

Cuestionarios psicológicos e investigación en Internet: Una revisión de la literatura

Francisco José Eiroá Orosa*, Irene Fernández Pinto y Pau Pérez Sales

Grupo de Acción Comunitaria. Centro de Recursos en Salud Mental y Derechos Humanos (Madrid, España)

Resumen: La creciente importancia de Internet ha cambiado el mundo de la investigación psicológica, iniciando un nuevo campo de exploración y abriendo nuevas perspectivas de muestreo. El presente artículo revisa aspectos metodológicos ligados al uso de cuestionarios en línea con fines de investigación. Existen innegables ventajas: reducción de costes, disminución de los efectos de deseabilidad social y acceso a grandes grupos de población. Pero también existen dificultades derivadas de potenciales sesgos muestrales y del modo de implementación, que podrían ser parcialmente compensados. En conjunto, las posibilidades de la implementación de cuestionarios son tan grandes como la propia red, pero solo a través del conocimiento de sus dinámicas, principios metodológicos y éticos.

Palabras clave: Cuestionarios en línea; muestreo; Internet; metodología; proyecto VIVO.

Title: Psychological questionnaires and research on the Internet: a literature review.

Abstract: The growing importance of Internet has changed the world of psychological research, initiating a new research field and opening new sampling perspectives. The present article reviews methodological aspects related to online questionnaires for research purposes. There are some clear advantages: cost reduction, decrease of social desirability effects and access to big population groups. But there are also difficulties derived from potential sampling bias and implementation methods which could be partially compensated. As a whole, the possibilities of questionnaire implementation are as big as the net, but only through the knowledge of its dynamics, methodological and ethical principles.

Key words: On line questionnaires; sampling; Internet; methodology; VIVO project.

Introducción

La introducción de medios telemáticos en el mundo de la investigación psicológica ha supuesto un impulso a las posibilidades de los investigadores. A pesar de las dificultades metodológicas que supone, la investigación en línea ha demostrado ser más eficiente en términos de rapidez y economía que los métodos tradicionales. Sin embargo, existen también multitud de aspectos que deben ser sometidos a mejora.

El inicio de la utilización de métodos computerizados en la investigación psicológica comenzó a final de los años 60 (Elwood, 1969; Miller, 1968) aunque se empezó a generalizar en los 70. Las mejoras introducidas incluían una manera controlada y estandarizada de presentar los estímulos, posibilidad de ramificar los ítems (*item-branching*), inmediatez en la entrada de datos, eliminación de respuestas omitidas, eliminación de errores de transcripción, disminución del sesgo del investigador y demás ventajas asociadas a la introducción directa de los datos por parte del sujeto de experimentación a través de una interfaz mecánica (Musch y Reips, 2000).

Pero lo que realmente ha revolucionado las posibilidades de los investigadores, ha sido la generalización del uso de Internet en los años 90. El impacto en nuestra disciplina es doble; por un lado, se abre un campo de investigación dedicado al nuevo tipo de redes humanas que se dan en el ciberespacio, y por otro, los investigadores tienen a su disposición un medio masivo para encontrar participantes en sus estudios (Skitka y Sargis, 2005).

Internet se perfila como un lugar donde recopilar información válida tanto con fines de investigación, como de inte-

rés práctico, incluso con fines clínicos (Chow, Treasure y Gallivan, 2007).

La investigación a través de internet no es aún considerada una disciplina por las propias asociaciones de investigadores en línea (Baym, 2005). Sin embargo, la clara utilidad de este medio en la construcción y administración de cuestionarios psicológicos, nos ha llevado a la elaboración de una revisión sobre sus características más destacadas.

Para la realización de este artículo hemos revisado la literatura disponible en las principales bases de datos incluyendo Academic Search Premier, MEDLINE, Psychological Abstracts, PsycARTICLES, PsycINFO y Google académico, además de manuales especializados, recopilando los fundamentos metodológicos para ofrecer una visión general de los procedimientos a utilizar en la implementación de un cuestionario psicológico a través de Internet.

Ventajas e inconvenientes

Multitud de manuales y artículos publicados en los últimos años recopilan las principales características, tanto positivas como negativas, referentes a la investigación en Internet. Antes de plantearse utilizar Internet como medio para realizar una investigación se debe plantear su idoneidad atendiendo a estas características.

Best y Krueger (2004) destacan las ventajas e inconvenientes de la recolección de datos a través de Internet. Las ventajas que describen son:

- *Nuevas posibilidades de muestreo.* Acceso a poblaciones de usuarios de ciertos servicios, miembros de organizaciones o personas que comparten intereses.
- *Flexibilidad en la administración.* Tanto en diseño, posibilidades multimedia, accesibilidad, aleatorización de las preguntas, etc.
- *Alternativas de recopilación.* Nuevas maneras de inducir a los participantes como prenotificaciones, recordatorios, incentivos. Posibilidad de enviar los datos al final del instrumen-

* Dirección para correspondencia [Correspondence address]: Francisco José Eiroá Orosa. Grupo de Acción Comunitaria. Centro de Recursos en Salud Mental y Derechos Humanos. C/ Escultores 3, 6ªA. 28760 Tres Cantos (Madrid, España). E-mail: fciroa@gmail.com

to o al acabar determinados ítems. Incluso métodos para autenticar los envíos y contrastar lo razonable de las respuestas.

Por otro lado, estos autores destacan las siguientes limitaciones potenciales:

- *Pobre accesibilidad* a determinados estratos económicos o zonas geográficas que no disponen de conexión.
- *Variabilidad entre los ordenadores* de los usuarios que nos puede hacer dudar de la uniformidad en la visualización del instrumento o el modo de aplicación.
- *Riesgo de atraer casos no elegibles y respuestas no válidas*. Al no poder supervisar la administración, pueden acceder personas al instrumento no adecuadas para la investigación o los sujetos pueden ignorar las instrucciones.

Skitka y Sargis (2005) añaden la creciente *eficiencia*. Esta característica se incrementa, ya que cada vez es más sencillo el uso de herramientas para implementar los cuestionarios y son más rápidas y baratas. Sin embargo, añaden las siguientes limitaciones:

- *Diferencias entre los usuarios de Internet con los no usuarios*. Ya que los usuarios de Internet son, en general, más jóvenes, tienen más recursos y mayor nivel cultural que la media de población general.
- *Dificultades para extraer muestras probabilísticas*. Exceptuando poblaciones en las que todos los usuarios tienen acceso a un correo electrónico como universidades o empresas. Como veremos más adelante, los cuestionarios en línea son adecuados para el muestreo no probabilístico, como el de conveniencia.
- *Menor probabilidad de participar*. La sobrecarga del medio ha causado que los posibles participantes estén sobresaturados de requerimientos para participar en estudios. Existen menos posibilidades de participación que a través de medios tradicionales como el teléfono o en persona.

Schonlau, Fricker y Elliott (2002) revisan en qué medida son ciertas las ventajas de la aplicación de cuestionarios por Internet. Encuentran apoyo en la evidencia respecto a cuatro de las cuestiones más planteadas sobre esta metodología:

- *Rapidez*: Se comprueba que la aplicación de cuestionarios en la red es más rápida que por métodos tradicionales. Tanto en la preparación de los instrumentos como en su reparto y cumplimentación.
- *Eficiencia*: Entre las mejoras que introduce se encuentran la inhibición de la deshabilidad social, el acceso a subconjuntos de población y la facilidad para introducir elementos interactivos.
- *Coste*: Para ellos no está tan claro que el uso de Internet sea más barato debido al coste de técnicos cualificados que ha sido ignorado por la literatura. Por ello, la inversión en una investigación de este tipo será menor si la muestra es mayor (en cuyo caso los gastos de medios tradicionales como el correo o los operadores telefónicos sobrepasaría el elevado coste técnico). Sin embargo, la aparición en los últi-

mos años de herramientas de fácil manejo, estaría acabando con esta dificultad.

- *Facilidad*: Las encuestas dirigidas por Internet necesitan más conocimientos técnicos y mayor número de ensayos previos.

Estos autores afirman que no se dispone de suficiente experiencia aún para dar respuesta a otras tres importantes cuestiones:

- *Diseño óptimo*: Aunque parece que la mayoría de los principios aplicables al papel son también aplicables a la pantalla del ordenador, esto debe ser investigado con mayor profundidad. Como veremos en siguientes secciones el diseño es de vital importancia para maximizar la tasa de respuesta.
- *Eficacia del uso de incentivos*: Estos autores encuentran insuficiente la investigación sobre el aumento de la tasa de respuesta que pueden suponer los incentivos en los estudios en la red. A pesar de ello ofreceremos algunas reflexiones y datos más adelante.
- *Eficacia de los métodos de ponderación*: A pesar de su uso habitual existe poca evidencia derivada de estudios que hayan minimizado los posibles sesgos. También estos aspectos serán revisados posteriormente.

Tipos de estudios en la red

Skitka y Sargis (2005) dividen en tres los tipos de investigaciones que se pueden realizar utilizando como medio la red:

- *Translacionales*: Aplicación de técnicas ya usadas con métodos tradicionales. Ofrecen la ventaja de que se pueden hacer estudios comparativos para evaluar la validez de criterio. Esta modalidad es sobre todo utilizada en investigaciones sobre personalidad, en las que se replican otros estudios ya realizados en formato papel.
- *Novelas*: Estudios que utilizan nuevas técnicas no utilizadas anteriormente. En esta modalidad destacan estudios que utilizan como ventaja el anonimato de la red para conseguir muestras. Ejemplos de ello son estudios en los que se incluyen variables étnicas, ya sea utilizando a minorías como muestra (Barry, 2001) o investigando actitudes violentas contra estas (Glaser, Dixit y Green, 2002).
- *Fenomenológicos*: Investigaciones cuyo objeto de estudio es precisamente el comportamiento y la interacción humana en Internet. Este área de investigación en crecimiento, explora como el uso de las nuevas tecnologías de la información afectan al bienestar psicológico (Gross, Juvonen y Gable, 2002) y al desenvolvimiento social (Kraut et al.; 1998, 2002).

Por otro lado, señalamos que la red puede también ser utilizada como un lugar para la recopilación de narrativas personales y grupales aptos para la investigación cualitativa (Adair, Marcoux, Williams, y Reimer, 2006).

Cuestiones de diseño

El diseño del cuestionario es la primera parte del proceso de implementación en línea del instrumento. Si bien los métodos de elaboración de preguntas no difieren de los utilizados en los cuestionarios en papel, es recomendable prestar especial atención a las cuestiones relacionadas con el diseño gráfico, ya que luego serán determinantes en la tasa de respuesta al instrumento. Para diseñar un cuestionario cibernético existen multitud de aplicaciones informáticas que hacen muy sencilla tanto los aspectos de diseño e instalación, como la conexión del cuestionario con una base de datos que almacene los resultados. Sin embargo, hay multitud de aspectos a tener en cuenta para hacer un buen diseño que conserve el rigor metodológico y nos de cómo resultado una alta tasa de respuesta. Según Dillman (2007), los principios fundamentales a tener en cuenta al diseñar un cuestionario alojado en una Web son:

- Introducir el cuestionario con una pantalla de bienvenida motivante, que enfatice la facilidad para responderlo y dé instrucciones específicas para ello.
- Proveer de un código de acceso cuando éste deba ser restringido a los miembros de la muestra.
- Presentar cada pregunta en un formato convencional similar al usado en los cuestionarios autoadministrados de papel.
- Contener el uso de los colores e imágenes para preservar la consistencia figura/fondo y la legibilidad. De este modo, la navegabilidad está libre de obstáculos y las propiedades de las preguntas se mantienen.
- Evitar las diferencias visuales que resultan de las diferencias técnicas entre usuarios (uso de diferentes configuraciones de pantalla, sistemas operativos, navegadores...).
- Proveer de las instrucciones específicas para realizar cada operación informática necesaria para responder al cuestionario y dar otras informaciones cuando sean necesarias.
- Usar los desplegables con moderación, considerando las implicaciones en el modo de respuesta e indicar cada uno con un texto "haz clic aquí".
- No exigir la respuesta a cada pregunta como requisito para pasar a la siguiente, a menos que sea parte de los objetivos de la investigación que el sujeto responda a todas las preguntas.
- Construir cuestionarios donde todas las preguntas sean visibles. En el caso de que este tenga varias partes o sea importante que se responda en un cierto orden, se proveerá de botones con flechas.
- Usar gráficos que indiquen el grado en el que se va completando el cuestionario sin que esto aumente excesivamente el uso de recursos.
- Restringir el uso de estructuras de respuesta con problemas de medición en cuestionarios en papel, como las preguntas abiertas o las preguntas de respuesta múltiple.

Best y Krueger (2004) ofrecen especificaciones técnicas para que el formato del instrumento sea óptimo. De acuerdo con estos autores debemos seguir los siguientes preceptos:

- La resolución en pantalla debe ser de 600X800 ya que de momento es la más utilizada. Aunque en los próximos años se producirá un aumento del uso de la resolución de 1024X768, la utilización de una resolución menor nos asegura que todos los usuarios pueden ver el cuestionario con corrección. Algunas aplicaciones más modernas pueden incluso adaptar el formato del cuestionario al del usuario.
- Limitar el uso de colores a aquellos que pueden ser visualizados en un sistema de 8 bits (256 colores), maximizar el contraste entre ellos, ser consciente de la interpretación que se puede hacer de ellos (como el rojo y su relación con el peligro) y evitar los contrastes entre rojo y verde ya que las personas daltónicas podrían responder de manera errónea.
- Utilizar fuentes convencionales (como Times New Roman o Arial), utilizar tamaños también convencionales (10, 11 o 12), utilizar espaciados proporcionales (tanto vertical como horizontalmente), no abusar de las letras mayúsculas y moderar el uso de negritas, subrayados o itálicas.
- Podemos usar un diseño estático, donde se accede directamente a todas las preguntas; o un diseño interactivo, en el que se requiere alguna acción para pasar por las diferentes partes del cuestionario o estas son aleatorias.
- Usar preferentemente casillas para responder a las preguntas cerradas.
- Tener en cuenta la longitud del instrumento (ya sea el número de preguntas o el tiempo que se tarda en responderlo) ya que hay evidencia de que el uso de cuestionarios muy largos afecta a la tasa de respuesta.

Muestreo

Para empezar a recopilar muestra debe tenerse en cuenta a qué tipo de población se puede acceder. Como ya hemos dicho antes, existe evidencia empírica de las diferencias existentes entre la población de usuarios de Internet y la población general. En España, por ejemplo, se está produciendo una gran evolución en el perfil de los internautas; aunque siguen siendo mayoritariamente menores de 50 años, estudiantes o profesionales con alto nivel educativo (Observatorio de las Telecomunicaciones y de la Sociedad de la Información, 2005).

Para que el cuestionario llegue a los posibles encuestados, debe estar alojado en una web y estar conectado a una base de datos. Aunque también existe la posibilidad de enviar cuestionarios adjuntos en un correo electrónico, estos no se ajustan a las características que describimos, ya que los clientes de correo actuales no están preparados para enviar los resultados a una base de datos. Solo podríamos enviar un documento de texto a cumplimentar, que deberíamos procesar de manera análoga a los de lápiz y papel.

La manera en que hagamos llegar el cuestionario a los miembros de la muestra determinará el tipo de muestreo utilizado. La multitud de maneras en las que los usuarios se conectan a Internet nos obliga a utilizar diversos métodos de contacto:

- *Contacto a través de correo electrónico*: Este es el único método en el que podemos acceder a muestras aleatorias. Se asemeja más a los métodos tradicionales de listín telefónico o correo convencional, pero con la desventaja de que no existen listines de correo electrónico, que muchas personas tienen varios y que no pueden ser generados aleatoriamente (como los números de teléfono). Sin embargo, este método de muestreo es válido en poblaciones que tienen listas de correo ya generadas como instituciones educativas, compañías, asociaciones, etc. En el correo electrónico el investigador envía una presentación del estudio que se está realizando y un enlace al lugar donde está alojado el cuestionario.
- *Contacto con usuarios de listas de correo*: Se plantean dificultades similares al contacto por email ya que no existen directorios completos y exhaustivos de las listas existentes. Localizarlas implica navegar por los portales web donde es más probable que estén identificadas y organizadas. También es difícil conseguir los datos demográficos de los usuarios de las listas. Sin embargo, es un buen método para entrar en contacto con poblaciones con intereses determinados, ya que esta es la razón por la que se afilian a las listas. La manera de envío es idéntica a la del correo electrónico ya que para enviar un mensaje a una lista de correo debemos enviar un correo a su dirección que, una vez autorizado, es enviado a toda la lista.
- *Contacto a través de una Web*: En este caso las personas que llegan al instrumento están de algún modo interesadas en la temática de la investigación y es la motivación principal para cumplimentarlo. En muchos casos hay páginas de investigadores o instituciones en las que se puede acceder a los instrumentos (p.e. las páginas Authentic Happiness o el departamento de Psicología de la Universidad de Hanover). Se puede llegar a través de anuncios en otras webs o buscadores. En este caso, dado el número de páginas disponibles, será el posicionamiento en los motores de búsqueda una de las mejores maneras para que los potenciales usuarios lleguen al instrumento. En otros casos, se utilizan webs de uso común para ofrecer responder el instrumento. Un ejemplo de ello es un estudio en el que se utilizó la web de un canal de televisión musical para realizar una investigación sobre la pérdida de oído en una muestra de jóvenes (Chung, Des Roches, Meunier y Eavey; 2005). Por otro lado, el muestreo a través de webs que ofrecen servicios o información ofrecen grandes resultados. Un estudio sobre pesadillas realizado a través de una Web sobre sueño y salud obtuvo casi 24000 respuestas (Nielsen, Stenstrom y Levin, 2006).
- *Contacto con usuarios de foros*: La utilización de foros para crear redes de comunicación entre usuarios de servicios o como punto de intercambio de intereses aumenta exponencialmente en la red. La utilización de este medio para conseguir muestra, implica una gran inversión de tiempo ya que la mayoría de foros exige registro. Sin embargo, la posibilidad de interacción y la gran variedad de foros, posibilita el contacto con poblaciones específicas o, incluso, muestreos

generalizados en los que se puede controlar el tipo de población a la que se dirige.

- *Contacto con usuarios de chats*: Es un método que consume gran cantidad de tiempo, ya que el investigador debe conversar con los usuarios del chat en el que va a anunciar el cuestionario. Sin embargo, como en el caso de los foros, es un método que nos permite acceder a poblaciones con características muy específicas y difíciles de acceder de otro modo. Además las posibilidades de contacto continuo y monitorización de las salas de Chat, permiten la recolección de información de tipo etnográfico (Mathy et al., 2002).

Como hemos dicho, el acceso a muestras probabilísticas está limitado a poblaciones en las que dispongamos de una lista de correos electrónicos que puedan dirigirles al cuestionario. Sin embargo, la red es el lugar apropiado para muestras no probabilísticas. Bradley (1999) recopila cinco tipos de muestreos no probabilísticos aplicables a muestras de internautas:

- *Conveniencia*: Se elige la muestra por “conveniencia”, o en otras palabras, porque es de fácil acceso.
- *Juicio*: Se elige la muestra respecto al “juicio” del investigador. En Internet podemos juzgar que los usuarios de un servicio muy común se ajusten a la población que queremos investigar.
- *Cuota*: Es la versión no probabilística del muestreo estratificado. Después de crear los estratos se utiliza el muestreo de conveniencia o de juicio para elegir una muestra.
- *Aumento (snowballing)*: Se utiliza a los usuarios que ya han cumplimentado el cuestionario para que lo envíen a otros posibles encuestados.
- *Auto-selección*: Los usuarios que completan el cuestionario son voluntarios y son ellos mismos los que acceden al instrumento, como en el caso de las webs de investigación.
- *Verosimilitud*: Se selecciona una muestra porque se considera plausible que se asemeje a la población real.

Schonlau et al. (2002) apuntan a las ventajas del muestreo de conveniencia. A pesar de no soportar inferencia estadística, este tipo de muestreo es útil para desarrollar hipótesis en los primeros pasos de una investigación, identificar asuntos que rodean el tema de la investigación, definir las categorías de respuesta para las preguntas de respuesta múltiple y otros tipos de información no inferencial.

Por otro lado, Couper (2000) habla de formas de muestreo basadas en métodos probabilísticos. Propone, además de la posibilidad de acceso a una lista de correos electrónicos, interceptar a usuarios de páginas web de manera aleatoria o la contratación de paneles de usuarios.

A pesar de la gran controversia que genera el estudio de la comparabilidad de las muestras que se obtienen a través de la red con muestras tradicionales; estudios comparativos muestran no solo la validez de estas, sino la mayor diversidad frente a las tradicionales muestras de estudiantes (Gosling, Vazire, Srivastava y John, 2004).

Posibles sesgos y control

Como vemos la manera más sencilla de trabajar en la red es a través de muestras no probabilísticas. Birnbaum (2004) destaca el sesgo de autoselección, ya que en el muestreo no probabilístico son los participantes quienes deciden su participación. Kyprí Stephenson, y Langley (2004) destacan los sesgos derivados de las bajas tasas de respuesta. Por otro lado, existen sesgos derivados de factores más técnicos como posible incompatibilidad del instrumento con algunos navegadores y el bloqueo del correo no deseado o de las ventanas emergentes en las que se puede anunciar el cuestionario (Hartford, Carey y Mendonca, 2007).

Debido a estos sesgos, se recomienda trabajar la información proveniente de muestreos no probabilísticos en base a formas de control post-hoc, con el objetivo de inferir estos datos a poblaciones generales. Destacan los siguientes métodos:

- *Ponderación después de estratificar (poststratification weighting;* Holt y Smith 1979). Este método consiste en darle un peso relativo a las partes de la muestra dependiendo de la proporción que se sabe que existe en población general.
- *Puntuación de propensión (propensity scoring;* Cochran, 1968; Rosenbaum y Rubin 1983). La puntuación por propensión consiste en calcular la probabilidad que tiene cada sujeto de estar en la muestra, para intentar inferir resultados en población general. Cada sujeto recibe una puntuación a través de la cual se ponderarán sus respuestas. Esta puntuación se halla ajustando la muestra a un modelo probabilístico mediante regresión logística, cuyo resultado indica la probabilidad de formar parte de la muestra de Internet. Las covariables del modelo incluyen datos sociodemográficos así como indicadores sobre la accesibilidad a Internet en la población a la que se quiere generalizar los datos (Best y Krueger, 2004).

A pesar del método elegido, no hay seguridad de que las amenazas sobre la cobertura de la muestra estén solventadas. Ambos métodos hacen dos asunciones muy cuestionables (Mitofsky, 1999). En primer lugar asumen que las variables usadas para el ajuste son las únicas variables relacionadas con las variables de interés. Esto es improbable considerando que ambos métodos se basan en el ajuste a través de un grupo limitado de datos sociodemográficos. Experiencias, creencias o actitudes, por ejemplo; pueden exagerar las variables de interés. Si las variables ponderadas están viciadas, la calidad de las medidas empeorará. Además, ambos métodos asumen que los sujetos de la muestra generan opiniones de la misma manera que aquellos que no están en la muestra.

La postestratificación solo mejora la precisión de los cálculos si la relación entre las variables ponderadas y las variables que se utilizan como control es equivalente entre los que están en la muestra de Internet y los que no. De manera similar, el método de puntuación por propensión entendiéndose que el modelo multivariante que derivaría de una

muestra probabilística es equivalente a la muestra de Internet. Sin embargo, las características y modos de generar opinión de la población que podemos encontrar en Internet pueden variar de la población general (Robinson, Neustadtl y Kestnbaum, 2002).

La calidad de los datos será mayor en los casos en los que hayamos registrado una baja tasa de abandonos, se haya cumplimentado el cuestionario entero y se haya comprobado la honestidad de las respuestas, en el caso que se hayan añadido preguntas de control (Schonlau *et al.*, 2002).

Best y Krueger (2004) también atienden el problema del tamaño de la muestra. Para estos autores el tamaño muestral está en función del número deseado de casos (necesarios para validar el estudio), la incidencia estimada de población diana en el marco de muestreo (número de personas de la muestra que realmente pertenecen a la población que queremos encuestar), la extensión de métodos inválidos de contacto (como direcciones de correo o enlaces en desuso) y finalmente, la cooperación proyectada (medida en la que se espera que respondan los sujetos válidos con los que se ha contactado).

Tasa de respuesta

Una vez el proceso de muestreo ha concluido y el cuestionario ha sido implementado, se obtiene una tasa de respuesta resultado de la proporción entre los contactos establecidos y las respuestas obtenidas. Como hemos dicho antes, la tasa de respuesta en este tipo de investigaciones es baja dado el congestionamiento de requerimientos en la red. Sin embargo, la tasa y velocidad de respuesta varía mucho dependiendo del ámbito y la temática de la investigación. La mayor parte de los trabajos disponibles sobre tasas de respuesta comparadas ha sido realizada en instituciones académicas donde existe un acceso universal a Internet y los cuestionarios son anunciados a través de correos electrónicos o convencionales. Una encuesta masiva a estudiantes (Carini *et al.*, 2003), en la que respondieron casi 60000 alumnos, apunta a mínimas diferencias en la tasa de respuesta en alumnos que respondieron a través de una Web. Kaplowitz, Hadlock, y Levine (2004) comparan las tasas de respuesta de un cuestionario enviado a 20000 alumnos. Encuentran que las tasas de respuesta del correo convencional y electrónico son similares cuando fueron precedidas de una notificación por correo convencional. Las características de estas investigaciones hacen poco generalizables los resultados, aunque apuntan a la similitud entre modos de administración.

Otros estudios utilizando cuestionarios sobre población general arrojan resultados contradictorios. Bälter, Bälter, Fondell, y Lagerros (2005), encontraron tasas de respuesta mayores en la primera mitad de un cuestionario administrado en papel mientras que la segunda mitad obtenía mayor tasa en su versión online.

Un metaanálisis realizado por Cook, Heath y Thompson (2000), indica que los factores más importantes para fomentar la tasa de respuesta en cuestionarios en red enviados a

través de correo son: un elevado número de contactos, personalizar los correos y hacer contactos preliminares.

El uso de incentivos puede suponer un aumento de la tasa de respuesta. Bosnjak y Tuten (2003) comprueban la eficacia de los incentivos en la respuesta a un cuestionario siempre que estos se paguen con antelación. En este estudio, el pago de incentivos con posterioridad no tuvo ventajas. Esto confirma los estudios que solo atribuían aumento de la tasa de respuesta tras un ofrecimiento inicial de incentivos en encuestas a través de correo convencional (Church, 1993).

Multitud de sujetos dejan de responder el cuestionario una vez ya han accedido. Aunque no se encuentra una respuesta clara a este problema, se apunta a cuestiones de diseño. Cuando este es excesivamente complejo, extravagante o incómodo (con barras de navegación que no permiten ver con claridad), con instrucciones confusas y pocas instrucciones de cumplimentación; se pueden producir más abandonos. Bosnjak y Tuten (2001) han intentado perfilar los patrones de conducta de las personas que no acaban de cumplimentar el cuestionario. Para ellos existen personas que cumplimentan el cuestionario en su totalidad, personas que solo cumplimentan partes y “merodeadores” (personas que antes de decidir su participación leen las preguntas). En base a este modelo podemos, por ejemplo, identificar partes del cuestionario especialmente conflictivas si se detectan muchos usuarios que dejan sin responder un bloque consistente de preguntas.

El uso de retroalimentación a través de un perfil suministrado al completar el cuestionario, es una buena manera de que los sujetos cumplimenten el instrumento hasta el final. Este método está siendo utilizado en el proyecto VIVO, un estudio sobre respuestas traumáticas realizado por los autores de este artículo a través de la web <http://www.psicosocial.info>, con excelentes resultados.

Validez

Según Krantz y Dalal (2000) hay dos formas principales de establecer la validez del método de investigación en Internet. En primer lugar, compararlo con otro estudio equivalente desarrollado con metodología tradicional. En el caso de los cuestionarios en Internet esto sería compararlos con los mismos cuestionarios administrados en papel. El otro es examinar teóricamente si los resultados siguen una determinada predicción de tendencias. El primer método es un tipo de validez convergente mientras que el segundo se define

como validez de constructo. También se ha investigado como diferentes formatos de presentación en línea pueden cambiar la estructura latente de las respuestas no encontrando diferencias significativas, como era el caso de la tasa de respuesta (Cook y Thompson, 2001).

Comparabilidad con estudios en papel

Este es el método más utilizado para comprobar la validez de los estudios realizados a través de la implementación de cuestionarios en línea, por ello le prestaremos especial atención.

Dillman (2007) indica mecanismos normativos y cognitivos que contribuyen a la diferencia entre ambos modos de administración. Los mecanismos normativos son: deshabilitación social, aquiescencia, orden de las preguntas y primacía de la opción de respuesta. Los mecanismos cognitivos hacen referencia a si los sujetos reciben el mismo estímulo según el método de administración y está en relación con la existencia de diferentes modos de procesamiento visual y auditivo o preferencias neurológicas.

Leung y Kember (2005) señalan como método aceptado para saber si los sujetos están respondiendo a las mismas preguntas, es decir, al mismo constructo, contrastar las estructuras de los factores entre los grupos o medidas usando modelos de ecuaciones estructurales. Si los modelos resultantes son equivalentes, puede inferirse que los sujetos evalúan los mismos constructos. Estos autores no hallan, tanto en su propia investigación como en otras revisadas, diferencia alguna entre los constructos medidos según la metodología de registro utilizada.

Por otro lado, la mayoría de investigaciones se han centrado en la posibilidad de que las puntuaciones registradas difieran. Por ejemplo, en psicología de la personalidad se han encontrado diferencias al pasar tests de medida de los cinco grandes factores de personalidad (Oswald, Carr y Schimdt, 2001). Pueden atribuirse esas diferencias a que ambas condiciones experimentales son administradas a diferentes muestras que, a pesar de ser grandes nunca se pueden equiparar. En una investigación al efecto, demuestran la equivalencia del papel e Internet utilizando metodología de validación test-retest (Salgado y Moscoso, 2003).

Sin embargo, en el resto de bibliografía consultada, los resultados indicaron la equivalencia entre métodos a través de análisis estadístico (ver Tabla 1).

Tabla 1: revisión de artículos sobre resultados comparados papel/Web.

Referencia	Resultados y conclusiones
Cronk, B. C., and West, J. L. (2002).	Se administró un cuestionario sobre visión de la moralidad en cuatro condiciones diferentes (en casa). No se encontraron diferencias significativas en las condiciones experimentales. Se encontraron diferencias en la tasa de respuesta según el método.
Epstein, Klinkenberg, Wiley, y McKinley (2001).	Se utilizaron dos muestras equivalentes para comparar la implementación en papel e Internet. No se encontraron diferencias en la muestra total aunque sí se encontraron diferencias en la muestra de Internet.
Fortson BL, Scotti JR, Del Ben KS y Chen YC. (2006).	Se utilizó un diseño factorial 2 x 2 x 2 de medidas repetidas siendo las variables dependientes las puntuaciones en los cuestionarios.

Hallfors D.; Khatapoush S.; Kadushin C.; Watson K.; Saxe L.(2000).	trauma. Se demostró que la estructura psicométrica apenas variaba. Se pretendía saber si la administración de un cuestionario sobre consumo de drogas por que utilizando métodos tradicionales, pero el nivel de consumo declarado no va
Vispoel, W. P. (2000).	Los resultados apoyan la comparabilidad de los resultados en papel o internet del . rencia por el uso de la versión computerizada.
Vispoel, W. P., Boo, J., and Bliceler, T. (2001).	Los resultados muestran pocas diferencias en la estructura psicométrica de la escala su aplicación vía web o papel.

Aspectos éticos

La posibilidad de utilizar la red para obtener muestras es un claro empuje a las posibilidades de los investigadores. Sin embargo, se debe utilizar teniendo en cuenta códigos éticos, y respetando a los propios usuarios susceptibles de contestar, ya que un foro o chat es un lugar donde muchas personas comparten y obtienen refuerzo social.

Azar (2000) revisa las comisiones éticas sobre investigación en línea creadas por El Consejo Científico de la Asociación Americana de Psicología y la Asociación Americana para el Avance de la Ciencia. Destacan aspectos como la falta de protocolos estandarizados de consentimiento de participación, el peligro de la violación de la confidencialidad (incluso por motivos ajenos al investigador, dada la carencia de seguridad absoluta en la red) o la participación de niños.

Recientemente el Concilio Americano de Organizaciones de Investigación y Encuestas ha elaborado un código ético y de estándares para la investigación a través de cuestionarios (Council of American Survey Research Organizations, 2004). En él se enfatiza el consejo de no utilizar correo no deseado para conseguir muestra y la prohibición de utilizar correos electrónicos a los que no se puede responder (método muy utilizado para la difusión de correos publicitarios no deseados).

Conclusiones y recomendaciones

Referencias

- Adair, C. E., Marcoux, G., Williams, A., y Reimer, M. (2006). The Internet as a source of data to support the development of a quality-of-life measure for eating disorders. *Qualitative Health Research*, 16, 538-546.
- Azar, B. (2000). Online Experiments: Ethically Fair or Foul? *Monitor on Psychology*, 31(4), 50-51.
- Bälter, K. A., Bälter, O., Fondell, E. y Lagerros, Y. T. (2005). Web-based and mailed questionnaires: A comparison of response rates and compliance. *Epidemiology*, 16, 577-579.
- Barry, D. T. (2001). Assessing culture via the Internet: Methods and techniques for psychological research. *CyberPsychology & Behavior*, 4, 17-21.
- Baym, N. K. (2005). Introduction: Internet research as it isn't, is, could be, and should be. *The Information Society*, 21, 229-239.
- Best, S.J. y Krueger, B.S. (2004). *Internet Data Collection*. London, Sage.
- Birnbaum, M. H. (2004). Human research and data collection via the Internet. *Annual Review of Psychology*, 55, 803-832.
- Bosnjak, M. & Tuten, T.L. (2003). Prepaid and Promised Incentives in Web Surveys - An Experiment. *Social Science Computer Review*, 21(2), 208-217.
- Bosnjak, M.M. y Tuten, T.L. (2001). Classifying Response Behaviors in Web-based Surveys. *Journal of Computer-Mediated Communication*, 6, 3.
- Bradley, N. (1999) Sampling for Internet surveys. An examination of respondent selection for Internet research. *Journal of the Market Research Society*, 41(4), 387-395.
- Carini, R. M., Hayek, J. H., Kuh, G. D., Kennedy, J. M. y Ouimet, J. A. (2003). College student responses to web and paper surveys: Does mode matter? *Research in Higher Education*, 44, 1-19.
- Chow, Y. W., Treasure, T., & Gallivan, S. (2007). Can the Internet be used to collect clinical data for research? *Health Informatics Journal*, 10, 77-82.
- Chung JH, Des Roches CM, Meunier J, Eavey RD. (2005). Evaluation of noise-induced hearing loss in young people using a web-based survey technique. *Pediatrics*, 115, 861-867.
- Church, A.H. (1993). Estimating the effect of incentives on mail survey response rates: A meta-analysis. *The Public Opinion Quarterly*, 57, 62-79.
- Cochran, W. G. (1968). The Effectiveness of Adjustment by Subclassification in Removing Bias in Observational Studies. *Biometrics*, 24, 205-213.
- Cook, C. y Thompson, R.L. (2001). Do Different Response Formats Change the Latent Structure of Responses? an Empirical Investigation Using Taxometric Analysis. *Educational and Psychological Measurement*, 61(1), 23-44.

- Cook, C., Heath, F. y Thompson, R.L. (2000). A Meta-Analysis of Response Rates in Web- or Internet-Based Surveys. *Educational and Psychological Measurement*, 60(6), 821-836.
- Council of American Survey Research Organizations. (2004). Code of Standards and Ethics for Survey Research. Consultado el 14 de diciembre de 2007 en <http://www.casro.org/codeofstandards.cfm>.
- Couper, M.P. (2000). Web Surveys: A Review of Issues and Approaches. *The Public Opinion Quarterly*, 64(4), 464-494.
- Cronk, B. C., and West, J. L. (2002). Personality research on the Internet: A comparison of Web-based and traditional instruments in take-home and in-class settings. *Behavior Research Methods, Instruments, & Computers*, 34(2), 177-180.
- Elwood, D. L. (1969). Automation of psychological testing. *American Psychologist*, 28, 694-699.
- Epstein, J., Klinkenberg, W. D., Wiley, D., & McKinley, L. (2001). Insuring sample equivalence across Internet and paper-and-pencil assessments. *Computers in Human Behavior*, 17, 339-346.
- Fortson B.L., Scotti J.R., Del Ben K.S. y Chen Y.C. (2006). Reliability and validity of an Internet traumatic stress survey with a college student sample. *Journal of Traumatic Stress*, 19(5), 709-720.
- Glaser, J., Dixit, J. y Green, D.P. (2002). Studying hate crime with the Internet: what makes racists advocate racial violence? *Journal of Social Issues*, 58, 177-93.
- Gosling, S. D., Vazire, S., Srivastava, S. y John, O. P. (2004). Should we trust web-based questions? A comparative analysis of six preconceptions about Internet questionnaires. *American Psychologist*, 59, 93-104.
- Gross, E., Juvonen, J. y Gable, S. (2002). Internet Use and Well-Being in Adolescence. *Journal of Social Issues*, 58(1), 75-90.
- Hallfors D.; Khatapoush S.; Kadushin C.; Watson K.; Saxe L. (2000). A comparison of paper vs computer-assisted self interview for school alcohol, tobacco, and other drug surveys. *Evaluation and Program Planning*, 23(2), 149-155.
- Hartford, K., Carey, R. y Mendonca, J. (2007). Sampling Bias in an International Internet Survey of Diversion Programs in the Criminal Justice System. *Evaluation & the Health Professions*, 30(1), 35-46.
- Holt, D., & Smith, T. M. F. (1979). Post stratification. *Journal of the Royal Statistical Society*, 142(1), 33-46.
- Kaplowitz, M. D., Hadlock, T. D., & Levine, R. (2004). A comparison of web and mail survey response rates. *The Public Opinion Quarterly*, 68(1), 94-101.
- Krantz, J.H. y Dalal R. (2000). Validity of Web-Based Psychological Research. En M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 61-85). San Diego, CA: Academic Press.
- Kraut, R., Kiesler, S., Boneva, B., Cummings, J., Helgeson, V., & Crawford, A. (2002). Internet paradox revisited. *Journal of Social Issues*, 58(1), 49-74.
- Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukopadhyay, T., & Scherlis, W. (1998). Internet Paradox: A Social Technology That Reduces Social Involvement and Psychological Well-Being? *American Psychologist*, 53(9), 1017-1031.
- Kypri, K. and Gallagher, S. J. (2003). Incentives to Increase Participation in an Internet Survey of Alcohol Use: A Controlled Experiment. *Alcohol & Alcoholism* 38(5): 437-441.
- Leung, D. y Kember D. (2005). Comparability of Data Gathered from Evaluation Questionnaires on Paper and Through the Internet. *Research in Higher Education*, 46(5), 571-591.
- Mathy, R.M., Schillace, M., Coleman, S.M., y Berquist, B.E. (2002). Methodological rigor with Internet samples: new ways to reach underrepresented populations. *CyberPsychology and Behavior*, 5, 253-66.
- Miller, E. (1968). A case for automated testing. *Bulletin of the British Psychological Society*, 21(71), 75-78.
- Mitofsky, W.J. (1999). Pollsters.com. *Public Perspective*, 10, 24-26.
- Musch, J. y Reips, U.D. (2000). A brief history of web experimenting. En M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 61-85). San Diego, CA: Academic Press.
- Nielsen, T.A., Stenstrom, P. y Levin R. (2006). Nightmare Frequency as a Function of Age, Gender, and September 11, 2001: Findings From an Internet Questionnaire. *Dreaming*, 16(3), 145-158.
- Observatorio de las Telecomunicaciones y de la Sociedad de la Información. (2006). Uso y perfil de usuarios de Internet en España. Consultado el 12 de diciembre de 2007 en: http://observatorio.red.es/estudios/documentos/uso_perfil.pdf.
- Oswald, F. L., Carr, J. Z., & Schmidt, A. M. (2001). The medium and the message: Dual effects of supervision and Web-based testing on measurement equivalence for ability and personality measures. Presentación a la decimosexta conferencia anual de la Sociedad de Psicología Industrial y Organizacional, San Diego, CA.
- Ployhart, R.E., Weekley, J.A., Holtz, B.C., & Kemp, C. (2003). Web-based and paper-and-pencil testing of applicants in a proctored setting: Are personality, biodata, and situational judgment tests comparable? *Personnel Psychology*, 56, 733-752.
- Robinson, J.P., Neustadt, A., Kestnbaum, M. (2002). The Online "Diversity Divide": Public Opinion Differences Among Internet Users and Nonusers. *IT & Society*, 1(1), 284-302
- Rosenbaum, P. R. y Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70, 41-55.
- Schonlau, M., Fricker, R.D. y Elliott, M.N. (2002). Conducting research surveys via e-mail and the Web. Santa Monica, CA: Rand.
- Skitka, L., y Sargis, E. (2005). Social psychological research and the Internet: the promise and peril of a new methodological frontier. En Y. Amichai-Hamburger (Ed.), *The Social net: Human behavior in cyberspace* (pp. 1-25). New York: Oxford University Press.
- Vispoel, W. P. (2000). Computerized versus paper-and-pencil assessment of self-concept: Score comparability and respondent preferences. *Measurement and Evaluation in Counseling and Development*, 33, 130-143.
- Vispoel, W. P., Boo, J., and Blieiler, T. (2001). Computerized and paper-and-pencil versions of the Rosenberg self-esteem scale: A comparison of psychometric features and respondent preferences. *Educational & Psychological Measurement*, 61(3), 461- 474.

(Artículo recibido fecha: 27-12-2007; aceptado: 21-4-08)