

# Big data y Psicología: ¿una oportunidad para el Internet de las personas?

Manuel Armayones, Beni Gómez-Zúñiga, Eulàlia Hernández & Modesta Pousada  
Universitat Oberta de Catalunya

Recibido: 10-6-2015

Aceptado: 29-9-2015

## Big data y Psicología: ¿una oportunidad para el Internet de las personas?

**Resumen.** En este trabajo se reflexiona sobre el concepto big data y su relación con la Psicología en ámbitos como la evaluación, la intervención, la investigación y la docencia. El concepto big data hace referencia literal al uso de grandes bases de datos, pero también a las tecnologías y metodologías que permiten recogerlas a través de dispositivos de diversos tipos, entre los que destaca el teléfono inteligente o smartphone. La Psicología no puede ser ajena al desarrollo de esta tecnología, y debe trabajar a varios niveles para hacer posible que los profesionales accedan lo antes posible a estas herramientas, y se pase de la mera fascinación a la concreción de estudios y a la investigación para que contribuyan al avance de la disciplina.

**Palabras clave:** big data; smartphone; app; EMI; TIC

## Big Data and Psychology: an opportunity for the Internet of people?

**Summary.** In this paper we discuss the Big Data concept and its relation to psychology in areas such as assessment, intervention, research and teaching. The concept of Big Data refers to the use of large databases, but also to the technologies and methodologies that allow us to collect them through different devices, among which Smartphones are the most relevant. Psychology cannot ignore the development of this technology and must work on several levels to enable professionals, as soon as possible, to access these tools and shift from the mere fascination with Big Data to the development of studies and research that can contribute to the improvement of the discipline.

**Keywords:** Big Data; smartphone; app's; EMI; ICT

### Correspondencia

Dr. Manuel Armayones

Grup de Recerca PSiNET. Universitat Oberta de Catalunya

Estudis de Psicologia i Ciències de l'Educació

Universitat Oberta de Catalunya

Rambla del Poblenou, 156

08018 Barcelona

marmayones@uoc.edu

933263892/635474745

## Introducción

El término *big data* empieza a ser recurrente en el ámbito científico y académico, en general, y en el particular de la Psicología. En este trabajo se reflexiona sobre el concepto *big data* desde la perspectiva de la Psicología, analizando hasta qué punto su generalización, en el ámbito tanto profesional como de la investigación y académico, puede contribuir a un cambio de era en esta disciplina. Es decir, cómo puede llevarnos de una Psicología que a menudo ha visto limitado su alcance por la dificultad de acceso a grandes muestras de individuos en su contexto natural, a una Psicología que disponga de métodos que permitan el acceso a cantidades de información generada en el contexto natural de las personas procedente de muestras de miles de ellas.

La literatura apenas comienza a recoger artículos en los que se delibera sobre cómo la tecnología *big data* puede contribuir a revolucionar la investigación psicológica y, concretamente, la Psicometría (Markowetz, Błaszkiwicz, Montag, Switala, & Schlaepfer, 2014). También se empiezan a publicar investigaciones con muestras de miles de personas que analizan cómo su comportamiento en redes sociales se relaciona con características de su personalidad (Eftekhar, Fullwood, & Morris, 2014; Kosinski, Stillwell, & Graepel, 2013). De hecho, ya se están llevando a cabo estudios en los que se utiliza *big data* para la investigación de la depresión (Markowetz et al., 2014). En otros estudios se utiliza *big data* para presentar que una evaluación de la personalidad, basada en la actividad de 86.220 personas en Facebook, es tan precisa como la evaluación de la personalidad realizada «tradicionalmente» (Youyou, Kosinski, & Stillwell, 2015). Y aún en otros se utiliza el análisis combinado de los «me gusta» de Facebook y las pruebas de personalidad para discriminar con un análisis de regresión logística: orientación sexual, uso de sustancias adictivas, edad, género, opiniones políticas y características de la personalidad (Kosinski et al., 2013). En otros trabajos (Schwartz, et al., 2013) se relacionan variables de personalidad con el lenguaje natural de 75.000 voluntarios, y se pone en evidencia la relación entre dimensiones de la personalidad y el uso de determinadas palabras y frases después de analizar corpus de más de 700 millones de estas.

Los ejemplos anteriores son solo una pequeña muestra de los estudios que se desarrollan. Tanto es así, que la propia American Psychological Association en su página web reflexiona sobre potenciales nuevos puestos de trabajo como el de psicólogo *big data* (American Psychological Association, 2014). Otras instituciones como la Association for Psychological Science han empezado a organizar eventos y a recopilar reflexiones y experiencias sobre cómo el uso de *big data* puede contribuir a la investigación y a la práctica profesional de los psicólogos (Jaffé, 2014).

Sin duda, *big data* no sería posible sin el elevado número de dispositivos existentes, como son los teléfonos y relojes inteligentes (*smartphone* y *smartwatch*),

ropa con sensores, miles de aplicaciones móviles distintas (*app*, en adelante) y, más genéricamente, el llamado «Internet de las cosas» (*Internet of Things*, IoT, en su acepción inglesa). El Internet de las cosas permite, gracias a sistemas de identificación únicos del individuo (Kopetz, 2011) y a los cientos de miles de objetos de nuestra vida cotidiana que pueden conectarse entre ellos, como pueden ser determinados dispositivos médicos (por ejemplo, un glucómetro), sensores de ambiente (por ejemplo, apertura de puertas o luces), del refrigerador (por ejemplo, para registrar la cantidad de azúcar de los alimentos), etc., generar datos que, integrados y analizados correctamente, contribuyan a la toma de decisiones, tanto por parte de los individuos como de los profesionales sobre distintos temas, entre ellos los relativos al estudio del comportamiento humano.

En este ámbito de la «medición total» tenemos el ejemplo del movimiento *quantified self* que reúne a miles de personas que, a modo de «donantes de datos», monitorizan a través de distintos dispositivos todos los datos que generan, tanto a nivel biológico, físico, de comportamiento, de movimientos, de consumos, del ambiente en el que se mueven, etc. y generan bases de datos de libre acceso para la investigación en distintos ámbitos (Swan, 2013).

Sin duda, parece que estamos llegando a una época de retos fascinantes, pero también de preguntas que entre todos deberemos responder más pronto que tarde. Preguntas como: ¿estamos preparados para un tránsito tan rápido desde una Psicología que durante muchos años ha basado sus teorías y su enfoque más experimental en muestras relativamente pequeñas y en entornos experimentales muy controlados, a una Psicología que puede gestar y validar teorías en –prácticamente– tiempo real? ¿Afectará la velocidad que impondrá la posibilidad de analizar en tiempo real miles de variables no solo al desarrollo, la redefinición, e incluso a los cambios drásticos en la metodología científica que se enseña en las facultades de Psicología? ¿Qué efecto psicológico, positivo y/o negativo tendrá/está teniendo el Internet de las cosas en las personas?

No son preguntas fáciles de responder en estos momentos; pero lo que es evidente es que desde la Psicología no podemos abstraernos de un debate en el que, al menos, ya hay una conclusión clara: *big data* continuará desarrollándose tanto en su acepción literal de grandes muestras de datos, como en la dimensión más paradigmática, en tanto que inicio de una nueva manera de trabajar que afectará a las personas a las que atendemos, a los compañeros con los que trabajamos y al contexto económico y social en el que lo hacemos.

El objetivo de nuestro trabajo es contribuir a la reflexión sobre la encrucijada entre *big data* y Psicología, pero centrándonos en el estudio del individuo, es decir, en las oportunidades que tanto para la evaluación como para la intervención individual nos va a proporcionar el uso de *big data*. Consideramos que, aunque ya existen estudios en los que de facto se trabaja con grandes

muestras de datos, no se está produciendo el proceso de reflexión, crítica y análisis que consideramos imprescindible en nuestra disciplina y al que pretendemos colaborar con este trabajo.

Para ello y tras definir y reflexionar sobre el propio concepto *big data* vamos a estructurar nuestro trabajo en tres partes. En una primera, reflexionaremos sobre el papel que dispositivos como los *smartphone* y las *app* de las que estos disponen son la «unidad básica» para la recopilación de datos y se configuran en verdaderos instrumentos de investigación. En una segunda, nos centraremos en la potencial utilización de *big data* en evaluación e intervención y del efecto que las distintas iniciativas que se están llevando a cabo pueden tener en el desarrollo de nuevas teorías, así como el uso de *big data* para la validación en entornos reales de algunas de las actuales teorías. En una tercera, nos preguntaremos sobre algunas implicaciones que *big data* puede tener en el ámbito de la formación de los profesionales y, en un último apartado, apuntaremos una serie de conclusiones finales, así como propuestas para futuros trabajos en este ámbito.

### ¿Qué es *big data*?

Dans (2011) recoge una acepción muy concisa y clara. Para el autor, *big data* hace referencia al tratamiento y análisis de enormes repositorios de datos. Estos son tan desproporcionadamente grandes, en relación a los que se utilizan en la mayoría de las investigaciones en Psicología, que es imposible tratarlos con las metodologías y herramientas convencionales.

La recogida de datos para la toma de decisiones en las organizaciones se viene realizando desde hace muchos años. El concepto *Business Intelligence* ha sido el término clásico que ha definido cómo las organizaciones recopilan información de las fuentes que tenían a su disposición. En la mayoría de los casos, estos datos se refieren a parámetros propios de la organización, como podrían ser las ventas, el número de clientes, la evolución de los datos de producción, etc. Genéricamente, estos datos reciben el nombre de datos estructurados.

Las tecnologías *big data* permiten también la recogida y el procesamiento de información llamada no estructurada, que es la que consideramos más útil en nuestra disciplina. Algunos ejemplos de información no estructurada que pueden recogerse con un *smartphone* serían: el lenguaje natural, datos procedentes del comportamiento de las personas en redes sociales, datos recopilados por dispositivos que pueden llevar las personas (*wearables*, en su acepción inglesa), datos procedentes de los sensores de presión, temperatura, movimiento, orientación, sonido ambiental, etc.

Esta información no estructurada que se puede recopilar a partir de grandes muestras de personas y variables está referida, en un amplio porcentaje, a comportamientos humanos en su contexto natural, información recogida en tiempo real, almacenable y explotable en períodos breves de tiempo, lo cual con-

sideramos que abre grandes posibilidades a la investigación psicológica.

De alguna manera, el uso de *big data* nos va a permitir acceder a una información sobre el comportamiento de las personas, tanto individualmente como grupalmente, en una escala hasta ahora inimaginable (O'Reilly, 2011). En este sentido, *big data* va a provocar una revolución en todos los ámbitos y disciplinas y, desde nuestro punto de vista, y más allá de los aspectos puramente tecnológicos, requiere una reflexión algo más profunda, ya que puede ser la punta del iceberg de un nuevo salto evolutivo en Internet, como lo fue en 2004 la irrupción de la Web 2.0 (O'Reilly, 2004).

La Web 2.0 permitió que los usuarios de Internet dejaran de ser meros consumidores pasivos de los contenidos generados por aquellos que tenían los conocimientos técnicos, presupuesto, intereses o todo lo anterior, para pasar a una Internet participativa en la que cualquier internauta puede generar su propio contenido, compartir sus conocimientos y convertirse en «agente» de Internet. Entre las aplicaciones de la Web 2.0, las más conocidas son las redes sociales y comunidades *online*, los blogs, los wikis y, en definitiva, cualquier aplicación que permita una conversación global entre internautas con una mínima mediación tecnológica.

En nuestra opinión, *big data* es un salto evolutivo en la red, similar en su magnitud a la irrupción de las redes sociales, y cuyo efecto aumenta con cada nueva persona que crea un perfil en una red social, comenta en un blog, o incorpora información a Wikipedia.

*Big data* llega en un momento clave para la historia de Internet porque puede estar en la base de la respuesta al: «¿Qué? ¿Cómo? ¿Quién? y ¿Para qué?» con el que podríamos explicar y explicarnos un fenómeno como el de Internet, que desde su aparición ha cambiado drásticamente nuestra forma de vivir, trabajar, relacionarnos, cuidar de nuestra salud, educarnos y cualquier otro ámbito de nuestras vidas.

Así, la propia aparición de Internet fue un «¿Qué?» rotundo, que se responde pensando en Internet como la tecnología que permite poder relacionar computadoras separadas por miles de kilómetros y, con ellas, personas, instituciones y gobiernos.

El «¿Cómo?» se iba a producir materializando la idea inicial de los creadores de Internet, y que se hizo realidad con la irrupción de la Web 2.0 en 2004. El «¿Cómo?» haría referencia a la manera en la que iba a permitir a las personas relacionarse y, sin duda, las aplicaciones de la Web 2.0 con las redes sociales a la cabeza son la mejor respuesta.

La respuesta al «¿Quién?» la referiríamos a cuáles son los actores de la Web 2.0.; y estos no son otros que las personas, los millones de usuarios que hora tras hora hacen que empresas como Facebook, Twitter, Instagram, Flickr y tantas otras generen millones de dólares de ganancias, coticen en bolsa y tengan un papel cada vez más relevante en la sociedad, cuando su «producto» no es otro que los datos, muchas veces comportamentales, que gratuitamente les proporcionamos.

Y el «¿Para qué?» ¿Cómo explicaríamos el «¿Para qué?», o al menos un «¿Para qué?» que consideramos muy relevante para la Psicología?

Desde nuestro punto de vista, el «¿Para qué?» se responde, al menos en parte, con el propio concepto de *big data*. Así, ante la pregunta de «¿Para qué?» nos puede servir la agregación de datos de cientos de miles de personas, producidos por ellas mismas y en un contexto real; la respuesta podría ser: «Para saber», para saber todo aquello que en cada ámbito de la Psicología consideremos que nos puede ayudar a explicar el comportamiento humano. Y es obvio que entrar en la cuestión del para qué queremos saber nos va a ofrecer toda suerte de situaciones. Algunas estarán alineadas con los objetivos, principios y responsabilidades de la Psicología y sus profesionales, y otras en las antípodas de un uso responsable, ético y legal. Sin duda, los aspectos éticos y legales en el uso de los datos relacionados con variables psicológicas deberán ser tema de trabajo y reflexión.

### Psicología, teléfonos inteligentes y aplicaciones móviles en la era *big data*

Después de reflexionar sobre el concepto *big data*, el siguiente paso es plantearnos cómo se puede empezar a trabajar con grandes muestras y, sobre todo, cómo se pueden recoger tal cantidad de datos dinámicos y agregarlos para su investigación. Una primera respuesta sería: a través del teléfono inteligente y sus dispositivos periféricos conectados a servidores seguros y con la participación de profesionales expertos en el análisis de datos y en su interpretación.

¿Y por qué afirmamos tan rotundamente que el teléfono inteligente será la herramienta básica en la investigación psicológica con *big data*? Para responder esta cuestión hemos de tener en cuenta que el número previsto de líneas de telefonía móvil superará los 4.600 millones en todo el mundo y que en España hace años que el número de líneas superó al número total de habitantes (Ortega, 2010). Ningún dispositivo tiene la penetración social de los teléfonos inteligentes y ofrece una manera tan sencilla para recopilar datos útiles para la investigación.

Cuando nos referimos a los teléfonos incluimos todos los dispositivos periféricos que en estos momentos ya están sincronizados con él (relojes inteligentes, pulsómetros, gafas como las *Google Glass*) en los que el teléfono inteligente es la «unidad central de procesamiento» y desde donde se aportan datos explotables tanto en evaluación e intervención como en investigación.

Ya en un artículo con el sugerente título de *Smartphone Psychology Manifesto* (Miller, 2012) se apoya la idea de la utilidad del teléfono inteligente en la investigación psicológica, y se detalla cómo la mayoría de teléfonos inteligentes disponen de sensores, como acelerómetros que pueden aportar información sobre los movimientos y el ritmo de desplazamiento de una persona en distintos momentos del día, sensores de

proximidad, sistemas de geoposicionamiento, dispositivos para capturar imagen y sonido, tanto a modo de fotografías como registrando vídeos, que ofrecen un marco inimaginable hasta estos momentos para combinar el control experimental con la validez ecológica de los datos. Obviamente, en función del diseño y tipo de experimento, el balanceo entre control experimental y validez ecológica se deberá ajustar a cada situación aunque en algunos casos no sea posible totalmente. Pero, sin duda, se abren oportunidades hasta ahora desconocidas, y más si tenemos en cuenta que la posibilidad de trabajar con muestras de miles de individuos que voluntariamente compartan sus datos, como por ejemplo los citados miembros de los grupos *quantified self*, contribuirá a mejorar la calidad de los estudios.

Pero los beneficios de los teléfonos inteligentes como tal no van a ser tan determinantes como lo puede ser el desarrollo de aplicaciones, que son las que ponen en valor sus prestaciones técnicas.

Consideramos que el desarrollo de aplicaciones móviles (*app*) puede contribuir a la investigación en Psicología de varias maneras. Una de las que estimamos más importante, y que quizás es la más sencilla, es el establecimiento de convenios de colaboración científica con desarrolladores de *app* relacionadas con tópicos de nuestra disciplina, y que pueden ir desde la promoción del ejercicio físico, los hábitos de vida saludable, la adhesión terapéutica, los programas de intervención psicológica para la reducción de la ansiedad, para mejorar la autoestima, etc. En este sentido, el reto no va a ser tanto identificar las *app* y los grupos que las han desarrollado (para ello, una simple inspección en Google Play o en Apple Store es suficiente), sino valorar las circunstancias que pueden afectar a la calidad de los resultados obtenidos, y plantearnos una serie de alternativas y posibilidades para mejorar el control experimental de su desarrollo.

Hay que tener en cuenta que muchas de las *app* han sido creadas por empresas que no tienen por qué haber contado para el diseño de sus aplicaciones con teorías psicológicas, como pusieron de evidencia Cowan et al. (2012) en el análisis de 127 *app* de la categoría de Salud y Fitness de Apple Store. Los autores analizaron la presencia de constructos teóricos de la teoría de la acción razonada/acción planificada, el modelo transteórico y la teoría cognitivo-social. Para ello, analizaron si en las *app* se tenía en cuenta y/o se intervenía sobre constructos como la autoeficacia percibida, la prevención de recaídas, las barreras percibidas, la actitudes ante el comportamiento, etc. La conclusión de los autores fue que la mayoría de estas *app* no tienen en cuenta teorías validadas en Psicología, lo cual, sin duda, abre un amplio campo de trabajo para futuras investigaciones.

Indudablemente, el establecimiento de convenios y acuerdos con los desarrolladores de estas *app* para que en su diseño, y en la explotación de sus resultados, se incorporen psicólogos podría mejorar tanto su calidad como los resultados obtenidos por sus usuarios. A la vez, sería una oportunidad excepcional para explotar

los datos, e incorporar los mecanismos necesarios para que esta explotación cumpla los imprescindibles requisitos éticos y legales.

De este modo, podría existir un grupo de «*app* psicológicas» que, convenientemente testadas y evaluadas, y si llega el caso «homologadas» por parte de colegios o asociaciones profesionales, puedan contribuir a la necesaria agregación de datos de investigación con todas las garantías técnicas, de contenidos y legales exigibles.

En este sentido, iniciativas como la de la Junta de Andalucía: «estrategia de calidad y seguridad en aplicaciones móviles en salud» para evaluar la calidad de las *app* de salud (Junta de Andalucía, 2014) son proyectos que desde la Administración aportan valor añadido al uso de este tipo de aplicaciones y ejemplos sobre los que basarse. El sello de AppSaludable que otorga la Junta de Andalucía implica haber superado positivamente la evaluación tanto de parámetros estrictamente técnicos y de seguridad informática, como de elementos de calidad de los contenidos y aspectos legales de la privacidad y tratamiento de los datos.

Otras iniciativas para velar por la calidad de las *app* están relacionadas con el análisis de las evidencias publicadas en la literatura científica que avalan su utilización. Así, el proyecto *Beacon* (National Institute for Mental Health Research at the Australian National University, 2014) ofrece datos sobre la base de una escala de cinco puntos, en la que el nivel inferior indica que no hay evidencias científicas publicadas sobre los resultados obtenidos con una *app* determinada, el nivel cuatro señala que existen evidencias robustas publicadas, y el nivel cinco la considera una *app* que debería ser de uso prioritario tanto por haber sido evaluada y «homologada», como por el hecho de que permite la agregación de datos por un tercero con objetivos estrictamente científicos.

Por su parte, el índice MARS (Stoyanov et al., 2015) de la Universidad de Brisbane ofrece 372 criterios para evaluar *app* en ámbitos como la calidad de los contenidos, la funcionalidad, la promoción del compromiso e implicación de los usuarios y la estética. Esta escala además puede ser útil tanto para la clasificación y evaluación de *app* como un sistema de comprobación durante el propio proceso de diseño y desarrollo de las mismas.

Cabe destacar en este punto que en el ámbito de la Medicina, en el que las *app* están integrándose a gran velocidad, se distingue entre *app* clínicas (aquellas que se consideran instrumentos médicos y, por lo tanto, exclusivos para profesionales) y *app* de salud/bienestar de uso generalizado. Por consiguiente, es necesario disponer de sistemas de acreditación y de análisis de los potenciales riesgos que el uso de *app* clínicas puede tener para usuarios y para profesionales, entre los que nos podemos encontrar los ya comentados como riesgos en la privacidad de la relación, la falta de calidad de contenidos o de base teóricos o los aspectos técnicos como un funcionamiento inadecuado (Lewis & Wyatt, 2014)

### Evaluación e intervención ecológica momentánea y big data: bajando al terreno aplicado y subiendo al teórico

Pero más allá del tipo de dispositivo y del *software* que se utilice, o que la aplicación móvil sea «recreativa» o «clínica», lo realmente importante para analizar grandes muestras de datos es saber qué datos hay que recoger y cómo queremos hacerlo. Y en la mayoría de los casos, los profesionales pretendemos recoger datos que nos permitan evaluar o intervenir sobre variables relevantes para nuestro objetivo.

El objeto de nuestra evaluación puede ser, desde recoger datos a modo de «heterorregistro continuo» sobre el comportamiento de una persona sin la participación activa de esta (por ejemplo, geoposicionando sus recorridos por la ciudad para evaluar el tiempo que dedica a hacer deporte), hasta la solicitud de que durante un período de tiempo y realizando un muestreo adecuado a los objetivos de la evaluación, la persona complete cuestionarios, confirme que ha realizado una tarea, marque sobre una escala visual analógica algún tipo de variable como el dolor percibido, la ansiedad, etc.

Exactamente de la misma manera, a distancia, de manera programada y con una recogida de datos en tiempo real, la persona podrá recibir en el marco de una intervención psicológica, y a través de su teléfono inteligente, indicaciones, instrucciones, mensajes, solicitud de tareas y cualquier otra actividad que el profesional haya considerado pertinente en el diseño de su intervención.

Tenemos, por tanto, un *hardware* que es el teléfono inteligente, un *software* que serían las *app*, y debemos añadir los modelos teóricos, las hipótesis, y/o los objetivos de evaluación e intervención que son los que nos deben permitir afinar en la identificación de aquellas variables relevantes para nuestro objetivo.

Pero detengámonos en este punto, en dos conceptos que no son nuevos ni en evaluación ni en intervención psicológica, como son la *Ecological Momentary Assessment* (Shiffman, Stone, & Hufford, 2008) (EMA, en adelante, y que podemos traducir por Evaluación Ecológica Momentánea) y la *Ecological Momentary Intervention* (Runyan et al., 2013) (EMI, en adelante, y que podemos traducir por Intervención Ecológica Momentánea).

Como su nombre indica, tanto la evaluación como la intervención «momentánea» y «ecológica» hacen referencia a la posibilidad de evaluar determinadas variables en distintos momentos temporales y en el ambiente natural de la persona, aumentando así la validez ecológica de los resultados y minimizando problemas tan habituales como pueden ser las dificultades para recordar y registrar emociones, cogniciones y/o conductas.

La EMA no es, desde luego, una nueva metodología de recogida de datos ni ha aparecido recientemente tal como consideramos que sí lo ha hecho la EMI, con la irrupción de teléfonos inteligentes y *app* que permiten

una intervención, en tiempo real y durante las 24 horas del día, hasta ahora impensable.

Los estudios en los que se utiliza la EMA se desarrollaron sobre la base de agendas electrónicas (Palm o PDA), y aunque no permitían una sincronización en tiempo real con servidores en línea, sí que favorecieron una recogida de datos muy distinta a la tradicional de papel y lápiz en muchos ámbitos de la Psicología (Shiffman et al., 2008).

En estos momentos, y bajo una óptica de trabajo con *big data*, la distinción entre la EMA y la EMI se diluye por la posibilidad de realizar una intervención en línea a través de una aplicación móvil que se conecte a dispositivos y permita una retroalimentación continua entre los datos recogidos en la fase de evaluación y durante la intervención. Así, se habla de *Smartphone Ecological Momentary Assessment/Interventions* (Runyan et al., 2013) incorporando la evaluación y la intervención, pero asociándolas ya al teléfono inteligente como herramienta natural para llevarla a cabo al mismo tiempo que se generan cantidades ingentes de datos.

Pero lo verdaderamente relevante en relación con nuestra reflexión sobre *big data* y Psicología no es tanto ni el *hardware* ni el *software* que se utiliza, sino el hecho de que los datos recogidos se almacenan en servidores remotos, conocidos popularmente como «la nube», y más técnicamente con el término «computación en la nube». Así, la convergencia entre dispositivos que recogen datos y que pueden compartirlos en tiempo real con profesionales e investigadores es, sin duda, el elemento diferencial entre las primeras aproximaciones de trabajo de la EMA a través de las ya olvidadas PDA, y los actuales dispositivos. Pero además, y este es otro aspecto diferencial, las intervenciones de la EMI permiten ya un nivel de personalización de la intervención hasta ahora imposible (Wangberg & Psychol, 2013), lo que, sin duda, puede tener una gran importancia en el ámbito de la evaluación e intervención, pero aún más en el campo de la investigación y, concretamente, en el desarrollo de modelos teóricos.

Las intervenciones que se basan tanto en la EMA como en la EMI se fundamentan, en su mayoría, en modelos teóricos de partida. Así, por ejemplo, la decisión de si a un determinado grupo se le envía  $n$  mensajes durante  $x$  semanas, y además dichos mensajes tienen como objetivo intervenir sobre algunas variables determinadas como podrían ser: aumentar la percepción de autoeficacia, instrucciones sobre cómo mantener una conducta, incrementar la motivación y el compromiso del paciente hacia la intervención, reforzar positivamente los logros, motivar para el cambio hacia un estadio diferente, etc., pueden estar basados en teorías determinadas, como podría ser la del aprendizaje social o el modelo transteórico de Prochaska y DiClemente.

En consecuencia, las decisiones «técnicas», como establecer el número, la cadencia, y el tipo de mensaje, se basarán en un modelo teórico previo que será el que

«guíe» nuestra intervención. La pregunta que llegados a este punto nos deberíamos hacer es: ¿serán válidos en una intervención en línea los modelos teóricos que hasta ahora se han utilizado? (Riley et al. 2011); gracias a la posibilidad de trabajar con *big data* (simultáneamente con datos de contexto, individuales y de interacción), ¿la información obtenida cuestionará modelos que se diseñaron y validaron con muestras infinitamente menores y que puedan controlar muchas menos variables?; ¿influirá en los datos el hecho de que las personas dispongan de herramientas tecnológicas impensables cuando se desarrollaron muchos modelos y teorías psicológicas? Sin duda, no son preguntas fáciles de contestar y sí generadoras de una gran controversia que solo se podrán contestar con datos de investigación en la mano.

Por otro lado, las teorías en las que se basan las intervenciones psicológicas van a precisar adaptarse a situaciones en las que estas intervenciones puedan llevarse a cabo a través de dispositivos como los teléfonos inteligentes, que no existían en el momento en el que fueron formuladas y que permiten a la persona una interacción mayor, en tiempo real, en cualquier lugar, e incluso una adaptación continua a sus características individuales. Valga en este sentido plantear la analogía con la Medicina, en la que ya se habla de una Medicina 4P (Hood & Flores, 2012), es decir, una Medicina personalizada, preventiva, participativa y predictiva gracias a la posibilidad tanto de monitorizar y explotar los datos de *big data* como de implicar mucho más activamente a los pacientes. ¿Puede ser el uso de *big data* un catalizador similar en Psicología, y más si nuestra disciplina ya ha considerado siempre la individualización de la intervención y la participación de los pacientes? ¿Puede contribuir a aumentar el nivel de personalización de las intervenciones, la participación de los pacientes y posibilitar una mayor predicción de la conducta? Creemos que para resolver estas cuestiones será necesario un trabajo de revisión de las distintas teorías para ver cómo se han operacionalizado en su tránsito a la intervención a través del *smartphone*, pero sobre todo en los resultados de investigación que se están generando. Es posible que algunas teorías vean reforzada su validez con tanto que la posibilidad de poder disponer de las ventajas de uso de teléfonos inteligentes y sus dispositivos permita una intervención de gran calidad. En otros casos, podría suceder que la explotación de grandes bases de datos haga emerger nuevas teorías en las que se incorporen no únicamente la presencia de nuevos dispositivos sino en las que determinadas variables, como podrían ser la percepción del logro, la norma subjetiva o cualquier otro tipo, adquieran mayor o menor importancia en una sociedad que no es la misma que la que vio emerger las teorías básicas en Psicología.

Yendo más allá y en un tono provocador, algunos autores sostienen que *big data* va a representar el final de las teorías e incluso del propio método científico tal y como lo conocemos hoy en día (Anderson, 2008). El posicionamiento extremo de Anderson fue rápidamente

te contestado por Pigliucci (2009) que apuntaba la necesidad de no sobrestimar *big data*, en el sentido de que las correlaciones espurias que pueden observarse en las grandes muestras de datos, puede que no tengan ningún tipo de valor explicativo de los fenómenos observados. De la misma manera, la mera acumulación de datos, por grande que esta sea, no implica la generación de conocimiento, ya que en realidad el objetivo de los científicos debe ir mucho más allá de encontrar patrones (un producto habitual en los análisis *big data*) si no de ser capaces de encontrar explicaciones a los fenómenos observados, lo cual, sin duda, va a precisar de psicólogos bien formados. Sea como sea, lo cierto es que en el ámbito de la llamada eSalud, es decir, de la incorporación de las nuevas tecnologías en el ámbito de la evaluación, intervención e investigación, ya hay autores que consideran que en estos momentos es fácil que los avances tecnológicos estén progresando a mayor velocidad que la de los investigadores para adaptarse a ellos (Potts, 2006).

### Big data en la formación de los futuros psicólogos

Nos equivocáramos si pensáramos que estamos ante un fenómeno que irrumpirá de manera brusca en nuestros laboratorios, en la academia y en la profesión. *Big data* ya es una realidad en Psicología (APA, 2014).

La figura del *Big Data Psychologist* se configura como una de las salidas profesionales para los psicólogos con sólida formación en metodología, psicometría y estadística aplicada a las ciencias del comportamiento. Es por ello que consideramos que en la formación de los futuros psicólogos se van a tener que incorporar contenidos y asignaturas que les permitan abordar las ingentes cantidades de datos con metodologías y técnicas que no eran necesarias hasta la actualidad. Así, habrá que incorporar nuevos diseños de investigación que permitan evaluar el cambio dinámico de los datos a tiempo real y a un tiempo el contexto físico, la conducta individual y la interacción de la persona; también definir nuevas variables (por ejemplo, la sesión de WhatsApp) o redefinir las existentes (por ejemplo, veracidad de la información en distintos dispositivos).

Pero más allá de los aspectos metodológicos y técnicos, que serán perfectamente abordados por los especialistas del área, creemos que el reto principal de los psicólogos será el de mantener no solo un alto nivel teórico para aportar los conocimientos que desde la Psicología permitan tanto hacer «las preguntas correctas a los datos» como, y sobre todo, para aportar una visión crítica frente a la fascinación que *big data* puede provocar.

Otros aspectos en los que se tendrá que profundizar en la formación serán el análisis que, desde la deontología profesional, se va a tener que hacer para, nuevamente huyendo de la fascinación por *big data*, saber diferenciar entre lo que podemos saber y lo que necesitamos saber de una persona para contribuir a la mejora de su salud mental y calidad de vida. Este objetivo inherente a la profesión, nos obligará a posicionarnos,

quizás de una manera mucho más crítica y firme que hasta ahora, en los límites entre los objetivos de determinadas organizaciones, y el velar por la integridad psicológica de la población. En este campo se deberá trabajar de manera interdisciplinar con profesionales tanto del ámbito jurídico como tecnológico, y sin duda el tan valorado trabajo interdisciplinar que se enseña, promueve y apoya desde las facultades de Psicología va a verse muy reforzado.

Ni en la evaluación, intervención o investigación el psicólogo va a poder trabajar ajeno a grupos de investigación en los que participen informáticos, pero también expertos en el análisis y la explotación de *big data*. Sin duda, creemos que dicho trabajo, más intenso aún del que se ha venido desarrollando hasta ahora, va a representar una nueva manera de avanzar en el conocimiento, la descripción y la explicación de la conducta humana.

### Conclusiones

Consideramos que los cambios que se van a dar en la profesión del psicólogo debido a la incorporación de *big data* tanto en su práctica profesional como en la investigación psicológica van a ser importantes. Gracias a la facilidad y flexibilidad en la recogida de datos, se va a poder estrechar la clásica brecha entre la psicología aplicada y la académica. Pero para ello va a ser necesario un trabajo conjunto que permita el avance disciplinar en temas tan concretos como la acreditación y homologación de *app* psicológicas que cuenten con todas las garantías científicas, tanto en lo relativo a su eficacia, eficiencia y utilidad como en su capacidad de contribuir a la agregación de datos.

Dicha agregación de datos, a su vez, va a permitir un desarrollo teórico a una velocidad y con una precisión desconocidas, pero sin olvidar que la agregación de datos que permite *big data* nunca podrá, por sí sola, explicar el comportamiento humano. Este, como no puede ser de otra manera, deberá ser analizado y explicado por los profesionales de la Psicología, cuyo objetivo ha de seguir siendo el mismo: comprender la conducta humana y explicarla sobre la base de datos significativos que nos permitan intervenir de manera adecuada con metodologías e instrumentos nuevos pero siempre manteniendo el mismo objetivo.

Estos profesionales deberán formarse en grado, y sobre todo, en posgrado, para ser capaces de incorporar a su bagaje metodológico nuevas herramientas y estrategias de trabajo. Así, y ante un base de datos como antes nunca había estado a disposición de los psicólogos, estos tendrán criterios técnicos que les permitirán «hacer las preguntas adecuadas a los datos», pero también serán capaces de moverse en la incertidumbre con una actitud de observación para saber discriminar, ante un conjunto de información de naturaleza cambiante, entre datos relevantes y ruido.

En el método científico, a una pregunta de investigación le seguía la formulación de una hipótesis, el diseño de una estrategia para validar (o no) dicha hi-

pótesis, y de los datos derivados se extraían unas conclusiones que a su vez generaban nuevas preguntas. En estos momentos, *big data* está poniendo frente a nosotros los datos, brutos en muchos casos, sin depurar y en cantidades difícilmente asumibles por equipos que no sean multidisciplinares. Al psicólogo corresponderá llegar a pulir los datos para que la información se transforme en conocimiento, sin atajos ni una fascinación desmesurada, pero sí con la firme convicción de que *big data* nos va a permitir, tanto como a obligar, cambiar nuestra manera de pensar, investigar y enseñar Psicología.

Sin duda, entramos en una era deslumbrante que va a marcar el desarrollo de nuestra disciplina. Un momento histórico en el que la posibilidad de «medirlo todo» puede generar nuevos miedos y nuevas ansiedades que deberán ser abordados. Creemos, en este sentido, que los profesionales de la Psicología tendremos que asumir nuestra parte de responsabilidad en conseguir que el Internet de las cosas, imparabile ya en su desarrollo, se convierta en el Internet de las personas.

## Referencias

- Anderson, C. (2008). The end of theory: The data deluge makes the scientific method obsolete. *Wired*. Retrieved June 2, 2015, from [http://archive.wired.com/science/discoveries/magazine/16-07/pb\\_theory/](http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory/)
- APA (2014, June 28). *Hot jobs: Big-data psychologists*. <http://www.apa.org>. Retrieved from <http://www.apa.org/gradpsych/2013/01/big-data.aspx>
- Cowan, L. T., Wagenen, S. A. Van, Brown, B. A., Hedin, R. J., Seino-Stephan, Y., Hall, P. C., & West, J. H. (2012). Apps of Steel: are exercise apps providing consumers with realistic expectations? A content analysis of exercise apps for presence of behavior change theory. *Health Education & Behavior, 40*(2), 133-139. doi:10.1177/1090198112452126
- Dans, E. (2011). Big Data: una pequeña introducción. Recuperado de <http://www.enriquedans.com/2011/10/big-data-una-pequena-introduccion.html>
- Eftekhari, A., Fullwood, C., & Morris, N. (2014). Capturing personality from Facebook photos and photo-related activities: How much exposure do you need? *Computers in Human Behavior, 37*, 162-170. doi:10.1016/j.chb.2014.04.048
- Hood, L., & Flores, M. (2012). A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory. *New Biotechnology, 29*(6), 613-624. doi:10.1016/j.nbt.2012.03.004
- Jaffé, E. (2014). What Big Data Means For Psychological Science - Association for Psychological Science. *Observer-Psychological Science, 27*(6). Retrieved from <http://www.psychologicalscience.org/index.php/publications/observer/2014/july-august-14/what-big-data-means-for-psychological-science.html>
- Junta de Andalucía (2014, September 15). Estrategia de calidad y seguridad en aplicaciones móviles de salud. Retrieved May 28, 2015, from <http://www.calidadappsalud.com/>
- Kopetz, H. (2011). *Real-Time Systems*. Boston, MA: Springer US. doi:10.1007/978-1-4419-8237-7
- Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences of the United States of America, 110*(15), 5802-5. doi:10.1073/pnas.1218772110
- Lewis, T. L., & Wyatt, J. C. (2014). mHealth and mobile medical Apps: a framework to assess risk and promote safer use. *Journal of Medical Internet Research, 16*(9), e210. <http://doi.org/10.2196/jmir.3133>
- Markowetz, A., Błaszczewicz, K., Montag, C., Switala, C., & Schlaepfer, T. E. (2014). Psycho-Informatics: Big Data shaping modern psychometrics. *Medical Hypotheses, 82*(4), 405-11. doi:10.1016/j.mehy.2013.11.030
- Miller, G. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science, 7*(3), 221-237. doi:10.1177/1745691612441215
- National Institute for Mental Health Research at the Australian National University. (2014). Beacon2.0. Retrieved May 29, 2015, from <https://beacon.anu.edu.au/pages/about>
- O'Reilly, T. (2004). What Is Web 2.0 - O'Reilly Media. Retrieved from <http://oreilly.com/web2/archive/what-is-web-20.html>
- O'Reilly, T. (2011). What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. *SSRN eLibrary*. Retrieved from [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1008839](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1008839)
- Ortega, J. M. C. (2010). Smartphone: toda la información al alcance de tu mano. Retrieved from <http://philpapers.org/rec/ORTSTL>
- Pigliucci, M. (2009). The end of theory in science? *EMBO Reports, 10*(6), 534. doi:10.1038/embo.2009.111
- Potts, H. W. W. (2006). Is e-health progressing faster than e-health researchers? *Journal of Medical Internet Research, 8*(3), e24. doi:10.2196/jmir.8.3.e24
- Riley, W. T., Rivera, D. E., Atienza, A. A., Nilsen, W., Allison, S. M., & Mermelstein, R. (2011). Health behavior models in the age of mobile interventions: are our theories up to the task? *Translational Behavioral Medicine, 1*(1), 53-71. doi:10.1007/s13142-011-0021-7
- Runyan, J. D., Steenbergh, T. A., Bainbridge, C., Daugherty, D. A., Oke, L., & Fry, B. N. (2013). A smartphone ecological momentary assessment/intervention «app» for collecting real-time data and promoting self-awareness. *PloS One, 8*(8), e71325. doi:10.1371/journal.pone.0071325
- Schwartz, H.A., Eichstaedt, J.C., Kern, M.L., Dziurzynski, L., Ramones, S.M., Agrawal, M., et al. (2013). Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. *PLoS ONE 8*(9): e73791. doi:10.1371/journal.pone.0073791
- Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of*

*Clinical Psychology*, 4(1), 1–32. doi:10.1146/annurev.clinpsy.3.022806.091415

Stoyanov, S. R., Hides, L., Kavanagh, D. J., Zelenko, O., Tjondronegoro, D., & Mani, M. (2015). Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps. *JMIR mHealth and uHealth*, 3(1), e27. doi:10.2196/mhealth.3422

Swan, M. (2013). The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery. *Big Data*, 1(2), 85–99. doi:10.1089/big.2012.0002

Wangberg, S. C., & Psychol, C. (2013). Personalized technology for supporting health behaviors. In *2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom)* (pp. 339–344). IEEE. doi:10.1109/CogInfoCom.2013.6719267

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences*, 112(4), 201418680. doi:10.1073/pnas.1418680112

### Big data i Psicologia: una oportunitat per a l'Internet de les persones?

**Resum.** En aquest treball es reflexiona sobre el concepte big data i la seva relació amb la Psicologia en àmbits com l'avaluació, la intervenció, la investigació i la docència. El concepte big data fa referència literal a l'ús de grans bases de dades, però també a les tecnologies i metodologies que permeten recollir-les a través de dispositius de diversos tipus, entre els quals destaca el telèfon intel·ligent o smartphone. La Psicologia no pot ser aliena al desenvolupament d'aquesta tecnologia, i ha de treballar a diversos nivells per fer possible que els professionals accedeixin al més aviat possible a aquestes eines, i es passi de la mera fascinació a la concreció d'estudis i a la investigació perquè contribueixin a l'avanç de la disciplina.

**Paraules clau:** big data; smartphone; app; EMI; TIC

