://doi.org/10.1146/annurev-psych-020821->

[114157](https://doi.org/10.1146/annurev-psych-020821-114157)

- Resnik, D. B. (2005). *The Ethics of Science: An Introduction*. Taylor and Francis.
- Resnik, D. B. (2018). *The Ethics of Research with Human Subjects: Protecting People, Advancing Science, Promoting Trust* (Vol. 74). Springer International Publishing.
<https://doi.org/10.1007/978-3-319-68756-8>
- Sacco, D. F., Brown, M., & Bruton, S. V. (2019). Grounds for Ambiguity: Justifiable Bases for Engaging in Questionable Research Practices. *Science and Engineering Ethics*, 25(5), 1321–1337. <https://doi.org/10.1007/s11948-018-0065-x>
- Shamoo, A. E., & Resnik, D. B. (2009). *Responsible conduct of research* (2nd ed). Oxford University Press.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, 22(11), 1359–1366.
<https://doi.org/10.1177/0956797611417632>
- Stefan, A., & Schönbrodt, F. D. (2022). *Big Little Lies: A Compendium and Simulation of p-Hacking Strategies* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/xy2dk>
- Steneck, N. H. (2006). Fostering integrity in research: Definitions, current knowledge, and future directions. *Science and Engineering Ethics*, 12(1), 53–74.
<https://doi.org/10.1007/s11948-006-0006-y>
- Tijdink, J. K., Horbach, S. P. J. M., Nuijten, M. B., & O'Neill, G. (2021). Towards a Research Agenda for Promoting Responsible Research Practices. *Journal of Empirical Research on Human Research Ethics*, 16(4), 450–460.
<https://doi.org/10.1177/15562646211018916>

Tuval-Mashiach, R. (2017). Raising the curtain: The importance of transparency in qualitative research. *Qualitative Psychology*, 4(2), 126–138.

<https://doi.org/10.1037/qup0000062>

Wigboldus, D. H. J., & Dotsch, R. (2016). Encourage Playing with Data and Discourage Questionable Reporting Practices. *Psychometrika*, 81(1), Article 1.

<https://doi.org/10.1007/s11336-015-9445-1>

World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. (2013). *JAMA*, 310(20), 2191.

<https://doi.org/10.1001/jama.2013.281053>