

R COMO ENTORNO PARA EL ANÁLISIS ESTADÍSTICO EN EVALUACIÓN PSICOLÓGICA

R AS THE ENVIRONMENT FOR DATA ANALYSIS IN PSYCHOLOGICAL EVALUATION

Ana María Ruiz-Ruano y Jorge L. Puga

UCAM Universidad Católica de Murcia

R es un entorno libre para el análisis estadístico de datos y para la generación de gráficos que está cobrando un papel clave en un amplio espectro de áreas de conocimiento. Pese a que el usuario ha de interactuar con el programa, principalmente, por medio de una consola de comandos; se están invirtiendo esfuerzos en proporcionar entornos gráficos amigables que aproximen al usuario novel a esta especie de lengua franca del análisis estadístico contemporáneo. Además, la flexibilidad y versatilidad del programa, en conjunción con una comunidad de desarrolladores y usuarios global, facilitan que R sea la herramienta elegida en muchos contextos básicos o aplicados. En este trabajo se presenta la utilidad potencial que supone R para la evaluación psicológica en general a la vez que se acentúa el rol que asume la estadística propiamente dicha como instrumento de progreso de la psicología al amparo del método científico.

Palabras clave: R, Evaluación psicológica, Análisis estadístico, Psicometría.

R is a free computing environment for statistical data analysis and graph creation. It is becoming a key tool in a wide range of knowledge domains. The interaction with the software is mainly based on a command line interface but efforts are currently being made to develop friendlier graphical user interfaces that will help novice users to become familiar with this programming language. R is a flexible and powerful system thanks to the development community that is working together to improve its capabilities. As a result, it is the chosen statistical software in many applied and basic contexts. This paper highlights the potential usefulness of R for psychological assessment and related areas. Additionally, the relevance of statistical data analysis is emphasised as an instrument that will boost the progress of psychology under the umbrella of the scientific method.

Key words: R, Psychological assessment, Statistical analysis, Psychometrics.

El análisis de datos estadístico ha cobrado un papel crítico en la cultura contemporánea. No es por ello extraño que se reclame cierto grado de alfabetización estadística en el conjunto de la sociedad con el objetivo de aspirar a nuevas cotas de desarrollo humano (p.e., Bond, 2009; Seldmeier y Gigerenzer, 2001). Para la ciencia, que aspira a conocer la naturaleza, y para la técnica, que anhela controlarla, la estadística se desdibuja como la herramienta clave que orienta y conduce el progreso de las comunidades humanas. Sin embargo, pese a que hoy en día disponemos de ingentes algoritmos de cómputo estadístico y que podemos implementarlos en computadoras lo suficientemente potentes como para manejar un volumen de datos considerablemente grande, parece que no somos capaces de sacar provecho de esta situación. Conviendría recordar aquel eslogan que popularizó Pirelli (la compañía italiana de neumáticos) a mediados de la década de los 90 del pasado siglo porque podría aplicarse claramente a esta situación: "la potencia sin control no sirve de nada" ("*power is nothing without control*"). En la actualidad contamos con ordenadores muy potentes, softwares informáticos de análisis de datos fáciles de manejar y, además, es relativamente sencillo recoger datos para realizar estudios. No obstante, y pese a todas estas facilidades con las que nos encontramos (*potencia*), es crucial realizar una buena planificación de lo que se quiere hacer y un buen diseño de la investigación (*control*), teniendo en consideración los análisis de datos que se realizarán *a posteriori*, para contrastar las hipótesis de partida (p.e., Bakan, 1966; Cohen,

1990; Dar, Serlin y Omer, 1994; Munafò et al., 2014; Rosnow y Rosenthal, 1989). Es decir, si nuestro estudio carece de esa planificación y diseño, aunque tengamos todas las herramientas y todos los datos a nuestra disposición, no nos servirán de nada. Y, contrariamente a como pudiese pensarse, el hecho de ser flexibles en el análisis de datos que se realiza tras una planificación anómala o inexistente no ayuda a que la investigación avance satisfactoriamente, sino que, más bien, conduce a unos resultados infructíferos o rebuscados y sin sentido científico (Ioannidis, 2005).

Estamos encontrando en los últimos años múltiples denuncias de malas prácticas en la investigación (p.e., Button et al., 2013; Cohen, 1990; Gigerenzer, 2004; Munafò et al., 2014; Trafimow y Marks, 2015; Vaux, 2012; Weissgerber, Milic, Winham y Garovic, 2015) que nos presentan un panorama, cuando menos, aterrador en lo relativo al análisis de datos estadístico. Incluso hay autores como Ioannidis (2005) que se atreven a asegurar que muchos de los trabajos de investigación que se presentan son falsos y que, de hecho, la mayoría de los descubrimientos que se realizan no son más que una evidencia de esos sesgos existentes en los diferentes sub-campos de investigación. En cuanto a las malas prácticas, por ejemplo, Munafò et al. (2014) critican los sistemas de calidad imperantes en ciencia, la baja reproducibilidad de los estudios, la existencia de cierto sesgo en las publicaciones científicas auspiciadas por las políticas editoriales de las revistas científicas, el falseamiento de los datos, la baja potencia estadística de los resultados que se presentan, el desarrollo de prácticas ilícitas en investigación y el uso de los valores *p* de manera generalizada como elemento crítico de inferencia estadística. Por su parte, Button et al. (2013) se centran en la denuncia de la baja potencia estadística que presentan los resultados de las investigaciones neurocientíficas y del reducido tamaño muestral que se utiliza en la mayoría de los estudios realizados. Además, resalta la importancia de tener en

Correspondencia: Jorge López Puga. Facultad de Ciencias de la Salud. Departamento de Ciencias de la Salud. UCAM Universidad Católica de Murcia. Ctra. de los Jerónimos s/n. 30107 Guadalupe - Murcia. España. E-mail: jpuga@ucam.edu

consideración las implicaciones éticas del sacrificio de animales (véase también p.e., Cressey, 2015; Editorial Nature, 2015) y de la inversión económica que se realiza para después no obtener, en la mayoría de los casos, resultados útiles para la sociedad.

Una de las temáticas que más atención sigue recibiendo en cuanto a su mal uso es el contraste de hipótesis basado en la hipótesis nula o NHSTP (*Null Hypothesis Significance Testing Procedure*) (p.e., Bakan, 1966; Cohen, 1994; Gigerenzer, 1998; Haller y Krauss, 2002; Halsey, Curran-Everett, Vowler y Drummond, 2015; Nieuwenhuis, Forstmann y Wagenmakers, 2011; Rosnow y Rosenthal, 1989; Trafimow, 2014; Trafimow y Marks, 2015) y el uso abusivo del p valor como único elemento que guía la inferencia estadística en los estudios científicos. Estas malas prácticas ya fueron resaltadas por Bakan (1966) aludiendo a que era algo que atentaba flagrantemente contra la lógica común y que “cuando llegamos a un punto en el que los procedimientos estadísticos son sustitutos en vez de ayudas al pensamiento, y concluimos absurdos entonces debemos retornar a las bases del sentido común” (p. 436). Igualmente, Cohen (1994) y Gigerenzer (1998) avisaban de la importancia de hacer un buen uso del contraste de hipótesis y no convertirlo en un ritual que se ha institucionalizado en una cultura científica que repite sistemáticamente los mismos errores a lo largo de la historia (Gigerenzer, 2004).

Una de las medidas más drásticas y controvertidas que se han tomado recientemente con relación al uso del contraste de hipótesis nulas fue la de la revista *Basic and Applied Social Psychology* que, en la editorial del año 2014 (Trafimow, 2014), propuso a los autores el uso de una metodología estadística diferente a la del procedimiento de la hipótesis nula que utilizara el valor p como único valor para guiar la inferencia estadística, y que prohíbe explícitamente en la editorial de 2015 (Trafimow y Marks, 2015). En este sentido, piden a los autores que exploten los datos a nivel descriptivo, y que se haga uso de otro tipo de análisis de datos que puedan llevar a alcanzar resultados de mayor calidad científica. De hecho, señalan que contrariamente a lo que muchos investigadores podemos pensar, una $p < 0,05$ es muy fácil de conseguir y algunas veces incluso sirve como excusa para publicar investigaciones de poca calidad. Existen evidencias que justifican un mal uso del p valor de un test estadístico. Por ejemplo, encontramos el trabajo desarrollado por Haller y Krauss (2002) en el que tras observar las malas interpretaciones que se hacían del valor p en psicología, concluyeron que tanto profesores como alumnos tenían deficiencias en la interpretación de la significación de los resultados de investigación. Sin embargo, Leek y Peng (2015) sugieren que esa prohibición no garantizará un incremento en la calidad de la *ciencia* que se haga, además de que el p valor supone “la punta del iceberg” dentro del mundo de la estadística o del análisis de datos.

La psicología como ciencia se ha ido proveyendo de una serie de herramientas de medida para desarrollar su quehacer en sus diferentes ámbitos de estudio que están estrechamente ligados al análisis estadístico de datos. Ya sea en la neuropsicología, la psicología clínica, educativa u organizacional, por ejemplo, es muy habitual el uso de tests. De hecho, como señalan Hernández, Tomás, Ferreres y Lloret (2015), los tests se consideran herramientas básicas en la evaluación psicológica que guían la toma de decisiones del psicólogo. Es por ello que deben cumplir con una serie de garantías psicométricas para poder ser usados de manera correcta que deben estar a disposición del usuario aplicador del test. Hernández et al. (2015) señalan que algunas de las

informaciones que no se proporcionan y que serían muy útiles son aquellas relacionadas con las evidencias sobre la validez de contenido, validez predictiva, fiabilidad como estabilidad en la medida, el análisis del sesgo o el funcionamiento diferencial de los ítems. A este respecto también señalan que es responsabilidad del profesional, a pesar de que su campo de trabajo no sea el de la psicometría, formarse e informarse sobre los avances psicométricos. No obstante, y al margen de las herramientas que se utilicen en la intervención o investigación psicológica, la psicología como ciencia necesita del método científico en la práctica diaria. En todas estas situaciones, que no son ajenas al profesional de la psicología, el análisis estadístico de datos cobra un protagonismo importante. Por ello, al igual que Dar et al. (1994), podríamos argumentar que las claves para la realización de un buen análisis de datos, independientemente del área de conocimiento psicológico que estemos considerando, radica en partir de una buena teoría de base y de realizar un buen diseño de la investigación que quiere llevarse a cabo con el fin de prevenir las debilidades que se achacan al procedimiento de contraste de hipótesis nulas.

El objetivo principal de este trabajo es presentar el software R como un entorno para la realización de análisis estadísticos que puede resultar muy útil para la psicología en general. Como ya veremos más abajo, este entorno puede favorecer el aprendizaje de la técnica estadística que estamos utilizando. A continuación encontramos un epígrafe sobre el entorno de R, en el que se hace una presentación del mismo desde un punto de vista histórico, filosófico y de la interacción del usuario con el mismo. Seguidamente, haremos un pequeño recorrido por algunas de las utilidades que ofrece R para la evaluación psicológica, presentando los distintos paquetes y la aplicabilidad de los mismos a esta disciplina. Por último, terminaremos con una reflexión sobre las posibles soluciones y mejoras para tratar de poner en marcha una investigación útil y de calidad para la sociedad en general con el apoyo del análisis estadístico donde R juega un papel crucial a día de hoy.

ENTORNO R

R es un entorno de trabajo para la ejecución de análisis estadísticos y la creación de gráficos (R Development Core Team, 2011). La interfaz gráfica del programa es una consola de comandos, es decir, que para interactuar con el mismo hay que escribir líneas de código y ejecutarlas. Podríamos trazar el origen de R hasta la década de 1970, cuando se acuñó el nombre de S para denominar a un lenguaje de programación de alto nivel destinado a realizar cálculos estadísticos. El cambio hacia R se produce en la década de 1990 coincidiendo con la expansión masiva de la versión comercial de S, S-Plus. Ross Ihaka y Robert Gentleman (*University of Auckland*, Nueva Zelanda) escribieron una versión educativa de la primera distribución de R, que fue dada a conocer a la comunidad científica general en 1996 (Ihaka y Gentleman, 1996). R se estructuró con base al lenguaje de programación S y Scheme, un dialecto de Lisp. Desde sus inicios, intentaron que fuese un lenguaje multiplataforma que pudiese ser utilizado bajo distintos sistemas operativos. En 1997 se crea el *R Core Team* como órgano responsable del desarrollo de la base del software y que proporciona apoyo continuo al desarrollo y distribución de R (Fox, 2009). Desde entonces, R además de seguir desarrollándose y creciendo, está consolidándose como la herramienta estadística de referencia por excelencia. De hecho, Tippmann (2015) señala que R será el software de análisis de datos más utilizado en 2015 a la par de, o incluso superando a, otros softwares comerciales.

El hecho de ser considerado un entorno de trabajo quiere decir que R es un lenguaje de programación que viene equipado con un conjunto de herramientas destinadas al cálculo y generación de gráficos estadísticos (Ihaka y Gentleman, 1996). Este entorno de trabajo fue hecho, y es, libre en el sentido de que no es necesario pagar por utilizar, copiar y distribuir el programa (Carleos y Corral, 2013). Además, está protegido frente a una posible privatización y/o comercialización del mismo bajo una licencia GNU. La filosofía GNU GPL (*General Public License*) es la base de la libertad y apertura de código, y cuyo propósito es proteger los derechos y libertades de los usuarios finales (<http://www.fsf.org>). Los pilares fundamentales en los que se sustenta R es el permitir usar, compartir, estudiar y modificar el código informático que hay a la base del programa. Es decir, cualquier usuario de R puede desarrollar una nueva aplicación del mismo, puede modificar lo ya existente, puede compartirlo y usarlo libremente.

R se compone de una estructura *base* que contiene una serie de funciones básicas para realizar cálculos y gráficos estadísticos. Por otro lado, se han creado una serie de complementos llamados *paquetes* por parte de la comunidad científica que realizan análisis estadísticos de datos más específicos. Dicha actividad de desarrollo informático puede considerarse que se realiza de una manera altruista y que atiende más a factores intrínsecos de la persona –como la satisfacción producida por dicha colaboración, la recompensa intrínseca de ese trabajo o el hecho de estar contribuyendo a un bien público–, que a factores externos como el reconocimiento científico o el dinero (Fox, 2009). Un paquete de R es, simplemente, “un conjunto de funciones que mantienen algún tipo de relación entre ellas” (Elosua, 2011, pp. 24-25). Por ejemplo, existen paquetes para el análisis estadístico de aspectos tan variados como el contenido de textos electrónicos, la distribución espacial de los núcleos urbanos o la activación neural registrada por una técnica de neuroimagen. Todos ellos están a libre disposición en Internet de manera gratuita. El CRAN (*Comprehensive R Archive Network*, <http://www.r-project.org>) es el lugar de la red desde donde se gestionan todas las descargas. Para un mayor detalle relativo a la descarga e instalación de R y de sus paquetes véase, por ejemplo, Elosua (2009) o López (2012, 2013).

La principal ventaja que presenta R es la libertad y la gratuidad de su uso frente a otros softwares de análisis estadístico. Otras ventajas que presenta el uso de este entorno de trabajo es el hecho de que favorece un aprendizaje básico de estadística (López, 2012; Tippmann, 2015) y que puede ser ejecutado en diferentes sistemas operativos (Fox, 2009) como Windows, Mac OS o Linux. Además, Huber et al. (2015) señalan que el uso de R como programa estadístico facilita la investigación y la innovación porque permite el desarrollo de prototipos rápidos, es flexible y funcional, permite la reproductibilidad, los cambios o modificaciones se realizan con rapidez, aporta facilidades gráficas y permite la interacción con otros lenguajes de programación como C y C++ o JavaScript para las aplicaciones web. Con relación a la interpretación de las salidas que genera tras la realización de un análisis estadístico podrían también señalarse ciertas ventajas. Dado que el usuario del software tiene que inmiscuirse más a fondo en la herramienta estadística que está utilizando desde el punto de vista técnico, acaba siendo más hábil para interpretar los resultados que generan los análisis de datos que aplica. Como inconveniente podemos señalar lo arduo de trabajar con un entorno de comandos frente a otros programas o interfaces donde únicamente hay que clicar (p.e., López, 2013; Tippmann,

2015). Sin embargo, como se verá más abajo, cada vez son más y mejores las interfaces gráficas de usuario que se han venido desarrollando para hacer más amigable la interacción con R.

La interfaz gráfica de R, como hemos comentado anteriormente, se basa en líneas de comandos. No obstante, para facilitar la interacción con este programa se han desarrollado interfaces gráficas más intuitivas para los usuarios. Estas interfaces pueden clasificarse en dos tipos, el primero de ellos consistente en menús tipo Windows (p.e., R Commander o RKWard) y el segundo en editores de código (p.e., Tinn-R, Emacs o RStudio). Las interfaces con menús tipo Windows son las preferidas por aquellos usuarios familiarizados con softwares comerciales (p.e., Elosua, 2009; Fox, 2005; López, 2013). El paquete R Commander desarrollado por Fox (2005) se presentó como la transición más natural entre los softwares de análisis estadísticos comerciales y R (Elosua, 2009). Su desarrollo perseguía poder proporcionar una plataforma fácil de usar para cursos básicos de estadística, intentar reducir la posibilidad de cometer errores no sensatos (p.e., calcular la media de una variable de tipo nominal) y, por último, para hacer visible la relación establecida entre lo que se selecciona en los menús y el código de R que habría que haber usado en la consola de comandos. Fox (2005) señala que este tipo de interfaces gráficas tienen aspectos positivos y negativos para los usuarios. Como positivo apunta a que no es necesario recordar los comandos y argumentos de las funciones, lo que reduce la probabilidad de cometer errores en la escritura de la sintaxis, y que es un entorno cómodo para el trabajo de usuarios noveles, infrecuentes o causales. Como aspectos negativos destaca el hecho de que se hace tedioso y difícil el reproducir ciertos análisis estadísticos por implicar la navegación a través de múltiples cuadros de diálogo y, el hecho de tener que incorporar innumerables análisis en la interfaz que pueden terminar convirtiéndola en un laberinto difícil de manejar (Fox, 2005). Una de las propuestas más atractivas que se han hecho en los últimos tiempos destinadas a facilitar la interacción del usuario con código base de R la podemos encontrar en el programa JASP (Love et al., 2015). JASP (<https://jasp-stats.org>) aspira a enriquecer la experiencia del usuario con una interfaz gráfica rápida, ágil y fácil de manejar. Por su parte, interfaces como la de RStudio (www.rstudio.com), que es un *Entorno Integrado o Interactivo de Desarrollo*, esto es, una aplicación que proporciona facilidades y herramientas a aquellos usuarios dedicados a la programación informática; es preferida por aquellos usuarios con conocimientos más avanzados de programación en R o que quieren aventurarse a intentar ser productivos con el código (López, 2013).

Por último, queremos destacar que la recurrencia a las fuentes de ayuda para trabajar con R es prácticamente inevitable independientemente del nivel de competencia que tengamos con su uso. R consta de una serie de manuales en PDF que están a disposición del usuario desde la propia consola. Además, proporciona otras fuentes de ayuda diferentes (en formato html, preguntas frecuentes, etc.). Otro modo de obtener ayuda, en este caso sobre paquetes específicos, es accediendo directamente al directorio en el que se encuentra alojado el paquete en la web o en el directorio local de instalación. El uso de buscadores tradicionales de la web pueden resultar de utilidad, el Quick-R (www.statmethods.net) o la propia página de la CRAN (López, 2013). Existen manuales tanto en inglés (p.e., Field, Miles y Field, 2012) como en español (p.e., Arriaza y cols., 2008; Elosua, 2011; Elosua y Etxeberria, 2012; López, 2012, 2013) para facilitar la interacción con este softwa-

re. Por tanto, además de estar arropados por una comunidad de desarrolladores y usuarios del software, los investigadores y profesionales que se decanten por el uso de R como herramienta de análisis estadístico de datos dispondrán de un voluminoso conjunto de recursos para hacer más satisfactoria su interacción con el software. En la Tabla 1 aparece un breve listado de recursos, accesibles libremente en la web, que pueden ser útiles para hacer más amigable la interacción con R.

UTILIDADES PARA LA EVALUACIÓN PSICOLÓGICA

En esta sección vamos presentar muy someramente algunos de los paquetes de R que cobran protagonismo en el ámbito de la evaluación psicológica y, principalmente, se comentarán algunos de los más relacionados con la tecnología psicométrica. Sería inviable tratar de abarcarlos todos y no es la pretensión de este artículo. En el momento en que se están escribiendo estas palabras existen 6695 paquetes disponibles para su descarga y es fácil que cuando este texto sea leído esta cantidad se haya incrementado en algún número sustancial dado el crecimiento cuasi-exponencial que ha experimentado este valor en los últimos años (Elosua y Etxeberria, 2012). Aunque sólo una proporción de los paquetes están directamente relacionados con la psicología, también es cierto que existen muchos paquetes que pueden ser objeto de uso por parte de nuestra disciplina en momentos puntuales pese a que hayan sido desarrollados en áreas tan variopintas como el reconocimiento de patrones topográficos. En cualquier caso, la CRAN contiene un listado de los paquetes disponibles junto con una documentación que especifica cómo pueden usarse de modo eficiente.

En el año 2007 Patrick Mair y Reinhold Hatzinger inauguraron la *Psychometric Task View* con el ánimo de aglutinar todas las iniciativas y progresos que se habían hecho hasta el momento en el ámbito de la psicometría (Mair y Hatzinger, 2007a). Ese mismo año apareció un volumen especial en la *Journal of Statistical Software* donde se presentaban formalmente algunos de los paquetes más relevantes para el análisis de datos psicométrico con R (de Leeuw y Mair, 2007). Agruparon los paquetes (127 en la actualidad) en cinco grandes grupos (más uno de miscelánea) orientados al modelado psicométrico bajo la Teoría de Respuesta a los Ítems, Análisis de Correspondencias, desarrollo de Modelos de Ecuaciones Estructurales, Escalamiento Multidimensional, y Teoría Clásica de Tests.

La mayor parte de los paquetes compilados por la *Task View* se concentran en el grupo destinado al modelado estadístico bajo la Teoría de Respuesta al Ítem (TRI). Este hecho podríamos encuadrarlo en el in-

tento que se está haciendo en los últimos años por avanzar en las ventajas que implica esta teoría de la medición psicológica en comparación con la Teoría Clásica de Test (Muñiz, 1997; 2010). Si tuviésemos que destacar alguno de los paquetes que se incluyen en este grupo, podríamos hacerlo con el *ltm* (Rizopoulos, 2006) ya que, además de incluir funcionalidades para desarrollar modelos TRI dicotómicos y polítómicos (de uno, dos o tres parámetros), también contiene funciones que han sido destacadas en contextos de aplicación diferentes a la TRI (Falissard, 2012). También convendría destacar los paquetes *eRm* (Mair y Hatzinger, 2007b) y *plRasch* (Anderson, Li, y Vermunt, 2007).

Por su parte, podríamos destacar cuatro paquetes esenciales para trabajar bajo la óptica de la Teoría Clásica de Test (Mair y Hatzinger, 2007a): *psychometric*, *psy*, *psych* y *MiscPsycho*. Estos cuatro paquetes aglutinan funciones básicas y avanzadas destinadas al análisis de ítems, al estudio de la validez y fiabilidad de tests así como funciones útiles frente al desarrollo de escalas al amparo de la filosofía clásica.

Otro grupo numeroso de paquetes es aquel destinado al análisis factorial, al análisis de componentes principales y al desarrollo de modelos de ecuaciones estructurales. El análisis factorial podría considerarse como una técnica de análisis estadística que, pese a surgir en el seno de la psicología, muestra su utilidad en áreas de conocimiento tan dispares como la biología o la economía (Ferrando y Anguiano-Carrasco, 2010). El análisis factorial clásico y el análisis de componentes principales vienen implementados en el paquete *stats* que está contenido en la distribución base de R. Aunque la versatilidad de las funciones *factanal()* y *princomp()* del paquete *stats* permite ejecutar análisis factoriales y de componentes principales básicos, otros paquetes han sido desarrollados para complementar y optimizar la experiencia del usuario de R en este sentido. Los Modelos de Ecuaciones Estructurales se perfilan hoy en día como una especie de evolución del análisis factorial y técnicas asociadas aunque implican una mayor complejidad desde el punto de vista de la estimación estadística (Ruiz, Pardo, y San Martín, 2010). En R existen algunos paquetes que han sido destinados al desarrollo de este tipo de modelos con una funcionalidad semejante a los paquetes comerciales habidos hasta la fecha. Por ejemplo, el paquete *sem* (Fox, 2006) y el paquete *lavaan* (Rosseel, 2012) son dos herramientas que gozan de gran popularidad para el desarrollo de modelos de ecuaciones estructurales que permiten la estimación de modelos factoriales de diversa índole al mismo tiempo que permiten la representación gráfica de los mismos. Paquetes, entre otros, como *polycor*, para la estimación de modelos que implican correlaciones po-

TABLA 1
ALGUNOS RECURSOS LIBRES EN LA WEB QUE PUEDEN SER DE UTILIDAD PARA
HACER MÁS AMIGABLE LA INTERACCIÓN CON R.

Descripción del recurso	Enlace
Breve manual que cubre análisis estadísticos básicos y comunes en Psicología con R Commander (Español)	http://hdl.handle.net/10835/1658
Uno de los mejores compendios de tutoriales sobre R que aborda tanto aspectos básicos como más avanzados del análisis estadístico para Psicología (Inglés)	http://ww2.coastal.edu/kingw/statistics/R-tutorials/
Esta página contiene un conjunto de enlaces a vídeo-tutoriales introductorios sobre R (Español)	http://www.jpuga.es/Docencia/mooc-r.html
Curso interactivo sobre el funcionamiento de la consola de R (Inglés)	http://tryr.codeschool.com/
Manual básico de psicometría con R Commander (Español)	https://web-argitalpena.adm.ehu.es/pasa_pdf.asp?File=UWLGPS5661

licóricas, **systemfit**, para alternativas estructurales no lineales, **pIs**, para la estimación mínimo cuadrática parcial permiten una gran flexibilidad y versatilidad en el espectro de técnicas estadísticas aplicables a datos generados en el ámbito de estudio de la psicología susceptibles de ser tratados bajo la perspectiva estructural.

Para terminar, podríamos destacar que diferentes procedimientos destinados a estimar y testar modelos de escalamiento multidimensional pueden encontrarse en paquetes como **MASS**, **MLDS**, **vegan**, **labdsv**, **ecodist** o **ade4**.

CONCLUSIONES

Como se ha tratado de poner de manifiesto, R es una herramienta considerablemente versátil y flexible que permite llevar a cabo análisis de datos estadísticos que son difíciles de conseguir utilizando cualquier otro software disponible en la actualidad. Además, la gran cantidad de paquetes que se asocian a R facilita que pueda utilizarse como software único, lo que evita el tener que cambiar de un programa a otro cuando aparecen necesidades específicas de cómputo. Por ello, la psicología podría beneficiarse de esta potencialidad en todos sus ámbitos de conocimiento. No obstante, también es cierto que requiere cierta cantidad de habilidad, que no se consigue rápidamente, para interaccionar eficientemente con el programa. En cualquier caso, cada vez son más numerosos los intentos de desarrollar interfaces gráficas de usuario que sean más amigables para el usuario general con el objetivo de que éste no se aterre al enfrentarse a la consola de comandos. No hay que olvidar, como venimos señalando desde el principio, que *la potencia sin control no sirve de nada* y que el análisis estadístico de datos solo proporciona un espectro de técnicas a aplicar que han de ser amparadas por sólidas teorías científicas.

Cohen, (1990) señaló que la inferencia estadística, en conjunción con el juicio informado de las personas de ciencia, es una herramienta muy útil. Sin embargo, también considera que la estadística no es la herramienta más importante de la ciencia sino que, más bien es el resultado de lo que la precede (teoría, planificación, hipótesis, etc.). Es por ello que tenemos que volver a recalcar la importancia de la planificación de la investigación como parte esencial del método científico tal como señalan otros muchos autores (p.e., Cohen, 1990; Munafò et al., 2014) para que obtengamos unos resultados que realmente sean de utilidad tanto para la comunidad científica como para la sociedad en general.

En este sentido, es necesario un cambio cultural en la comunidad científica (p.e., Morrison, 2012; Munafò et al., 2014; Weissgerber et al., 2015) que nos permita obtener trabajos de calidad y fieles a una ética profesional adaptada a los tiempos que corren, además de una modificación en las políticas editoriales de las revistas científicas (Weissgerber et al., 2015). Este cambio cultural en parte se verá beneficiado por un entrenamiento y formación de los investigadores (Weissgerber et al., 2015), tal como señalan Leek y Peng (2015) cuando señalan que la "educación es el comienzo". En cuanto a los cambios en las políticas editoriales, al menos en psicología, tenemos el ejemplo, en cierto modo extremo, de la revista *Basic and Applied Social Psychology* que prohibió el envío de trabajos en los que se hiciera uso del p valor (Trafimow, 2014 y Trafimow y Marks, 2015).

Igualmente, y para aportar una mayor calidad a los resultados científicos que se presenten, algunos autores apuntan a que se enfaticen los datos exploratorios, es decir, que se haga un mayor uso de la estadística descriptiva (p.e., Cohen, 1990, 1994; Leek y Peng, 2015; Trafimow

y Marks, 2015; Weissgerber et al., 2015). También sería interesante aportar más resultados gráficos (Cohen, 1994), información relativa al tamaño del efecto (p.e., Cohen, 1994), intervalos de confianza o el uso de la estadística bayesiana (p.e. Bakan, 1966; Cohen, 1992, 1994; Haller y Krauss, 2002; Puga, Krzywinski y Altman, 2015a, 2015b; Trafimow y Marks, 2015). Otro aspecto importante señalado por algunos autores es realizar más replicaciones de aquellos estudios que encuentren resultados que puedan ser susceptibles de ser considerados débiles desde un punto de vista metodológico (p.e., Hasley et al. 2015; Huber et al., 2015; Munafò et al. 2014; Rosnow y Rosenthal, 1989). Además, Ioannidis (2005) apunta a que no se persiga la obtención de resultados significativos y que se tenga en cuenta la información previa existente al respecto del tema a investigar. Esto último redundaría en lo interesante del uso de la estadística bayesiana, cuyos fundamentos se sustentan en la información previa existente.

La credibilidad de la ciencia psicológica está en nuestras manos y el análisis estadístico de datos es un aliado esencial en el que confiar para optimizar el devenir del progreso científico. R se presenta en la actualidad como la lengua franca del análisis estadístico y de la generación de gráficos en un amplio espectro de áreas de conocimiento. Es cuestión de tiempo que comprobemos si R finalmente se convierte en algo que podría considerarse como patrimonio de la humanidad (López, 2012) o que simplemente su apogeo es fruto de una moda pasajera.

AGRADECIMIENTOS

Este trabajo ha sido desarrollado al amparo del Ministerio de Economía y Competitividad (proyecto PSI2014-53427-P).

REFERENCIAS

- Arriaza, A. J., Fernández, F., López, M. A., Muñoz, M., Pérez, S., y Sánchez, A. (2008). *Estadística básica con R y R-Commander*. Cádiz: Servicio de Publicaciones de la Universidad de Cádiz.
- Anderson, C. J., Li, Z., y Vermunt, J. K. (2007). Estimation of models in a Rasch family for polytomous items and multiple latent variables. *Journal of Statistical Software*, 20, 1-36.
- Bakan, D. (1966). The test of significance in psychological research. *Psychological Bulletin*, 66, 423-437. doi: 10.1037/h0020412
- Bond, M. (2009, Octubre). Decision-making: risk school. *Nature*, 461, 1189-1192. doi: 10.1038/4611189a
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., y Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Neuroscience*, 14, 365-176. doi: 10.1038/nrn3475
- Carleos, C. y Corral, N. (2013). Paquetes estadísticos con licencia libre (I). *Revista Electrónica de Metodología Aplicada*, 18, 12-37.
- Cohen, J. (1990). Things I have learned (so far). *American Psychologist*, 45, 1304-1312. doi: 10.1037/0003-066X.45.12.1304
- Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, 1(3), 98-101. doi: 10.1111/1467-8721.ep10768783
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 997-1003. doi: 10.1037/0003-066X.50.12.1103
- Cressey, D. (2015, Abril). UK funders demand strong statistics for animal studies. *Nature*, 520, 271-272. doi:10.1038/520271a
- Dar, R., Serlin, R. C., y Omer, H. (1994). Misuse of statistical tests in three decades of psychotherapy research. *Journal of Consulting and*

- Clinical Psychology*, 62, 75-82. doi: 10.1037/0022-006X.62.1.75 de Leeuw, J. y Mair, P. (2007). An introduction to the special volumen on "psychometrics in R". *Journal of Statistical Software*, 20, 1-5.
- Nature Editorial (2015, Abril). Numbers matter. *Nature*, 520, 263-264. doi:10.1038/520263b
- Elosua, P. (2009). ¿Existe vida más allá de SPSS? Descubre R. *Psicothema*, 21, 652-655.
- Elosua, P. (2011). *Introducción al entorno R*. Bilbao: Euskal Herriko Unibertsitateko Argitalpen Zerbitzua / Servicio Editorial de la Universidad del País Vasco.
- Elosua, P. y Etxeberria, J. (2012). *R Commander. Gestión y análisis de datos*. Madrid: La Muralla.
- Falissard, B. (2012). *Analysis of questionnaire data with R*. Boca Raton, FL: Taylor & Francis Group.
- Ferrando, P. J. y Anguiano-Carrasco, C. (2010). Factor analysis as a research technique in psychology. *Papeles del Psicólogo*, 31, 18-33.
- Field, A., Miles, J., y Field, Z. (2012). *Discovering statistics using R*. Londres: SAGE.
- Fox, J. (2005). The R commander. A basic-statistics graphical user interface to R. *Journal of Statistical Software*, 14, 1-42.
- Fox, J. (2006). Structural equation modeling with the sem package in R. *Structural Equation Modeling*, 13, 465-486.
- Fox, J. (2009). Aspects of the social organization and trajectory of the R project. *The R Journal*, 1/2, 5-13.
- Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics*, 33, 587-606. doi: 10.1016/j.socec.2004.09.033
- Gigerenzer, G. (1998). We need statistical thinking, not statistical rituals. *Behavioral and Brain Sciences*, 21, 199-200. doi: 10.1017/S0140525X98281167
- Haller, H. y Krauss, S. (2002). Misinterpretations of significance: A problem students share with their teachers? *Methods of Psychological Research Online*, 7, 1-20.
- Hasley, L. G., Curran-Everett, D., Vowler, S. I., y Drummond, G. B. (2015). The fickle *P* value generates irreproducible results. *Nature Methods*, 12, 179-185. doi: 10.1038/nmeth.3288
- Hernández, A., Tomás, I., Ferreres, A., y Lloret, S. (2015). Tercera evaluación de tests editados en España. *Papeles del Psicólogo*, 36, 1-8.
- Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., ... y Morgan, M. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods*, 12, 115-121. doi: 10.1038/nmeth.3252
- Ihaka, R. y Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5, 299-314. doi: 10.1080/10618600.1996.10474713
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLOS Medicine*, 2, 696-701. doi: 10.1371/journal.pmed.0020124
- Leek, J. T. y Peng, R. D. (2015, Abril). P values are just the tip of the iceberg. *Nature*, 520, 612. doi: 10.1038/520612a
- López, J. (2012). *Introducción al análisis de datos con R y R Commander en psicología y educación*. Almería: Editorial de la Universidad de Almería. Disponible en <http://hdl.handle.net/10835/1658>.
- López, J. (2013). *Primeros pasos analizando datos estadísticos con R. Un texto para acompañarte*. Murcia: Universidad Católica San Antonio.
- Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., Ly, A., Gronau, Q. F., Smira, M., Epskamp, S., Matzke, D., Wild, A., Rouder, J. N., Morey, R. D. y Wagenmakers, E.-J. (2015). JASP (Version 0.7)[Programa Informático].
- Mair, P. y Hatzinger, R. (2007a). Psychometrics task view. *R News*, 7, 38-40.
- Mair, P. y Hatzinger, R. (2007b). Extended rasch modeling: The eRm package for the application of IRT models in R. *Journal of Statistical Software*, 20, 1-20.
- Morrison, R. P. (2012). Reforming science: methodological and cultural reforms. *Infection and Immunity*, 80, 891-896. doi: 10.1128/IAI.06183-11
- Munafò, M., Noble, S., Browne, W. J., Brunner, D., Button, K., Ferreira, J., ... y Blumenstein, R. (2014). Scientific rigor and the art of motorcycle maintenance. *Nature Biotechnology*, 32, 871-873. doi:10.1038/nbt.3004
- Muñiz, J. (1997). *Introducción a la Teoría de Respuesta a los Ítems*. Madrid: Pirámide.
- Muñiz, J. (2010). La teoría de los tests: Teoría Clásica y Teoría de Respuesta a los Ítems. *Papeles del Psicólogo*, 31, 57-66.
- Nieuwenhuis, S., Forstmann, B. U., y Wagenmakers, E. (2011). Erroneous analyses of interactions in neuroscience: a problem of significance. *Nature Neuroscience*, 14, 1105-1107. doi: 10.1038/nn.2886
- Puga, J. L., Krzywinski, M., y Altman, N. (2015a). Points of Significance: Bayes' theorem. *Nature Methods*, 12, 277-278. doi:10.1038/nmeth.3335
- Puga, J. L., Krzywinski, M., y Altman, N. (2015b). Points of Significance: Bayesian statistics. *Nature Methods*, 12, 377-378. doi: 10.1038/nmeth.3368
- R Development Core Team. (2011). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rizopoulos, D. (2006). ltm: An R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software*, 17, 1-25.
- Rosnow, R. L. y Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist*, 44, 1276-1284. doi: 10.1037/0003-066X.44.10.1276
- Rosseeel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48, 1-36.
- Ruiz, M. A., Pardo, A., y San Martín, R. (2010). Structural equation models. *Papeles del Psicólogo*, 31, 34-45.
- Seldmeier, P. y Gigerenzer, G. (2001). Teaching bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General*, 130, 380-400. doi: 10.1037/0096-3445.130.3.380
- Tippmann, S. (2015, Enero). Programming tools: Adventures with R. *Nature*, 109-110. doi:10.1038/517109a
- Trafimow, D. (2014). Editorial. *Basic and Applied Social Psychology*, 36, 1-2. doi: 10.1080/01973533.2014.865505
- Trafimow, D. y Marks, M. (2015). Editorial. *Basic and Applied Social Psychology*, 37, 1-2. doi: 10.1080/01973533.2015.1012991
- Vaux, D. (2012, Diciembre). Know when your numbers are significant. *Nature*, 492, 180-181. doi:10.1038/492180a
- Weissgerber, T. L., Milic, N. M., Winham, S. J., y Garovic, V. D. (2015). Beyond bar and line graphs: Time for a new data presentation paradigm. *PLOS Biology*, 13, 1-10. doi: 10.1371/journal.pbio.1002128