

# Retos metodológicos de la investigación psicológica a distancia

Baena, A.<sup>1,2</sup>, Fuster, H.<sup>2</sup>, Carbonell, X.<sup>2</sup> y Oberst, U.<sup>2</sup>

<sup>1</sup> Unitat de Tractament del Tabaquisme. Hospital Universitari de Bellvitge.

<sup>2</sup> Facultat de Psicologia, Ciències de l'Educació i l'Esport. Universitat Ramon Llull

## Resumen

*Resumen: Las Tecnologías de la Información y la Comunicación pueden ser un medio altamente eficaz para el desarrollo de investigaciones en general y la recolección, análisis e intercambio de datos en particular. La relativa facilidad y bajo coste del uso de aplicaciones a distancia, así como la potencial accesibilidad a grandes muestras, pueden marcar la estandarización de una nueva metodología de investigación en auge. Pero no todo son ventajas, también surgen nuevas dificultades y casuísticas que deben considerarse, principalmente relacionadas con la seguridad de la gestión de datos. Por ello el presente trabajo repasa la mayoría de preguntas, analiza el estado de la cuestión en todas sus vertientes, ya sean técnicas, éticas o, principalmente, de validez con el fin de arrojar luz y aportar soluciones prácticas sobre un tema que en nuestro país ha sido hasta ahora tratado de forma parcial o incluso marginal.*

**Palabras clave:** Internet, investigación a distancia, diseño de investigación.

## Autor/a de correspondencia:

**Antoni Baena**

Unitat de Tractament de Tabaquisme

Hospital Universitari de Bellvitge

Av. Feixa Llarga s/n

08907 Hospitalet de Llobregat

[antonibaena@tabaquisme.cat](mailto:antonibaena@tabaquisme.cat)

## Introducción

Las tecnologías de la información y de la comunicación (TIC) están revolucionando la forma en la que las personas se comunican en todos los ámbitos sociales imaginables, en todas las esferas y en todas las actividades sociales: la productividad, el comercio, la educación, el ocio, el entretenimiento... (Shapiro, 2001). Una parte de la población accede a Internet de forma natural para satisfacer su necesidad de conocimiento, o mejorar su salud, lo que está provocando una reacción en los terapeutas que han traspasado, habitualmente con éxito, naturalidad y fiabilidad, la terapéutica tradicional o el uso de herramientas de diagnóstico como los tests a un protocolo a distancia (Ritter, Lorig, Laurent y Matthews, 2004). Este uso de las TIC se remonta a los años setenta, cuando se desarrollaron programas de ordenador para incrementar la conciencia de las personas sobre sus hábitos alimentarios o el tratamiento computarizado de fobias a las arañas (Gallego, 2006). Pero los campos que han ido aumentando considerablemente son el estudio del potencial adictivo de la red (Estallo, 2001), la personalidad de los internautas y la realización de experimentos en línea (Bermejo, 1999; Steffens, 2001; Riso, 2001).

En cuanto al uso de pruebas de diagnóstico a distancia, los resultados han sido muy positivos ya que pueden ser estandarizados más fácilmente que los tests rellenos presencialmente. Las instrucciones dadas y el tiempo empleado para completarlos pueden ser monitorizados con mucha más facilidad y los resultados añadidos automáticamente a una base de datos y devueltos de inmediato. El uso a distancia comparte con el uso presencial la preocupación por los derechos de autor, pero se repite, para tranquilidad de los autores, que un test sin la clave y la interpretación de la respuesta por un profesional cualificado, es simplemente un texto sin sentido. Copiar un archivo PDF o fotocopiarlo es exactamente lo mismo.

Pero no todo son buenas noticias ya que, si bien hay estudios que han demostrado que el uso presencial o a distancia de tests psicológicos son equivalentes, el estudio de las propiedades psicométricas de éstos frecuentemente presenta problemas metodológicos en relación con la representatividad de la muestra o con la incorrecta asignación aleatoria, lo que dificulta la generalización y confianza de los resultados (Vallejo, Jordán, Díaz, Comeche y Ortega, 2007).

Es lógico pensar que unas herramientas que permiten acceder a múltiples servicios o profesionales sin contacto previo o intermediario, también deberían ser fácilmente aprovechadas por los investigadores para fo-

mentar la participación de abundantes muestras en estudios a distancia, mejorando incluso la validez externa (Murray *et al.*, 2009). Por eso surgen más preguntas que repuestas en cuanto a si este nuevo escenario global para la investigación va a complementar o mejorar la investigación tradicional, o qué dificultades deben superarse o controlarse para alcanzar los objetivos previstos en la investigación a distancia. Una buena noticia es que existe cierta evidencia de que Internet también se está convirtiendo en un escenario cada vez más aceptable para los encuestados como método de recogida de información, particularmente en el caso de los hombres (Dillman *et al.*, 2001).

Cabe mencionar que el uso de Internet se está convirtiendo en una actividad cotidiana para los casi 1.500 millones de usuarios alrededor del mundo, una cuarta parte de la población mundial. Según el último informe de la Fundación Telefónica (2009), en España, el 90% de los usuarios de Internet se conecta semanalmente y el 70% a diario; prueba de que el uso de la red es una conducta afianzada. Pese a todo, existe una desigualdad en la edad, formación y clase social ya que hay un marcado predominio de usuarios jóvenes, estudiantes de clase media alta que se conectan en casa (77,8%) con banda ancha; aunque sigue aumentando el segmento de usuarios de mediana edad. Es importante considerar que sólo un 8,8% de los españoles encuestados sobre el uso de las TIC consideraron que Internet es una actividad complicada.

En cuanto al uso, las principales aplicaciones que le dan los usuarios son las de búsqueda de información, correo electrónico, intercambio de vídeos y redes P2P de intercambio de archivos, junto con un considerable auge de las redes sociales en las que participan más de 13 millones de españoles.

### **Características definitorias del uso de las TIC**

No es posible acceder al uso de las TIC sin un ordenador, una conexión de red, cierta familiaridad con el uso y un navegador instalado. Estos prerequisites son necesarios pero no suficientes ya que, aunque nos encontraremos con usuarios que posean las herramientas físicas citadas, no serán capaces de utilizarlas eficazmente y de extraerles provecho si no han sido mínimamente alfabetizados en el mundo digital; en otras palabras, si no han aprendido a usarlas de forma adecuada. De modo que esta división entre usuarios hábiles y no hábiles está provocando una brecha digital entre los individuos, hogares, negocios, organizaciones y áreas geográficas de diferentes niveles socioeconómicos, dependiendo de la oportunidad que tienen de acceder a las TIC y el uso que dan a Internet en diferentes tipos de actividad (Obando, 2007).

En definitiva, esta necesidad de aprendizaje obliga a que cualquier herramienta de investigación a distancia sea lo más útil y simple posible para poder generalizar al máximo su utilización.

Por todo esto, el presente estudio a través de una revisión de la literatura científica existente sobre la investigación a distancia, persigue el objetivo de evaluar toda la información necesaria para iniciarse en la metodología a distancia, dar respuesta a la mayoría de preguntas que suscita una forma de investigar tan novedosa y, además, analizar los aspectos más prácticos de la misma, como la descripción de herramientas concretas o las medidas necesarias para cumplir adecuadamente con la legislación y/o código deontológico que la psicología en su condición de ciencia aplicada debe cumplir.

### **Características de la investigación a distancia. Retos metodológicos para mejorar la calidad de la intervención**

Creemos firmemente en la investigación a distancia, pero no en un uso indiscriminado o no crítico de la misma, de sus aspectos éticos, técnicos o metodológicos, cuyo análisis es el principal objetivo de este artículo.

Si bien, al usar Internet, el investigador principal puede tener un control total sobre la intervención, especialmente importante cuando participan múltiples terapeutas o centros, ya que la monitorización de la intervención es más fácil y el usuario, en comparación con el presencial, tiene mayor libertad en cuanto al número, frecuencia y duración del acceso a la información, herramientas utilizadas y participación más o menos activa en el estudio.

Así pues, el primer mito que hay que desechar es el que considera que las intervenciones a distancia son fáciles o que sólo consisten en una traslación del uso presencial al uso a distancia, ya que hay que considerar múltiples características funcionales.

### **Aspectos éticos y confidencialidad de los datos**

Una investigación a distancia no puede nutrirse únicamente de los códigos deontológicos existentes para los profesionales de la salud, ya que no contempla muchos aspectos importantes diferenciales del uso de las TIC. Por eso, diversas entidades han ido paulatinamente desarrollando una serie de códigos deontológicos recomendados cuando se traslada la intervención presencial a la de distancia. Los ejemplos más fiables son: Health on the Net ([www.hon.ch](http://www.hon.ch)), el programa de certificación Web

Médica Acreditada del Colegio Oficial de Médicos de Barcelona ([wma.comb.es](http://wma.comb.es)) o las “Recomendaciones y criterios deontológicos de la intervención psicológica mediante Internet” del Colegio Oficial de Psicólogos de Catalunya ([www.copc.cat/content/view/409/451/](http://www.copc.cat/content/view/409/451/)). Por otra parte, en España existe una legislación muy restrictiva sobre protección de datos (Ley Orgánica 15/1999) lo que puede dificultar su cumplimiento especialmente en un servicio a distancia. Entre sus principios está el recoger datos que no sean excesivos o innecesarios para la labor para la cual han sido recogidos; no recoger datos para otros fines; informar previamente al interesado de dicha recogida y sobre el derecho de toda persona al acceso, cancelación y rectificación de dichos datos.

En el caso de la recogida de datos de salud hay un apartado específico, el 6 del artículo 7, que permite la recogida de datos por parte de profesionales de la salud cuyo código deontológico recoja el secreto profesional, lo que no evita que sea necesario tener el consentimiento manifiesto, libre, inequívoco, específico e informado de la persona que aporta los datos. En cualquier caso todo fichero debe ser registrado en la Agencia Española de Protección de Datos e incluir un responsable del mismo que será la persona que velará por el cumplimiento escrupuloso de la ley, especialmente en referencia al acceso a la información recogida.

Hay que tener en cuenta que el principal problema relacionado con el uso de la comunicación a distancia, en todas sus modalidades, es la seguridad en el intercambio y el almacenamiento de la información. Éste es un tema primordial para todo investigador ya que, según todos los indicios, los participantes que se inhiben de participar en un estudio a distancia perciben la falta de seguridad en este proceso (Paul, Seib y Prescott, 2005). En cambio, entre quienes participan hay gran aceptación de las demandas de confidencialidad y anonimato, incluso cuando existe la posibilidad real de que la contraseña y la encriptación del estudio puedan ser descubiertos o de que el Protocolo de Internet (IP) sea reconocido (Granello y Wheaton, 2004).

En definitiva, participantes e investigadores, necesitan confiar en que la información incluida en un formulario electrónico y en la comunicación por correo electrónico no será interceptada y hecha pública (Paul, Seib y Prescott, 2005), por lo que deberá valorarse la utilización de protocolos seguros o emails encriptados, por una parte, y el seguimiento escrupuloso de la legislación vigente en cuanto a la protección de datos por otra, dada la importancia que tiene.

Así pues, es esencial que las personas que participen en una investigación a distancia conozcan todos y cada uno de los posibles riesgos que el envío de información pueden provocar y den el consentimiento informado de los mismos, especialmente cuando estos datos deben ser reenviados o compartidos con terceras personas. El cómo, es un tema a tener muy presente ya que, según la información demandada, el asentimiento simple puede no ser suficiente y la inclusión de un procedimiento de consentimiento por correo postal tiene dos impactos negativos. Primero, se puede influir en la recepción de la muestra, especialmente si la población a la que va dirigida valora muy positivamente el anonimato. Segundo, enviar un consentimiento por correo y esperar su retorno puede producir un retraso significativo en el proceso de inclusión. También debe estarse seguro de que durante el seguimiento o en cualquier momento de la investigación, todos los participantes saben cómo retirarse fácilmente del estudio cuando lo deseen.

Por otro lado, cualquier investigación que evalúe psicológicamente a las personas debe considerar los posibles efectos negativos que puedan resultar de la propia intervención o de los resultados de los tests administrados (Wood, Griffiths y Eatough, 2004). En cualquier caso, la simple presencia de riesgos no excluye necesariamente el uso de una intervención, a distancia o no, si está justificado por los beneficios potenciales. Por lo tanto, es responsabilidad de los investigadores entender la naturaleza del riesgo y minimizarlo, así como la naturaleza de los posibles beneficios, para poder informar de todo ello a los participantes en el estudio, de modo que puedan tomar una decisión meditada (Childress, 2000).

### **Aspectos técnicos de la investigación online**

Debemos tener en cuenta que el primer requisito necesario para la creación de una herramienta de valoración basada en una página web consiste en contar con el conocimiento necesario para su programación, tarea no siempre fácil para los profesionales de la psicología, que puede obligar a contratar los servicios externos de profesionales de la informática y a incluir el coste en el presupuesto general del estudio desde el principio. Cuando se trate de una investigación pequeña sin presupuesto suficiente, el coste puede ser demasiado caro y dificultar la investigación.

Aunque hay que decir que algunos autores han evaluado en un 50% la reducción del coste de una investigación cuando se realiza a distancia, a través de una página web en vez de por teléfono, y en un 20% menos si se realiza por correo ordinario (Granello y Wheaton, 2004). Cosa que

puede rentabilizar el coste de implementación de la web. No olvidemos contabilizar ciertos costes ocultos, como la búsqueda de información sobre el modo de financiarla, las herramientas para el diseño y desarrollo de un protocolo de investigación a distancia o la fácil diseminación de los resultados de la investigación (Paul, Seib y Prescott, 2005). Otros factores que suelen abaratar el coste del estudio son una menor necesidad de personal de administración, una reducción de los informes en papel, seguridad y copias de seguridad, distribución de la información del estudio entre los participantes o centros y los formularios de recogida de datos en una única localización (simplifican la diseminación de los resultados y evitan la pérdida accidental de información).

Aunque puede encarecer mínimamente el estudio es muy recomendable la compra de todos los dominios relacionados con los estudios para no caer en el llamado *cybersquatting*. Éste consiste en que una empresa ajena se aproveche de la publicidad que supone la investigación para atraer gente a sus páginas, utilizando dominios engañosos o muy parecidos, habitualmente con finalidad mercantil o incluso de captación ilícita de datos. Por ejemplo, podemos incluir una investigación en una dirección web como [www.miestudio.es](http://www.miestudio.es) y no tener el dominio [www.miestudio.com](http://www.miestudio.com) con lo cual confundir una con otra sería muy fácil.

Una vez superados estos aspectos, si la red donde se implementa el estudio es adecuada, comparada con una página web, el email u otras herramientas de comunicación, permite a los investigadores utilizar diseños flexibles en cuanto al color, los gráficos, las pantallas, las animaciones, los vídeos, el sonido, la realidad virtual, la realidad aumentada, etcétera, multiplicando considerablemente las posibilidades (Granello y Wheaton, 2004).

### **Softwares de diseño de cuestionarios online y otras herramientas**

Las herramientas para la investigación *online* (de bajo coste o gratuitas) permiten la creación de cuestionarios y la gestión de las respuestas. Los paquetes básicos son ideales para la creación de cuestionarios sencillos, son de bajo coste y fácil manejo. Perfectos para la creación de cuestionarios sencillos en los que no se necesitan preguntas complejas, lógica de respuesta o análisis integrado de datos. Algunos ofrecen un mayor abanico de posibilidades, como la personalización del diseño o la integración dentro de un website propio, entre otros. Entre los más utilizados están:

- SurveyMonkey ([www.surveymonkey.com](http://www.surveymonkey.com)): uno de los servicios más populares. En su versión gratuita permite la creación ilimitada de cuestionarios, con un máximo de 10 preguntas y 100 respuestas cada uno. Así mismo, el paquete básico ofrece escasa personalización del cuestionario y no permite la descarga de datos/informes. En su versión de pago más allá de eliminar los límites en preguntas y respuestas, se añade la posibilidad de crear cuestionarios multi-página y con lógica de respuesta, así como la descarga de datos para ser analizados en otras aplicaciones (PDF, HTML, XML y CSV).
- Zoomerang ([www.zoomerang.com](http://www.zoomerang.com)): similar en muchos aspectos al anterior en su versión gratuita o de bajo coste, pero ofrece un paquete profesional con mayores prestaciones. Pese a no tener una *interface* intuitiva, su sistema de informes y gestión de datos es muy completo con gestión multiusuario, envío de cuestionarios a teléfonos móviles, etcétera.
- SurveyGizmo ([www.surveygizmo.com](http://www.surveygizmo.com)): la versión gratuita permite un máximo de 250 respuestas mensuales. Otras versiones añaden funciones e incrementan el número de respuestas.

Aparte de éstas, existen otras herramientas integradas que ofrecen, además del diseño y uso de los cuestionarios como los anteriormente comentados, el envío masivo de emails.

- Constant Contact ([www.constantcontact.com](http://www.constantcontact.com)): pese a tratarse de una empresa de marketing por correo electrónico ofrece una herramienta de creación de cuestionarios.
  - Form Site ([www.formsite.com](http://www.formsite.com)): Incluye versiones con diferentes limitaciones. Sus principales características son la personalización del cuestionario y la posibilidad de integrarlo en una página web propia.
- Moodle ([www.moodle.com](http://www.moodle.com)): se trata de un sistema de código abierto que integra la gestión de contenido web con la gestión de cursos, incluyendo cuestionarios y test.
- CATI (*Computer Assisted Telephone Interviewing*): se trata de una técnica de entrevista a través del teléfono en la que el entrevistador sigue un guión suministrado por un software informático, que garantiza una gestión estructurada de la información y agiliza el proceso de captación y uso de los datos estadísticos. El software es capaz de ajustar el cuestionario según las respuestas recibidas y la información ya conocida.

Por último están los paquetes avanzados que, por su mayor complejidad, son los menos accesibles. Pese a todo, son ideales para investigaciones a



gran escala, gracias a la posibilidad que ofrecen de crear preguntas más complejas así como análisis de datos más exhaustivos.

- SurveyZ ([www.surveyz.com](http://www.surveyz.com)): ofrece una lógica avanzada de respuestas y la posibilidad de análisis complejos, más orientado al ámbito académico. Es posible conseguir una versión gratuita con limitaciones, o convenir un presupuesto según las necesidades del usuario.
- Key Survey ([www.keysurvey.com](http://www.keysurvey.com)): se trata de la herramienta más eficaz, con un paquete completo de características avanzadas como la gestión de redes de usuarios, soporte para cuestionarios “alumno/profesor”, opciones multimedia, etcétera.
- Lime Survey ([www.limesurvey.org](http://www.limesurvey.org)): una herramienta gratuita de código abierto instalable en un servidor propio, ideal para el ámbito académico. Entre sus cualidades se cuenta la personalización completa del cuestionario, el soporte para 40 idiomas distintos en un mismo cuestionario, lógica avanzada de respuestas, el sistema de gestión por correo, etcétera. Tiene una amplia comunidad de investigadores que participan activamente en su desarrollo y cuenta con constantes actualizaciones. El principal inconveniente radica en su dificultad de uso, configuración e instalación, que requiere conocimientos de programación PHP/MySQL.

## El reclutamiento

Éste es a menudo el principal desafío de las investigaciones convencionales, llegando incluso a afirmarse que una tercera parte de las investigaciones fallaron a la hora de conseguir el tamaño de muestra deseable. Si lo trasladamos al tema que nos ocupa, lo primero que debe preocuparnos es si el teórico potencial de usuarios por enrolar es una realidad y, en segundo, si la muestra conseguida poseerá suficiente validez interna y externa.

En cuanto a lo primero, para alcanzar al máximo de la población se suelen utilizar diversas vías de comunicación: invitaciones por correo electrónico, publicidad *online* con *banners* en cualquier web o en aquellas específicas para la investigación, invitaciones en foros o grupos de discusión y el *lurking* en grupos de discusión (Murray 2009). El *lurking* es un fenómeno típico de los foros y grupos de usuarios entre los que hay un gran número de registrados que consultan habitualmente lo que el grupo genera, pero pocas veces participa. Es una audiencia silenciosa.

Sin embargo, según algunos autores, si utilizamos el número de visitantes como denominador de la participación, con frecuencia la investiga-

ción a distancia presenta una tasa de respuesta muy baja (Eysenbach, 2004). Aunque no siempre es así, ya que se han encontrado resultados similares o incluso mejores al comparar el correo postal o un protocolo web (Schonlau, 2004), por lo que no podemos llegar a conclusiones certeras sobre este aspecto. En lo que sí parece haber cierto acuerdo es en el hecho de que una de las formas para maximizar el número de participantes en un estudio consiste en la explicación detallada de quién investiga y por qué se realiza la investigación (Wood, Griffiths y Eatough, 2004).

La muestra seleccionada tendrá un sesgo desde el inicio que provendrá principalmente de la naturaleza no representativa de la población general ya que ha sido reclutada *online*, es decir, no incluye a personas no conectadas y, además, la mayoría de veces se da una autoselección de los participantes (efecto del voluntarismo). Se ha visto que las investigaciones que utilizan Internet como escenario tienen una muestra cuya media de usuarios es hombre, blanco, con trabajo, buena educación y un status socioeconómico alto (Paul, Seib y Prescott, 2005). Además muchas veces el grupo de usuarios de Internet seleccionado pertenece a la comunidad universitaria (Vallejo, Jordán, Díaz, Comeche y Ortega, 2007), muestra explotada tradicionalmente en exceso en la investigación psicológica que, ahora, también puede serlo, provocando una validez externa muy criticable por la ya comentada autoselección y un reclutamiento en grupo cerrado (Riva, Teruzzi y Anolli 2003). Este problema puede convertirse en una oportunidad única, teniendo en cuenta que los jóvenes son las personas que generalmente adoptan las nuevas tecnologías más temprana y naturalmente, personas que difícilmente sean accesibles de otra forma (Koo y Skinner, 2005).

Encontramos otra desviación en el hecho de que obviamente los usuarios reclutados en la web tienen una mayor probabilidad de responder a una investigación web que la población en general. Eso sí, el coste de los servicios y un ordenador, junto con la necesidad de una alfabetización digital seguirá haciendo muy difícil la inclusión de la población con nullos o escasos conocimientos, teniendo en cuenta que el nivel lingüístico necesario para el uso de Internet es avanzado (Birru *et al.*, 2004).

Por último, una investigación a distancia proporciona acceso a aquellas personas socialmente poco hábiles, que no suelen tomar parte en estudios presenciales (Wood, Griffiths y Eatough, 2004), dado que la ausencia de anonimato contribuye a eliminar el efecto desinhibitorio o a aumentar el efecto de deseabilidad social, o a aquellas personas que

son reacios a facilitar una dirección postal para recibir el consentimiento informado o evitar la participación múltiple.

En resumen, podemos mejorar el reclutamiento si desarrollamos una estrategia multicanal, monitorizamos para determinar las tasas adecuadas y consideramos la validez interna y externa de la muestra conseguida. Además no restemos importancia al hecho de que el proceso aleatorio está totalmente automatizado y es fiable, ya que se basa en la generación por ordenador de números aleatorios (Murray *et al.*, 2009).

### La participación múltiple

Merece un punto y aparte la posible participación múltiple de un mismo usuario. Que una misma persona se inscriba y responda más de una vez puede alterar considerablemente los resultados si esa circunstancia se repite metódicamente. Si como hemos visto, el anonimato es el punto fuerte de algunas investigaciones, éste también reduce el control sobre los datos ya que impide poder contrastarlos adecuadamente. Aunque, según algunas estimaciones, la participación duplicada se encuentra por debajo del 3% en la mayoría de estudios (Reips 2002), lo que no debería ser un problema para la calidad de la información recogida de los estudios basados en Internet; incluso así hay una serie de acciones correctivas que deben tenerse en cuenta.

En España, la mayoría de usuarios tiene una IP dinámica que simplemente con reiniciar el módem podrían cambiar fácilmente sin ser detectados, porque aparecen como un usuario totalmente diferente. Por otra parte, es cada vez más habitual que una misma IP sea compartida por diferentes personas o empresas de una misma red, por lo que borrar directamente las entradas duplicadas de IP no es recomendable ya que no necesariamente corresponden a un mismo usuario.

Si aún conociendo la baja incidencia de la duplicación de datos creemos necesario incluir mecanismos de detección, deberemos valorar los pros y los contras de las posibles opciones. En primer lugar podemos implementar un mecanismo de validación de usuario por correo electrónico que nos permita verificar los datos a posteriori, provocando una pérdida total o parcial del anonimato; una medida no del todo fiable que puede influir directamente en la participación. No es fiable porque en la actualidad es sumamente fácil activar diferentes emails de los proveedores (Hotmail, Gmail, Yahoo, etcétera), dado que un único individuo puede poseer múltiples identidades y participar en un estudio muchas veces (Koo y Skinner, 2005).

Otras opciones paralelas pueden ser el uso de *cookies* y/o la eliminación de incentivos por participar en la investigación. En cuanto a las *cookies* podemos decir que sólo son efectivas si se intenta acceder dos veces al estudio con el mismo ordenador aunque, sólo con borrarlas, se evita fácilmente la función para la que se han diseñado.

### La calidad de los datos recogidos

Deben tenerse en cuenta principalmente tres cuestiones: la validez de los resultados como ya hemos dicho; la alteración de las características psicométricas de la administración *online* de cualquier consulta y la cantidad de ítems no contestados.

En relación con la validez de los datos se dan dos implicaciones importantes. Algunos datos, como las características demográficas, edad o género, no pueden ser verificadas con absoluta fiabilidad. Además, la mayoría de formularios habitualmente han sido diseñado para ser rellenados con lápiz y papel. Por lo tanto, cualquier modificación en el modo de entrega de los resultados puede modificar las propiedades psicométricas.

Ya que el software a distancia puede obligar a los participantes a responder todas las preguntas, lo que debe preocuparnos es el número de personas que prefieren no rellenar el formulario a contestar todas las preguntas, aunque no hay estudios sobre este particular.

### La tasa de respuesta-retención

La primera pregunta que debe formularse es cómo es la tasa de respuesta de una investigación a distancia. Los primeros estudios han devuelto unos resultados más bajos de lo inicialmente esperado, con una tasa de respuesta de entre el 10-25% lo cual convierte la retención en el principal reto de los estudios a distancia; particularmente cuando todo el seguimiento se realiza también a distancia. Aunque de forma arbitraria se ha valorado como punto de corte una tasa de respuesta de entre el 70-80% para que los resultados sean valorados con mucho escepticismo (Eysenbach, 2004).

Si como vemos, la tasa de respuesta web es menor incluso que la de la investigación que utiliza el correo postal, no debería darse por un hecho que el gran potencial de extensión y facilidad de uso de las TIC se traducirá en una mayor proporción de respuestas y, por lo tanto, debería dedicarse mucha más energía y recursos con el fin de reducir al máximo este problema.

Según todos los indicios, entre las principales acciones que mejoran la tasa de respuesta está la utilización de un sistema de múltiples recordatorios. Una investigación web normalmente alcanza una proporción mayor de respuestas cuando la muestra es contactada por email en lugar de por correo postal. Su utilización llega incluso en algunos estudios a doblar la tasa de respuestas; aunque en otros, el resultado ha sido más modesto. Lo económico y fácil de su implementación aconsejan su utilización. Gracias a los múltiples canales de comunicación que las TIC brindan, algunos autores han podido también mejorar las tasas de retención y respuesta usando recordatorios *offline* (cartas o teléfono) o incentivos financieros (Murray *et al.*, 2009). En definitiva, aquellos estudios que han utilizado métodos mixtos, incluyendo el correo postal o los recordatorios por teléfono junto con el email, han alcanzado tasas más elevadas de seguimiento. Sin embargo, hay que recordar que este tipo de seguimiento es considerablemente más caro y consume más tiempo.

Los recordatorios pueden presentar algunos problemas técnicos ya que la preocupación por los virus hace que muchos usuarios de Internet borren directamente los emails no solicitados o con remitentes desconocidos. Además se da el caso de que los propios filtros de los correos electrónicos puedan interpretarlos como *spam* y los eliminen.

Por otra parte, cualquier recordatorio sea por la vía que sea, puede convertirse en un elemento intrusivo para los participantes (Koo y Skinner, 2005) por lo que es importante tener un protocolo de antemano para decidir el número de recordatorios que serán utilizados en cada fase de la investigación pero, sobre todo, sus limitaciones y la posibilidad de cancelación en cualquier momento por parte de los mismos sujetos investigados.

La herramienta de recogida de datos es un factor que debe valorarse, ya que influye considerablemente sobre la tasa de respuesta. Podemos conseguir herramientas con diseños innovadores y atractivos que nos convertirán en la envidia del resto de investigadores, pero quizás el valor que damos a la estética nos haga perder parte de su aprovechamiento, ya que nos obliga a utilizar lo último en software, cosa que, hasta que su uso se haya generalizado, complica el acceso a los usuarios no avanzados. Los estudios con diseños sofisticados (tablas HTML, colores múltiples, movimiento, sonido) tienen una tasa de respuesta un 5% menor que aquellos que utilizan una plantilla o formato más simple. Ese diseño fácil y funcional del que estamos hablando debe incluir: una página principal de bienvenida; formatos que sean similares al lápiz

y papel; líneas limitadas para evitar la necesidad de usar las barras de desplazamiento lateral y el uso de instrucciones en cada página en lugar de un único listado al principio, como puntos principales (Dillman *et al.*, en Granello y Wheaton, 2004). En resumen, no se ha demostrado que el diseño actúe positivamente sobre la tasa de respuesta, pero sí negativamente.

Por último, se ha visto que cuando a los participantes se les explica inicialmente –en la web, email o en ambos soportes– las características más importantes del cronograma de la investigación como el tiempo de duración de la investigación, cuándo y cuántos emails recibirán (ordinarios o recordatorios), si recibirán o no contraseñas automáticas de acceso, etcétera, la tasa de respuesta mejora (Granello y Wheaton, 2004).

### La gestión de la información

Si llegados a este punto hemos diseñado una buena y funcional herramienta de recogida de datos, hemos alcanzado una muestra representativa de la población que queríamos investigar, hemos evitado la participación múltiple y tenemos una tasa de respuesta-retención suficiente, nuestro foco de atención se centrará en qué y cómo gestionar la información.

Hasta el momento se solía emplear mucho tiempo en la reintroducción de los datos y en la depuración de los mismos para su posterior análisis estadístico. Esto puede ser historia en la investigación a distancia ya que la tecnología te permite planificar una base de datos centralizada a la cual pueden acceder todos los investigadores desde su lugar de trabajo, con lo que se facilita la investigación multicéntrica. Con esa base de datos ya no es necesaria la reintroducción de unos datos que ya ha introducido el propio sujeto de estudio, a quien, además, hemos forzado a responder a todas las preguntas para finalizar el cuestionario (No deje esta pregunta en blanco). Otra ventaja añadida es que esta agilización de todo el proceso de análisis permite la utilización de herramientas estadísticas *online* y la obtención de resultados casi de forma inmediata.

Una cuestión adicional es la proporción de participantes que reciben una intervención durante el estudio. Ya que suele darse una baja retención, o lo que es lo mismo, una proporción sustancial de sujetos no recibirá la intervención. La forma de eliminar este riesgo pasa por realizar un análisis de “intención de tratamiento” en el que los resultados de todos los participantes son analizados según el tratamiento al cual fueron

asignados. El efecto real de la intervención se subestima, pero se acerca más al efecto real de la misma.

Por lo tanto, en una investigación a distancia podemos encontrar ciertas limitaciones, como ya hemos dicho, en la representatividad o tamaño de la muestra, en el tipo de información recogida, en la ausencia de comunicación no verbal, en la calidad de los datos, en la ausencia de muestra sin acceso a Internet, en las bajas tasas de respuestas, en la mayor pérdida de información (un error en el email al registrarse hace imposible recuperar al usuario), etcétera, pero de lo que podemos estar seguros es de que el proceso de análisis de la información lo ha convertido en algo infinitamente más fácil e inmediato.

## Conclusiones

Es obvio que cualquier tecnología virtual no sustituye a la actividad real, sino que la complementa, incluso puede que ayude a estimular las actividades “reales” existentes. Por eso encontramos importantes y fundamentadas razones metodológicas para utilizar diseños a distancia en investigación. El fácil acceso a una amplia muestra, la automatización de recogida de datos, la reducción del coste, etcétera, son ventajas que no podemos dejar pasar. Eso sí, se tendrá que evaluar y seleccionar adecuadamente a los sujetos objeto de estudio ya que en una investigación *online* fácilmente puede alterarse lo aleatorio y la validez interna, dificultando el control y la validez estadística de la investigación. Además, no debería darse por hecho que el potencial acceso a una muestra muy amplia se traducirá en una mayor tasa de participación.

Por otra parte, las características de los investigadores que decidan participar o dirigir una investigación a distancia deben ampliarse. Ya no es suficiente con que conozcan su propia materia de estudio y la metodología necesaria para investigar, ahora además deberán ser competentes y estar capacitados en el uso de todas las herramientas tecnológicas que deben usar.

En resumen, puede que las TIC en general e Internet en particular no se conviertan en la herramienta principal de la investigación en un futuro cercano, pero es casi seguro que van a formar parte de un enfoque integrado. Irán fortaleciendo su papel dentro de la investigación a medida que vayan creando metodologías y modelos propios.

Categoría	Variable	Explicación
Diseño	Descripción del diseño	- Descripción del diseño de la investigación (muestra, representatividad, número de usuarios necesarios), etc.
	Usabilidad	- ¿Es el programa de recogida de datos simple, fácil de leer y comprensible?
	Adaptación / discapacidad	- World Wide Web Consortium (W3C) o la Web Accessibility Initiative, (WAI).
Ética, deontología y protección de datos	Aprobación del comité ético	
	Consentimiento informado	- Forma de recogida, aceptación expresa o por escrito.
	Ley española de protección de datos	- Seguir lo expresado en la legislación - Registrar la base de datos - Tener un responsable de la seguridad
	Código deontológico a distancia	- HON, WMA y COPC
	Información de los límites temporales, legislativos, éticos y deontológicos del estudio	- Duración del estudio - Protección de los datos: tiempo y lugar de almacenamiento de los datos, cesión a terceros, etc. - Responsable y código de registro de la base de datos en la AGPD - Número de emails generales y recordatorios
Reclutamiento	Estudio abierto vs. cerrado	- Estudio accesible a todos los visitantes de la página web o estudio accesible por usuario y contraseña, previamente conocidos por los investigadores. - Limitación geográfica
	Modo de contacto	- Email - Teléfono - Correo postal - Canal múltiple - Ninguno si se valora el anonimato
	Publicidad del estudio	- Dónde, formato, tiempo de visibilidad, etc.
	Prevención multientrada	- Cookies - Control IP - Registro de email - Otros mecanismos: correo postal, teléfono, etc.



Categoría	Variable	Explicación
	Incentivos	- Sí/No - Formato y limitaciones del mismo
Administración del estudio	Web / email/ correo postal / teléfono	- Determinar si se utilizará un protocolo por email o basado en una página web, junto con la determinación de cómo la información será introducida en el ordenador. - Determinar otros elementos multicanal como el teléfono y/o el correo postal
	Obligatorio/voluntario	
	Hora, fecha y zona geográfica	
	Aleatorización de los ítems	- Sí/No
	Número de ítems idóneo	
	Número de páginas web	
	Comprobación de la integridad de las respuestas	- Facilitar o no la posibilidad de corregir y reevaluar las respuestas dadas antes de enviar el cuestionario
Análisis de los datos	Manipulación de los cuestionarios incompletos	- ¿Serán totalmente desechados o sólo parcialmente?
	Cuestionarios enviados con una fecha y hora atípica	- Cuándo podrán ser aceptados o no.
	Correcciones estadísticas	- Análisis de intención del tratamiento - Generalización de los resultados - Valoración de la validez interna y externa
	Tasa de respuestas	- Visitante único - Tasa de visión de la web - Tasa de participación - Tasa de finalización
Evaluación de la calidad de la investigación	Medidas preventivas	- Valorar qué aspectos pueden mejorarse - Implementar soluciones a problemas teóricos
	Medidas correctivas	- Desarrollar un pretest - Evaluación constante del proceso - Implementar las medidas y volver a evaluar el impacto de éstas

## REFERENCIAS BIBLIOGRAFICAS

Bermejo, A. (1999). Adicción a Internet. *Revista de Información Psicológica*, 70, 27-30.

Birru, M.S., Monaco, V.M., Charles, L., Drew, H., Njie, V., Bierria, T., Detlefsen, E. & Steinman, R.A. (2004) Internet Usage by Low-Literacy Adults Seeking Health Information: An Observational Anlaysia. *Journal of Medical Internet Research*, doi:10.2196/jmir.6.3.e25.

Childress, C.A. (2000) Ethical Issues in Providing Online Psychotherapeutic Interventions. *Journal of Medical Internet Research*, doi:10.2196/jmir.2.1.e5.

Dillman, D.A., Phelps, G., Tortora, R., Swift, K., Kohrell, J. & Berck, J. (2001) *Response Rate and Measurement Differences in Mixed Mode Surveys Using Mail, Telephone, Interactive Voice Response and the Internet*. Trabajo presentado en Annual Conference of American Association for Public Opinion Research, Montreal, Canada.

Estallo, J.A. (2001). Usos y abusos de Internet. *Anuario de Psicología*, 32, 2.

Eysenbach, G. (2004) Improving the Quality of Web Surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *Journal of Medical Internet Research*, doi:10.2196/jmir.6.3.e34.

Fundación Telefónica (2009). *La sociedad de la Información en España 2009. Colección Fundación Telefónica*. Madrid: Ariel.

Gallego, M.J. (2006). *Un estudio controlado que compara un tratamiento autoadministrado vía Internet para el miedo a hablar en público vs. el mismo tratamiento administrado por el terapeuta*. Disertación doctoral no publicada, Universitat Jaume I, Castelló de la plana, España.

Granello, D.H. & Wheaton, J.E. (2004) Online Data Collection: Strategies for Research. *Journal of Counselling & Development*, 82(4), 387-394.

Koo, M. & Skinner (2005) Challenges of Internet Recruitment: A Case Study with Disappointing Results. *Journal of Medical Internet Research*, doi:10.2196/jmir.7.1.e6.

Murray, E., Khadjesari, Z., Whita, I.R., Kalaitzaki, E., Godfrey, C., McCambridges, J., Thompson, S.G. & Wallace, P. (2009) Methodological Challenges in Online Trials. *Journal of Medical Internet Research*, doi:10.2196/jmir.1052.

Obando, C. (2007). Las tecnologías de la información y la comunicación (TIC): un nuevo escenario para el Desarrollo Local de las Comunidades (Estudio de Caso: Comunidad Ómnia-Barrio el Raval, Barcelona). Disertación doctoral no publicada, Universidad Ramon Llull, Barcelona, España.

Paul, J., Seib, R. & Prescott, T. (2005) The Internet and Clinical Trials: Background, Online Resources, Examples and Issues. *Journal of Medical Internet Research*, doi:10.2196/jmir.7.1.e5.

Reips, U. (2002) Standards for Internet-based experimenting. *Experimental Psychology*, 49(4), 243-256.

Risso, A. (2001). ¿Experimentos psicológicos a través de Internet? *Anuario de Psicología*, 32(2), 109-116.

Ritter, P., Lorig, K., Laurent, D. & Matthews, K. (2004) Internet Versus Mailed Questionnaires: A Randomized comparison. *Journal of Medical Internet Research*, doi:10.2196/jmir.6.3.e29.

Riva, G., Teruzzi, T. & Anolli, L. (2003) The Use of The Internet in Psychological Research: Comparison of Online and Offline Questionnaires. *Journal of CyberPsychology & Behavior*, 6(1), 73-80.

Schonlau, M. (2004) Will Web Surveys Ever Become Part of Mainstream? *Journal of Medical Internet Research*, doi:10.2196/jmir.6.3.e31.

Shapiro, A.L. (2001). *El mundo en un clic*. Barcelona: Grijalbo Mondadori.

Steffens, K. (2001). Self-regulation and computer based learning. *Anuario de Psicología*, 32(2), 77-94.

Vallejo, M.A., Jordán, C.M., Díaz, M.I., Comeche, M.I. & Ortega, J. (2007) Psychological Assessment via the Internet: A reliability and Validity Study of Online (vs Pape and Pencil) Versions of the general Health Questionnaire-28 (GHQ-28) and the Symptoms Chek-List-90.Revised (SCL-90-R). *Journal of Medical Internet Research*, doi:10.2196/jmir.9.1.e2.

Wood, R. Griffiths, M.D. & Eatough, V. (2004) Online Data Collection from Video Game Players: Methodological Issues. *Journal of CyberPsychology & Behavior*, 7(5), 511-518.

### Resum

Les Tecnologies de la Informació i la Comunicació poden ser un mitjà eficaç per el desenvolupament d'investigacions en general i per la recollecció, anàlisi i intercanvi de dades en particular. La relativa facilitat i baix cost de l'ús d'aplicacions a distància, així com el potencial d'accessibilitat a grans mostres, poden marcar l'estandardització de una nova metodologia d'investigació en apogeu. És per això que el present treball repassa la gran majoria de preguntes, analitza l'estat de la qüestió en totes les seves vessants, ja siguin tècniques, ètiques o de validesa principalment, amb la finalitat d'aportar llum i solucions pràctiques sobre un tema que en el nostre país ha estat fins ara tractat de forma parcial o inclús marginal.

**Paraules clau:** Internet, investigació a distància, disseny d'investigació

### Abstract

Information and Communication Technologies can be an effective means to develop research in general and to collect, analyse and exchange data in particular. The relative easiness and low cost involved in using long distance applications, as well as the potential to access big samples, may mark the standardization of a new research method. For this reason, this article reviews most questions, analyses the state of the art in all its perspectives, whether technical, ethical or validity, with the aim of shedding some light and contributing with practical solutions about a topic that in our country has been dealt with partially or even marginally so far.

**Key words:** Internet, online research, methodology, research design.