

Compartir datos: hacia una investigación más sostenible

Juan Botella Ausina* y Carmen Ortego Maté**

*Universidad Autónoma de Madrid y **Universidad de Cantabria

En Psicología hay poca tradición de compartir los datos directos de nuestras investigaciones. Esta falta de tradición, más otras circunstancias, reflejan algunas debilidades de la psicología como ciencia. Se exponen las razones por las que hay poca disposición a compartir los datos y se presentan contraargumentos para defender los beneficios que tendría un cambio en esta actitud. Aparte de fortalecer el carácter científico de la psicología, este cambio redundaría en un desarrollo más sostenible de la investigación, permitiendo prácticas como el análisis secundario y el reciclado de datos. Se proponen vías de actuación para facilitar este cambio.

Data sharing: Towards a more sustainable research. In psychology, there is little tradition of sharing raw data from one's research. This lack of tradition, along with other circumstances, reflects some weaknesses of psychology as a science. The reasons for which this scarce willingness to share data and the counter-arguments in favour of the benefits that a change in attitude would involve are discussed in this article. Such a change would not only strengthen the scientific nature of psychology, but also result in a more sustainable development of the research, allowing practices such as recycling and secondary analyses of raw data. Some ways of action to facilitate this change are also suggested.

Existe hoy una fuerte sensibilización para reconducir las actividades humanas hacia parámetros de sostenibilidad que deberíamos extender a la investigación científica en general, y a la psicología en particular. Desde este marco vamos a defender la necesidad de desarrollar actitudes y estrategias que potencien el intercambio de datos. Estas acciones, aunque aparentemente sencillas, pueden tener un fuerte efecto en la reorientación hacia una ciencia más sostenible y, en particular, hacia una psicología que sea más sostenible que la actual, en la que casi cada investigación exige la generación de datos nuevos.

Nuestros principales argumentos giran en torno al cumplimiento de un principio ético, a la transparencia y la calidad en la investigación, a la posibilidad de hacer una acumulación más eficiente del conocimiento, a la eventual reutilización de los datos para responder a preguntas antes insospechadas (análisis secundario) y a la protección contra las amenazas a la fiabilidad.

Pero eso no es todo. En otras disciplinas también hay resistencias a compartir los datos, pero los psicólogos ofrecemos una especial resistencia. No sólo no tenemos la cultura ni la tradición de poner al libre acceso nuestros datos brutos, sino que solemos ser reacios a proporcionárselos a quien nos los solicita directamente. Aprovecharemos esta peculiaridad para señalar algunas idiosincrasias sobre su origen y naturaleza, comprobando que al analizar las causas de esta actitud la imagen que nos devuelve el espejo desde esta inusual perspectiva no es muy reconfortante.

Nuestro objetivo en este artículo es doble. Por un lado, fomentar y defender la actitud de compartir datos. Por otro lado, proponer estrategias que favorezcan estos comportamientos.

La débil cultura de compartir datos de la Psicología

Muchas revistas de Psicología han adoptado el código ético de la APA (2001a) para la publicación de información científica. Uno de sus estándares es que los autores deben mantener a disposición de la comunidad científica los datos en los que se basan las investigaciones que publican, permitiendo el acceso a los mismos a quien lo solicite (www.apa.org/science/data).

Hace unos años la prestigiosa revista científica generalista *Nature* (2006) dedicó un editorial a los usos, costumbres y actitudes que prevalecen en la psicología en torno a esta cuestión. Tras señalar nuestra poca tradición con esta práctica, narra el caso de Wicherts y sus colegas (Wicherts, Borsboom, Kats y Molenaar, 2006), quienes querían valorar la robustez de los análisis estadísticos de las investigaciones publicadas en varias revistas de la APA. A pesar de la directriz mencionada anteriormente, su carta a *American Psychologist* narra la larga trayectoria que siguieron, incluyendo centenares de mensajes electrónicos, para acabar desistiendo. Aunque el caso se trató en la reunión de editores de la APA de 2007, acordando que “algo había que hacer”, lo cierto es que por el momento las cosas han cambiado poco (Wicherts, comunicación personal). De hecho, experiencias similares habían sido denunciadas ya en los años 60 y 70 (Craig y Reese, 1973; Wolins, 1962), pero al parecer la situación no ha cambiado mucho a pesar de las enormes facilidades actuales de comunicación.

En una reciente investigación hemos podido comprobar estas dificultades en primera persona. Nuestro objetivo era realizar un

Fecha recepción: 16/2/2009 • Fecha aceptación: 29/6/2009

Correspondencia: Juan Botella Ausina

Facultad de Psicología

Universidad Autónoma de Madrid

28049 Madrid (Spain)

e-mail: juan.botella@uam.es

análisis integrado de algunos de los datos contenidos en 109 estudios primarios incluidos en 4 meta-análisis (Ortego y Botella, 2010). Tras solicitar a los autores principales de estos 109 estudios las puntuaciones directas de algunas de sus variables y a los autores principales de los 4 meta-análisis información adicional, tan sólo conseguimos las puntuaciones de 13 estudios primarios y la información adicional de 2 meta-análisis.

Pero esta dificultad a la hora de obtener los datos directos se extiende a peticiones de informaciones más generales. Nuestra experiencia y la de la mayoría de los investigadores que han realizado o replicado un meta-análisis lo ponen de manifiesto. En estos casos al meta-analista solo le hacen falta resúmenes numéricos (medias, desviaciones típicas, correlaciones y tamaños muestrales) y algunas características específicas de cada estudio; ante la ausencia de esta información estadística en los informes publicados el meta-analista se ve obligado a solicitarla al autor. A pesar de ello, los autores con frecuencia se resisten a proporcionar lo solicitado.

Nuestra conclusión es que los psicólogos suelen ser notablemente reacios a cooperar entre sí desinteresadamente, al menos en cuanto a compartir sus datos con desconocidos. Sería injusto decir que siempre es así. Junto a los casos mencionados hay otros que lo hacen con prontitud y diligencia, mostrando una actitud de franca colaboración. Pero son los menos.

El defectuoso carácter acumulativo de la Psicología

Una de las características tradicionalmente atribuidas al conocimiento científico es su carácter acumulativo (Bem y Looren de Jong, 1997). Esto solo es posible si el conocimiento que genera está relativamente consensuado por científicos que comparten una misma forma de entender su objeto de estudio. Las teorías y explicaciones que en cada momento se asumen como las más sólidas son aquellas que se fundamentan en las evidencias empíricas disponibles. Las nuevas hipótesis, principales o secundarias, orientan la búsqueda de evidencias adicionales que, a su vez, se integrarán en las bases empíricas para suscitar nuevas hipótesis en los mismos u otros investigadores. Este ciclo se repite indefinidamente, constituyendo la columna vertebral del proceso de acumulación. La ciencia puede ser vista como una tarea colectiva de los investigadores.

Si bien es cierto que hay “revoluciones” (Kuhn, 1970), los paradigmas científicos, en el sentido kuhniano, establecen unos límites y marcan unas pautas durante su vigencia. Un cambio de paradigma puede obligar a una revisión completa de la visión dominante pero al menos durante su vigencia el conocimiento puede considerarse acumulativo. Esto permite agotar las posibilidades de la propuesta conceptual más verosímil de cada momento.

Si este panorama puede aplicarse a otras disciplinas, el carácter supuestamente acumulativo en períodos de “normalidad” no es tan evidente en psicología. En nuestra disciplina el proceso de acumulación ha sido tradicionalmente muy defectuoso.

¿Es una casualidad que el Meta-análisis naciera en la Psicología?

El defectuoso proceso de acumulación del conocimiento en Psicología tuvo una curiosa consecuencia en los años 70. El meta-análisis, una metodología ideada para acumular evidencias, nace precisamente en la intersección entre la psicología y la educación. Lejos de ser una casualidad, este hecho refleja las características

que hacían de la Psicología el candidato ideal para que germinase y prosperase. Este hecho es, sin duda, ilustrativo para nuestra argumentación (Botella y Gambará, 2006). El meta-análisis nació ante la demanda de procedimientos para hacer revisiones más rigurosas y sistemáticas, capaces de proporcionar conclusiones más precisas (Lipsey y Wilson, 1993), frente a la opacidad y poca validez de su alternativa tradicional, la revisión narrativa (Botella y Gambará, 2002).

Una de las razones por las que cada vez era más difícil sintetizar conclusiones precisas era el tamaño de la investigación (Botella y Gambará, 2006). Pero reducir la explicación a un problema meramente cuantitativo, de tamaño de la literatura científica, sería demasiado ingenuo. El primer meta-análisis publicado, origen de una fuerte polémica, analizaba y comparaba el éxito terapéutico de intervenciones en el ámbito de la psicología clínica realizadas desde las perspectivas de diversas escuelas (Smith y Glass, 1977). Más que aclarar la situación, aquel meta-análisis reavivó el árido enfrentamiento que había estado presente en las décadas anteriores (Eysenck, 1978).

En realidad había una desagradable sensación de pesimismo y frustración a la vista de que prácticamente todos los campos estaban llenos de estudios con resultados inconsistentes. Pero si no es casualidad que el meta-análisis naciera en la psicología, tampoco creemos que lo sea que el primer meta-análisis estuviera dedicado a la eficacia de las intervenciones en psicología clínica. Las circunstancias específicas de la psicología (especialmente de la psicología clínica) favorecieron el nacimiento de una metodología que no hace otra cosa que responder a una sentida necesidad (Botella y Gambará, 2006). Esta necesidad surge de la defectuosa acumulación que se hacía en su seno, asociada a una fuerte carga de dogmatismo. Una consecuencia muy negativa para la psicología ha sido que con frecuencia no ha sabido ofrecer respuestas claras a las preguntas que se le han hecho desde fuera. La sociedad demanda orientaciones sobre una variedad de problemas, pero la respuesta de los psicólogos con frecuencia es contradictoria, fragmentaria y confusa. En el origen del meta-análisis estaba la necesidad de ofrecer respuestas más claras a las demandas externas. Empezaba a haber una clara conciencia de que la acumulación no se hacía sola. Había que implicarse activamente en esta tarea y hacerlo con el mismo rigor que se exigía a la investigación primaria (Chalmers, Hedges y Cooper, 2002).

Tal y como señalaba uno de los revisores que leyeron este manuscrito (a quien agradecemos sus sugerencias) a veces en psicología las variables son definidas “ad hoc” para un estudio concreto, lo que arroja dudas sobre la utilidad de compartirlas. Siendo esto cierto, también es verdad que en la mayoría alegan constructos previamente definidos por otros autores. Nuestro argumento aquí es que el hecho de que el meta-análisis naciera en la psicología es una prueba palpable de la necesidad de integrar resultados que los propios autores relacionan con un mismo constructo.

Lamentablemente, los males de fondo que dieron lugar al nacimiento del meta-análisis y que hacen que los psicólogos no compartan sus datos siguen en parte presentes. A ello dedicamos el siguiente apartado.

¿Por qué los psicólogos son reacios a compartir sus datos?

El mencionado editorial de *Nature* expone algunos de los argumentos tradicionales para justificar las negativas a compartir datos. En primer lugar, está la *cuestión de la confidencialidad*. No

vamos a argumentar que la conveniencia de compartir datos deba prevalecer sobre el derecho a la confidencialidad, pero tampoco creemos que ésta pueda ser una excusa para justificar el estado de cosas del que estamos hablando. En la mayoría de los casos los datos se refieren a informaciones que los propios participantes compartirían sin reservas. Excluir los “datos sensibles” probablemente dejaría fuera apenas un porcentaje mínimo de los estudios.

Además, hay procedimientos que permiten poner al alcance de otros los datos separando o eliminando los que permitirían una identificación personalizada de los participantes. Por ejemplo, en una encuesta un análisis secundario puede simplemente elaborar Tablas de Contingencia entre variables, en combinaciones que no han sido ofrecidas en los análisis primarios. De esta forma se pueden obtener nuevos factores de riesgo o estimaciones de probabilidades condicionales. Para esto no es necesario conocer la identidad de los individuos. Con códigos que atribuyan números de orden a los casos se pueden hacer identificaciones relativas, intra-muestra, suficientes para cualquier exploración posterior. Estos procedimientos preservan la confidencialidad al eliminar la posibilidad de identificar a los participantes.

Una segunda razón es el *esfuerzo invertido*. Los autores se resisten a permitir el acceso de otros a sus datos porque les ha costado mucho esfuerzo obtenerlos y no quieren que otros se aprovechen de su trabajo.

Otra razón tiene que ver con el *cuestionamiento de su competencia profesional*. Un re-análisis podría llevar a conclusiones diferentes, poniendo en duda su competencia para explotar sus propios datos o, incluso, para cuestionar su integridad. Sin embargo, los debates en este sentido probablemente serían fructíferos.

Otro motivo esgrimido en el editorial tiene que ver con la *prevención ante usos malintencionados de los datos*, pues podrían ser malinterpretados por alguien con una determinada agenda política. Es cierto, pero quien quiere hacer esto también lo puede hacer con los resúmenes estadísticos publicados.

Por último, una razón pocas veces reconocida es que algunos investigadores son incapaces de registrar y presentar sus datos en un formato no ambiguo, fácil de leer y archivable.

Sin embargo, creemos que al catálogo de razones de *Nature* para no compartir datos habría que añadir alguna más. En la psicología hay un margen mayor de interpretabilidad de los datos que en otras ciencias más “duras”. Esta es una de las causas de la deficitaria acumulación de la Psicología. Mientras en las ciencias más duras los datos tienen poca ambigüedad, de forma que cualquier investigador del campo entiende esencialmente lo mismo, en psicología el margen de interpretación es mucho mayor. Una consecuencia inevitable es que los psicólogos confían menos en los datos que no han generado ellos mismos. A veces tampoco confían mucho en los suyos propios, lo que se puede convertir en un disuasorio para compartirlos. Esta característica lastra notablemente el desarrollo de la psicología científica y, de hecho, es una de las causas de que el meta-análisis naciera en la Psicología. Hoy la vemos también como una de las razones por las que compartimos poco nuestros datos. La necesidad de redoblar nuestros esfuerzos por reducir la ambigüedad de los datos aparece como inevitable consecuencia.

A esto hay que añadir que la versatilidad de los procedimientos estadísticos proporciona un margen de actuación considerable. No sería justo decir que con el análisis estadístico se pueda llegar a la conclusión que se quiera. Sin embargo, cualquiera que tenga una cierta pericia en el campo sabe que unos datos que muestren

una cierta inconsistencia, o estén al límite entre dos patrones interpretables, siempre se pueden analizar empleando técnicas y procedimientos que conducen más fácilmente a lo que se busca.

Todo esto nos lleva a pensar que, además de la cuestión de la confidencialidad y de la resistencia a que otros se aprovechen de nuestro trabajo, el resto de los argumentos de *Nature* esconden algo más fundamental. Al levantar las alfombras vemos una cierta debilidad de la psicología, que en algunos campos ha estado levantando edificios con cimientos de barro. Nos hemos dedicado durante décadas a desarrollar procedimientos que mejoren la fiabilidad de los datos, es decir, a tener ciertas garantías de que la medición es razonablemente estable y consistente. Sin embargo, lo que realmente le hace falta hoy en día a la Psicología es un impulso definitivo a la validez (Messick, 1988). Es decir, conseguir que nuestros registros se refieran a lo que dicen referirse y a nada más.

Respecto al tratamiento de los datos se podría decir algo parecido. Disponemos de técnicas de análisis desproporcionadamente potentes para la calidad de los datos a los que se aplican. Procedimientos tan complejos como el ajuste de modelos de ecuaciones estructurales, impensables antes de disponer de ordenadores potentes al alcance de la mano, hoy son moneda corriente. Sin embargo, con frecuencia vemos cómo estas técnicas se aplican a datos que proceden de sencillos cuestionarios aplicados en condiciones dudosas. Los participantes responden a preguntas en formato cerrado a las que no estamos seguros de que hayan hecho una interpretación correcta o completa y con niveles de validez inciertos. La aparatosidad estadística a veces camufla la pobreza de los datos.

El lector podría pensar que nuestra conclusión es que nos hemos empeñado en una tarea, la de construir una Psicología científica, que en realidad es imposible. Muy al contrario, creemos que haciendo explícitos estos argumentos podemos reafirmarnos en nuestro empeño, reforzando lo que estamos haciendo bien y reconduciendo lo que no. En este segundo grupo está la actitud de los psicólogos ante la propuesta de compartir sus datos. En el siguiente apartado desarrollamos nuestros argumentos.

Pero antes de abandonar este apartado no queremos dejar de señalar también que cuando los psicólogos no conseguimos definir bien los conceptos y constructos, no debemos ocultarlo. Señalando nosotros mismos los puntos débiles estamos poniendo límites a lo que es razonable concluir a partir de los datos. Ocultar esas debilidades como si fueran vergüenzas familiares, en nada ayuda al desarrollo de la psicología.

¿Por qué los psicólogos deberíamos compartir nuestros datos?

Comenzaremos exponiendo tres argumentos para promover la actitud de compartir datos que serían defendidos desde cualquier disciplina: *es un principio ético, ayuda a prevenir el fraude* y *protege contra algunas amenazas a la fiabilidad*.

Compartir datos y material procedente de una investigación es un *principio ético* para los investigadores de cualquier disciplina. Como ya hemos mencionado, instituciones como la APA exhortan a los investigadores a guardar y compartir los datos y material con otros investigadores, al menos durante 5 años después de su publicación. Además, es un indicador de transparencia.

Con respecto al *fraude* en la ciencia, su origen está en que el científico, como todo ser humano, está sujeto a ciertas debilidades. Suele ser especialmente vulnerable a la ambición y la vanidad. La

posibilidad de un gran reconocimiento de la comunidad científica o de la sociedad en general, amén de ventajas para ascender profesionalmente, sin olvidar las consecuencias económicas, son algunas de las razones que pueden llevar a cometer un fraude consciente y planificado. En otros casos el origen no está en los propios científicos, sino en instituciones o empresas con intereses económicos. Para realizarlo necesitan la colaboración de científicos que presten conscientemente su nombre a un fraude dirigido específicamente a generar un beneficio económico o a apoyar una posición ideológica.

Compartir los datos supone un obstáculo adicional a las tentaciones de incurrir en prácticas fraudulentas. Hay muchas formas de realizar un fraude científico, de las que la más obvia es inventarse los datos. Si hubiera que publicarlos ya no se podría enviar a la revista un simple resumen que describiera las conclusiones de unos datos inexistentes. No se podría elaborar un informe ficticio sin más.

Otra forma más sofisticada es la invención de los datos directos, antes de la elaboración del informe que se envía a la revista. Esto implica un trabajo adicional de generación de unos datos creíbles, tanto en las regularidades buscadas como en las anomalías puntuales que cualquier investigador conoce (datos perdidos, valores atípicos, etc). Otra forma de fraude puede tomar unos datos reales e incidir en una fase posterior, mediante un estudio maquillado. Otra consiste en aplicar análisis estadísticos tendenciosos, elegidos conscientemente para favorecer la legitimación de las conclusiones que los autores persiguen. Si bien todo esto puede ocurrir en cualquier ciencia, en las ciencias más “duras” la capacidad de replicación es lo suficientemente alta como para que el fraude no tarde mucho en ser detectado. En Psicología no tenemos una capacidad de replicación tan grande, pero la publicación de los datos directos, aparte del informe en forma de artículo, es un obstáculo adicional.

La cuestión de la *fiabilidad* es de naturaleza diferente. Desde actitudes honestas se pueden producir errores en la grabación, tratamiento y análisis de los datos. Al ponerlos a disposición de la comunidad científica cualquiera puede, desde actitudes bien intencionadas, replicar los análisis y detectar los posibles errores (Craig y Reese, 1973). Aunque todos podemos cometer errores, la prudencia y la minuciosidad son valores que deben acompañar al trabajo científico. La posibilidad de un re-análisis sería una garantía adicional de confianza porque habría una mayor probabilidad de detectar los errores y porque fomentaría el reforzamiento de las medidas de control que adoptan los propios autores (Freedland y Carney, 1992).

A estos tres argumentos, perfectamente aplicables a cualquier disciplina, vamos a añadir algunos más específicos de la psicología.

Nuestro cuarto argumento se refiere a la posibilidad de realizar nuevos análisis o, como se suelen llamar, *análisis secundarios* (Glass, 1976; Law, 2005). No se trata solo de un análisis con procedimientos alternativos, para valorar las mismas hipótesis. Desde luego, cualquier científico puede idear formas de análisis de los datos que exploten mejor la información que contienen. Ello implica argumentar esa supuesta superioridad y realizar un análisis a posteriori de sensibilidad, ilustrando las consecuencias que tienen sobre las conclusiones la elección de las técnicas estadísticas, las transformaciones previas, etc.

Pero un análisis secundario es algo más. Es frecuente que en psicología aparezcan nuevas teorías y modelos desde las que los

datos se miran de otra forma. Ya no se trata de aplicar procedimientos alternativos para valorar más eficientemente las mismas hipótesis. Más bien, se trata de valorar hipótesis completamente nuevas con los datos de investigaciones anteriores que a posteriori resultan pertinentes, aunque no fueran recogidos originalmente con esa intención. De hecho, es fácil que esos nuevos modelos y teorías hayan sido desarrollados después de la generación de los propios datos (tampoco se debe confundir el análisis secundario con el llamado “undiscovered public knowledge”; Spasser, 1997; Swanson, 1987). Por ello a este uso basado en el principio “reducir, reutilizar, reciclar” (Law, 2005) podría denominarse *reciclado de datos* y conduce directamente al concepto de *ciencia sostenible*.

Cada vez hay una mayor conciencia social de que el nuestro es un mundo con recursos limitados. La filosofía del reciclado se aplica a numerosos contextos cada vez con mayor efectividad y alcanzando mayores niveles de conciencia social. En el caso de la investigación en psicología la pregunta es si podemos permitirnos el lujo de recoger nuevos datos siempre que queramos poner a prueba nuevas hipótesis, cuando en los archivos de otros investigadores hay almacenados datos que pueden darnos la respuesta. Es verdad que con frecuencia esos datos no son exactamente los que necesitamos, pero otras veces el análisis de esos datos parciales permite acotar mejor la búsqueda y ahorrar recursos.

No nos resistimos a mencionar un ejemplo que ilustra bien lo que estamos exponiendo. Su origen está en los acontecimientos que se produjeron en el observatorio astronómico de Greenwich en 1796, a propósito de las discrepancias entre los registros de Maskelyne y Kinnebrook, que llevaron al primero a despedir al segundo. La existencia de diferencias sistemáticas entre distintos observadores condujo en el siglo XIX al estudio de las características personales y la llamada ecuación personal (Boring, 1980), eximiendo a Kinnebrook de toda responsabilidad. Recientemente Mollon y Perkins (1996) han desempolvado aquellos datos y los han re-analizado desde una perspectiva muy distinta: la de los modernos modelos cognitivos de generación de números pseudo-aleatorios (e.g., Baddeley, Emslie, Kolodny y Duncan, 1998; Towse y Cheshire, 2007). Su sorprendente conclusión, que vuelve a condenar a Kinnebrook por el poco rigor de sus anotaciones, solo ha sido posible porque los datos de los registros astronómicos son tradicionalmente guardados de forma sistemática y porque nadie duda de lo que los datos significan.

A los argumentos expuestos para compartir datos en la psicología queremos añadir tres: el desarrollo de la psicología como una ciencia más sostenible, el impulso de la práctica basada en la evidencia y el uso pedagógico de las bases de datos.

En torno al concepto de *sostenibilidad* ha surgido la cultura del reciclado, que fomenta la reutilización antes del uso de nuevos recursos. En esta filosofía se enmarca la cultura del compartir datos y para la que la identificación del genoma humano se ha convertido en modelo de referencia. Compartir datos hace de la psicología una disciplina más sostenible. La reutilización de datos es más barata. Ya está relativamente instalada la filosofía de la integración a partir de los resúmenes estadísticos que aparecen en los informes, tal y como hace el meta-análisis, pero sin duda hay que ir más allá.

No debemos esperar un cambio espontáneo en los científicos. Por diversas razones muchos de ellos son y serán reacios a compartir sus datos, aunque creemos que respetando los principios éticos que deben guiar cualquier investigación el compartir los datos no debería suponer un conflicto entre investigadores. Solo a partir de las necesidades expresadas desde fuera los científicos cambia-

rán realmente su valoración de las diferentes contribuciones que pueden hacer al conocimiento científico. Además de valorar más las revisiones, hay que fomentar las investigaciones que se derivan de la evidencia ya acumulada y racionalmente construida hasta el momento.

Nuestro siguiente argumento se refiere a la *práctica basada en la evidencia*. Una de las fuerzas que dinamizan nuestra ciencia es el incremento de la demanda externa de evidencia científica, por parte de los gestores, políticos y profesionales, para su empleo como base de sus propias decisiones. Este es otro de los vectores que están incidiendo en la reciente reorientación de la investigación científica en psicología. En los últimos años hemos asistido en España a debates, a veces muy acalorados, acerca de temas como la adopción por parte de parejas homosexuales, la educación segregada por sexos o la posibilidad de rehabilitación de determinados tipos de delincuentes; obviamente, siempre son temas con un fuerte componente ideológico y, por tanto, muy expuestos al dogmatismo. Queremos creer que nuestra sociedad anhela despojarse de criterios dogmáticos y pretende incorporar cada vez más criterios científicos. Mucha gente se queda perpleja al comprobar que posiciones completamente enfrentadas respecto a estos temas pueden estar avaladas por expertos que aportan evidencias que favorecen sus puntos de vista. La credibilidad de la psicología no se ve precisamente favorecida por estos incidentes. A muchos nos gustaría disponer de evidencias directamente interpretables sin discursos ideológicos y posiciones dogmáticas. No quiere decir que creamos en una psicología completamente despojada de ideología, pero sí que se puede hacer una separación más explícita y clara entre los hechos y las valoraciones desde postulados ideológicos.

La ausencia de buenas síntesis ha tenido unos efectos muy adversos para la psicología, al reducir sus aportaciones a la sociedad y limitando la presencia de la psicología científica en la vida cotidiana (Botella y Gambará, 2006). El problema no ha sido tanto el no contar con datos relevantes, sino más bien que los resultados son contradictorios y confusos, e incluso a veces tendenciosos.

A lomos de esta tendencia ha ido creciendo la corriente conocida como “práctica basada en la evidencia”, nacida como “medicina basada en la evidencia” (e.g., Sackett, Rosenberg, Gray, Haynes y Richardson, 1996; Sackett, Straus, Richardson, Rosenberg y Haynes, 2000). Su principio es: a falta de un criterio mejor, más vale tomar como base la evidencia conocida y fundamentar las intervenciones en los estudios que valoran la eficacia de las intervenciones, unido a la experiencia clínica del médico. En este principio se han inspirado, entre otras, la “psicología basada en la evidencia” (APA, 2006b).

Nuestro último argumento tiene que ver con el *uso pedagógico* de las bases de datos. Los estudiantes de psicología de todos los niveles formativos (licenciatura, postgrado y cursos de especialización) podrían disponer de datos reales en su proceso de aprendizaje. No hace falta resaltar que para los docentes de metodología serían de gran utilidad, dado que permitirían replicar el proceso de análisis de trabajos ya publicados, mostrando cómo se pueden detectar las dificultades, cómo resolverlas y cómo manejarlas en el proceso de publicación. Igualmente, podrían explorar el grado de sensibilidad de las conclusiones a las decisiones adoptadas en el proceso, tales como las eventuales transformaciones de datos, su depuración previa, la elección de la técnica estadística de análisis o, incluso, la del *software* empleado.

Pero por muy positivo que nos parezca que los psicólogos compartan sus datos, no podemos ocultar que también tiene algunos

inconvenientes. Entre ellos está, por ejemplo, el hecho de que un análisis repetido de unos mismos datos puede acabar produciendo una excesiva representación de los mismos en la literatura. Cualquiera particularidad de los mismos, que se diluiría en un colectivo de muestras diferentes como efecto de las fluctuaciones aleatorias, persistiría una y otra vez en cada nuevo análisis. En este sentido es importante destacar bien el origen de los datos de cada nuevo análisis o informe, para no acabar creyendo que las nuevas conclusiones se basan en nuevas evidencias.

Algunas vías de acción

Completaremos nuestra exposición con algunas propuestas prácticas. Compartir datos no es la panacea que resuelve los problemas de la investigación en psicología, pero creemos que avanzar en esa dirección tendría efectos positivos. Aquellos a los que hayamos podido convencer de sus beneficios estarán preguntándose cómo pueden hacer para ponerlo en práctica (Johnson y Sabourin, 2001). En primer lugar, no estaría de más que los psicólogos españoles se informasen de cómo está la cuestión en otras disciplinas o en la psicología de otros países. Comprobarían que hay muchas formas de hacerlo y que no es tan obvio como parece.

En segundo lugar, habría que ayudar a cambiar las actitudes en nuestros entornos de trabajo, dado que esta filosofía está poco desarrollada en la psicología porque los psicólogos tienen actitudes en contra basadas en un conjunto de creencias y desconfianzas poco fundamentadas. Nos gustaría haber contribuido con estas páginas a cambiar estas actitudes. Por un lado, debilitando las razones negativas y, por otro, reforzando las positivas que hubiera y añadiendo otras nuevas que esperamos hayan sido convincentes.

En tercer lugar, proponemos dos vías concretas de actuación: la creación de herramientas prácticas y las alianzas estratégicas.

Con la *creación de herramientas prácticas* nos referimos a la generación de instrumentos que permitan extender estas prácticas sin incrementar significativamente la carga de trabajo del investigador. Pero a la vez tienen que ser útiles. David H. Johnson, un inagotable defensor de la causa, ha hecho contribuciones muy destacadas a la misma. Además de un fino análisis de la psicología del investigador de la psicología (Johnson, 2001), ha mencionado diversas formas en las que se podría llevar a cabo esta tarea. La más simple es la que los particulares realizan cuando se envían datos, como consecuencia de una relación bilateral. Una algo más desarrollada sería la implantación, por parte de las instituciones, de una política que fomente esta práctica, aunque el almacenamiento sería aún realizado por los autores de cada trabajo; naturalmente, exigiría una combinación de archivos que incluiría tanto el de los datos propiamente dichos como algunos otros explicativos de las variables y su métrica. Debería ir acompañado de un trabajo de homogeneización de los formatos. También las revistas podrían acoger en sus páginas web, en concepto de material suplementario, los archivos asociados a cada artículo (Wicherts et al, 2006), como ya hacen o han hecho algunas revistas (véanse, por ejemplo, los casos de *European Journal of Psychological Assessment* o las revistas que publica la *Psychonomic Society*) y en otras instancias (en la Universidad Autónoma de Madrid estamos ya trabajando en la creación de un lugar donde compartir bases de datos psicométricos).

La principal limitación de estas dos formas es que los autores tendrían que buscar los archivos de datos cuya existencia han descubierto por otros medios, especialmente la lectura del artículo co-

rrespondiente. Un procedimiento más rico consistiría en volcar los archivos de datos en bases cuyos elementos fueran a su vez bases de datos (archivos relacionales). Desde ella se podrían localizar las bases disponibles sobre el tema deseado. Esta vía implicaría que alguien hiciera el trabajo.

Naturalmente, la cuestión de la confidencialidad ha de ser resuelta satisfactoriamente. A veces se ha sugerido la conveniencia de crear una jerarquía de categorías, que podría estar formada, por ejemplo, por categorías como: acceso no restringido, acceso con restricciones moderadas, acceso con altas restricciones, no accesible.

Con respecto a las *alianzas estratégicas*, se trataría de comprometer a algunas personas y entidades clave, especialmente dos colectivos. Por un lado, los editores de revistas científicas o sus consejos editoriales, que podrían establecer esta condición para la publicación de un artículo (en una primera fase sería sólo como sugerencia). Por otro, algunas de las principales entidades financiadoras de investigaciones en psicología podrían poner condiciones similares. Los proyectos del Plan Nacional y los de cobertura autonómica son el primer ejemplo (después de todo, se han obtenido mediante financiación pública; Law, 2005). Naturalmente, si estas entidades no proporcionan los medios para depositar los datos se ven obligados a confiar en que cada autor lo haga por sus propios medios (por ejemplo, páginas web de sus propias instituciones).

Ya que la psicología fuera de España no es un terreno completamente yermo en este aspecto, podríamos aprovechar las enseñanzas que se derivan de otras experiencias. Sumarnos a los proyectos internacionales que ya están en marcha es una opción con indudables ventajas (Johnson y Sabourin, 2001).

Conclusiones

Alguien podría pensar (como señala uno de los revisores de este artículo) que compartir es lo contrario de competir y que "...la competencia entre científicos es la principal fuerza que ha desarrollado la ciencia". Sin duda la competencia siempre es un vector de desarrollo, pero ésta debe restringirse al ámbito de las ideas, los modelos y las explicaciones alternativas, pero no debe aplicarse a la publicidad de los datos, especialmente una vez que hayan sido ya explotados por quien los ha generado.

Para la Psicología sería muy beneficioso fomentar una actitud de franca cooperación entre los investigadores, que incluya compartir abiertamente sus datos. Aparte de guiar nuestra investigación por un principio ético, ser un estandarte de transparencia y calidad y proteger de las posibles amenazas a la fiabilidad, creemos que esta acción fortalece el desarrollo de la Psicología como ciencia. Para ello hay que estar convencido de sus beneficios, pero también hay que proporcionar los medios y las oportunidades para hacerlo sin que suponga un incremento significativo de la carga de trabajo para los investigadores. Los editores de las revistas españolas y las entidades financiadoras son elementos clave para actuar en esta línea.

Agradecimientos

Este trabajo ha sido realizado, por parte del primer autor, en el marco del proyecto SEJ2006-12546 del Ministerio de Educación y Ciencia y de España.

Referencias

- American Psychological Association (2001). Ethical standards for the reporting and publishing of scientific information. In *Publication manual of the American Psychological Association* (5th ed., pp. 387-396). Washington, DC: Author.
- American Psychological Association (2006b). Evidence-Based Practice in Psychology. *American Psychologist*, 61(4), 271-285.
- Baddeley, A.D., Emslie, H., Kolodny, J., y Duncan, J. (1998). Random generation and the executive control of working memory. *Quarterly Journal of Experimental Psychology*, 51A, 819-852.
- Bem, S., y Looren de Jong, H. (1997). *Theoretical Issues in Psychology*. Londres: Sage Publications.
- Boring, E.G. (1980). *Historia de la Psicología Experimental*. México: Trillas [traducido de la 2ª edición inglesa, 1957].
- Botella J., y Gambará H. (2002). *Qué es meta-análisis*. Madrid: Biblioteca Nueva.
- Botella J., y Gambará H. (2006). El meta-análisis: una metodología de nuestro tiempo. *Infocop*, 29 mayo.
- Craig, J.R., y Reese, S.C. (1973). Retention of raw data: A problem revisited. *American Psychologist*, 28, 723.
- Chalmers, I., Hedges, L.V., y Cooper, H. (2002). A Brief History of Research Synthesis. *Evaluations and the Health Professions*, 25, 12-37.
- Eysenck, H.J. (1978). An exercise in mega-illness. *American Psychologist*, 33, 517.
- Freedland, K.E., y Carney, R.M. (1992). Data management and accountability in behavioural and biomedical research. *American Psychologist*, 47, 640-645.
- Friedman, S.L. (2007). Finding treasure: Data sharing and secondary analysis in developmental science. *Journal of Applied Developmental Psychology*, 28, 384-389.
- Glass, G. (1976). Primary, secondary and meta-analysis of research. *Educational research*, 5, 3-8.
- Johnson, D.H. (2001). Sharing data: It's time to end psychology's guild approach. *American Psychological Society Observer*, 14, 8, 1, 38-39.
- Johnson, D.H., y Sabourin, M.E. (2001). Universally accessible databases in the advancement of knowledge from psychological research. *International Journal of Psychology*, 36(3), 212-220.
- Johnson, R.W. (1964). Retain the original data. *American Psychologist*, 19, 350-351.
- Kuhn, T.S. (1970). *The structure of scientific revolutions* (2ª ed). Chicago: University of Chicago Press.
- Lipsey, M.W., y Wilson, D.B. (1993). The efficacy of psychological, educational and behavioural treatment: Confirmation from meta-analysis. *American Psychologist*, 48, 1181-1209.
- Mann, C.C. (1994). Can meta-analysis make policy? *Science*, 266, 961-963.
- Messick, S.M. (1988). The once and future issues of validity: Assessing the meaning and consequences of measurement. En H. Wainer y H. Braun (Eds.): *Test validity*. Hillsdale, NJ: Lawrence Erlbaum.
- Mollon, J.D., y Perkins, A.J. (1996). Errors of judgement at Greenwich in 1796. *Nature*, 380, 101-102.
- Nature (2007). A fair share (editorial). *Nature*, 444, 653-654.
- Ortego, C. y Botella, J. (2010). The hard but necessary task of gathering order-one effect size indices in meta-analysis. *Psicológica*, 31, 291-315.
- Petrosino, A., Boruch, R., Soydan, H., Duggan, L., y Sánchez-Meca, J. (2001). Meeting the challenges of evidence-based policy: The Campbell Collaboration. *Annals of the American Academy of Political & Social Science*, 578, 14-34.

- Sackett, D.L., Rosenberg, W.M., Gray, J.A., Haynes, R.B., y Richardson, W.S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal*, 312, 71-72.
- Sackett, D.L., Straus, S.E., Richardson, W.S., Rosenberg, W., y Haynes, R.B. (2000). *Evidence-Based Medicine: How to practice and teach EBM* (2ª ed). Londres: Churchill Livingstone.
- Smith, M.L., y Glass, G.V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32, 752-760.
- Spasser, M.A. (1997). The enacted face of undiscovered public knowledge. *Journal of the American Society for Information Science*, 48(8), 707-717.
- Swanson, D.R. (1987). Two medical literatures that are logically but not bibliographically connected. *Journal of the American Society for Information Science*, 38(4), 228-233.
- Towse, J., y Cheshire, A. (2007). Random number generation and working memory. *European Journal of Cognitive Psychology*, 19, 374-394.
- Wicherts, J.M., Borsboom, D., Kats, J., y Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *American Psychologist*, 61, 726-728.
- Wolins, L. (1962). Responsibility for raw data. *American Psychologist*, 17, 657-658.