

Encuestas por Internet y nuevos procedimientos muestrales

MILLÁN ARROYO MENÉNDEZ Y LUCILA FINKEL*

RESUMEN

Este artículo presenta una panorámica sobre las encuestas por Internet, partiendo de la evolución del fenómeno y explicitando sus principales aplicaciones, ventajas e inconvenientes, así como sus especificidades metodológicas.

Asimismo, se analizan las posibilidades y limitaciones de los distintos tipos de muestreo y sus estrategias de recogida de información. Se concluye proponiendo la utilización de muestreos no probabilísticos por cuotas en los que, tomando ciertas precauciones en la planificación del trabajo de campo y ejerciendo ajustes por ponderación, se consiguen muestras parecidas a las que se obtendrían por muestreo probabilístico.

1. EVOLUCIÓN Y SITUACIÓN ACTUAL DE LAS ENCUESTAS ONLINE

Aunque existen referencias de actividad de encuestas *online* desde los años noventa del siglo pasado (Deutskens, Ruyter y Wetzels, 2006), su despegue en España se sitúa en la segunda década del presente siglo. A partir de ese momento, el acceso a Internet de la pobla-

* Departamento de Sociología: Metodología y Teoría, Facultad de Ciencias Políticas y Sociología, Universidad Complutense de Madrid (millan@cps.ucm.es, lfinkel@ucm.es).

ción empieza a ser mayoritario, el correo electrónico alcanza una amplísima penetración social y comienza a crecer la utilización de redes sociales y aplicaciones de mensajería instantánea. Al mismo tiempo, el largo periodo de recesión económica vivido en nuestro país conlleva una importante disminución de las encuestas por entrevista personal o telefónica y un incremento de las *online*.

De acuerdo con datos facilitados por ANEIMO (Asociación Nacional de Estudios de Mercado y Opinión Pública), en España la expansión de las encuestas *online* se produce principalmente a lo largo de la última década, pasando de representar el 3,6 por ciento del total de encuestas en 2008 al 20,7 por ciento en 2017. Se aprecia también un cierto retraso en relación a los datos mundiales, si bien las distancias se han acortado mucho en los últimos años: en 2008 la modalidad *online* era el tipo de encuesta principal en el mundo, con un 20 por ciento de penetración, pasando al 25 por ciento en 2017 (ANEIMO, 2018).

Este crecimiento es concomitante a la rápida penetración de Internet y con esta, a la digitalización de la sociedad, que también evoluciona muy rápidamente hasta alcanzar en la actualidad a la mayoría de la población, mientras que hace diez años se había consolidado el uso mayoritario del correo electrónico, y estaba

comenzando a expandirse el uso de redes sociales que era elevado entre los más jóvenes, pero no había alcanzado aún al resto de la población. En 2018, el 85 por ciento de la población española usa Internet, el 80 por ciento utiliza teléfonos móviles, el 58 por ciento usa redes sociales, y entre ellos, el 50 por ciento es usuario activo de dichas redes sociales (We are Social, 2018).

El *Estudio general de medios* (AIMC, 2019) situaba la cifra de españoles que usa de forma frecuente Internet (tomando como indicador el día anterior a la encuesta) en el 38 por ciento de la población en 2010 y el 54 por ciento en 2013, cifra que en 2018 alcanzaba ya al 78 por ciento. Estos datos explican que sea a partir de los últimos diez años cuando se aprecia una fuerte expansión de las encuestas *online*, de forma paralela a cómo las encuestas telefónicas se generalizaron siguiendo el ritmo de implantación de los teléfonos en los hogares.

La digitalización de la sociedad está permitiendo que el reclutamiento y participación en este tipo de encuestas esté aumentando y mejorando, dado que cada vez hay más personas que tienen acceso a Internet a través de distintos soportes y aplicaciones de mensajería instantánea o redes sociales.

Junto con su progresiva difusión, se suceden las opiniones que cuestionan la calidad de las respuestas obtenidas, sugiriendo que la información proporcionada es menos precisa y menos completa que la de las encuestas administradas por otros medios. Por ello, entendemos que empieza a ser el momento de plantearse seriamente qué posibilidades tiene esta modalidad de encuesta de mostrar a colectivos a los que tradicionalmente se ha representado mediante muestreos polietápicos y entrevistas personales o telefónicas.

En este artículo se abordan algunas de las principales características de la investigación mediante encuestas realizadas a través de Internet (o de la web), que en la práctica se conocen mediante el vocablo anglosajón de encuestas *online* y no tanto por lo que podría ser su traducción más directa al castellano de "encuestas en línea". Por ello, a efectos de simplificación, en este texto se utilizará el término de encuesta por Internet y encuesta *online* de forma indistinta.

2. VENTAJAS E INCONVENIENTES DE LAS ENCUESTAS POR INTERNET

Como acabamos de ver, las encuestas por Internet están siendo cada vez más populares y su uso se está extendiendo, en parte porque sus ventajas frente a otras formas de indagación resultan muy evidentes (Evans y Mathur, 2005).

La principal ventaja y causa de esta expansión es su bajo precio, comparado con las encuestas personales e incluso con las telefónicas, que en su día proliferaron principalmente porque eran mucho más baratas que las personales. La reducción de costes que implican las encuestas por Internet es sustancial, dado que no es necesario contar con entrevistadores, aunque no puede perderse de vista el coste que supone poner en marcha y mantener un panel de personas activo (Duffy *et al.*, 2005).

A ello se une el ahorro económico de lo que suponían las tareas de depuración manual de los cuestionarios en papel y la grabación de datos con su posterior depuración para detectar errores de grabación. Estas tareas han desaparecido con la digitalización, dado que en las encuestas *online* es el entrevistado quien se encarga de marcar las respuestas, que se graban automáticamente. Como las preguntas que van apareciendo responden a filtros predeterminados, y el número máximo de opciones de respuesta también está prefijado, hay nulas posibilidades de errores de grabación y de respuesta.

Otra ventaja es el ahorro de tiempo, que en su momento también supuso una gran ventaja de la encuesta telefónica sobre la personal. Las encuestas *online* por lo general suelen requerir tiempos más cortos de recogida de datos que el necesario en las encuestas telefónicas.

La facilidad de uso, junto con la simplicidad logística que brinda la digitalización, son los factores que explican, aparte del precio, su creciente éxito. El hecho de que no se requiera de conocimiento experto para implementarlas resulta muy atractivo. En la actualidad, existen infinidad de aplicaciones sencillas que permiten a cualquier persona diseñar un cuestiona-

rio para su uso en línea¹. Las versiones básicas de los distintos programas suelen ser gratuitas, aunque si se desean utilizar características adicionales como un mayor número de variables o casos, exportar los datos a distintos formatos, quitar los logos de la aplicación e incluir los propios o mayores posibilidades de formato de preguntas, hay que comprar licencias anuales que generalmente resultan bastante accesibles. Entre las aplicaciones más conocidas, cabe citar a SurveyMonkey, Google Forms, SoGoSurvey o Encuestafácil. El programa *Limesurvey*, basado en *software* libre, ofrece mayores posibilidades de personalización que otras aplicaciones, porque además de permitir la programación de nuevas funcionalidades al usuario avanzado, cuenta con comunidades de desarrolladores en todo el mundo.

Lamentablemente, no todo son ventajas, y son muchas las aportaciones que ponen de manifiesto los principales inconvenientes y desafíos que pueden aparecer en la investigación a través de Internet (Baltinic *et al.*, 2002; Duda y Nobile, 2010; Duffy *et al.*, 2005; Díaz de Rada, 2012).

En primer lugar, nos encontramos con el problema de la validez muestral, que tiene que ver con la dificultad de contar con marcos muestrales representativos donde cada sujeto tiene igual probabilidad de ser seleccionado en el estudio. A menudo no se cuenta con listados de direcciones de correo electrónico de la población global que pudieran equivaler al listado maestro de teléfonos que se utiliza en las encuestas telefónicas. Por tanto, en este tipo de encuestas se usan mayoritariamente muestras no probabilísticas (Díaz de Rada, 2012).

En segundo lugar, el problema de la no respuesta, derivado de la autoselección, o la exclusión automática de aquellos sin acceso a Internet, constituyen fuentes importantes de sesgo identificadas por los metodólogos cuando reflexionan sobre la calidad de las encuestas *online* (Pavía *et al.*, 2011; Díaz de Rada, 2012). Es evidente que el que contesta una encuesta por Internet probablemente esté mucho más dispuesto e interesado en el tema y sea más propenso a encontrar el estudio en el propio

¹ Ver Wright, 2005 para una descripción detallada de las principales características de varias aplicaciones de encuestas *online*.

proceso de búsqueda de información sobre la materia. Sobre la exclusión de los que no tienen acceso a Internet, como se ha expuesto, la situación está cambiando rápidamente. Hace tan solo unos años la brecha digital era mucho mayor que en la actualidad, pero aún existen segmentos de población que no tienen acceso o no acceden regularmente a Internet.

Para analizar el error derivado de la no respuesta y su relación con la calidad de los datos recogidos, se han venido tomando como principales indicadores la tasa de respuesta y la tasa de abandono del cuestionario. En distintas investigaciones se ha constatado que las tasas de respuesta en las encuestas *online* son tan bajas como las que se obtenían con las encuestas postales, y que los factores más importantes para conseguir mayor participación tienen que ver con la personalización de la contactación (que se realiza a través de un correo electrónico previo si se cuenta con un listado de potenciales participantes o "marco muestral") y la persistencia del equipo investigador, que se cuantifica a través del número de recordatorios de participación que se remiten a cada potencial participante. Cuando se trabaja con bases de datos de encuestados, existe la posibilidad de seguimiento o *tracking*, gracias a que los ficheros que se crean automáticamente permiten identificar quiénes han abandonado la encuesta sin haberla completado².

En tercer lugar, en tanto cuestionario autorrellenado, debido a la falta del control que puede ejercer un encuestador, está la peor calidad de cumplimentación del cuestionario, que puede derivar en falta de respuesta parcial (no responder a una parte de las preguntas) o no responder con la diligencia debida. Para abordar la no respuesta parcial se están adoptando medidas más finas que permiten establecer y mejorar de forma más precisa la calidad de las respuestas. Así, como complemento a la tasa de abandono, se trabaja con la tasa de compleción del cuestionario o la proporción de preguntas contestadas sobre el total (Deutskens *et al.*, 2004), el número de ítems contestados (sin considerar los "no sabe" o los "no contesta"), la implicación y esfuerzo realizado por la per-

² Hay cierto consenso en que el abandono se produce generalmente tras la visualización de las dos primeras pantallas de los cuestionarios, lo cual implica que los participantes tienden a hacer una exploración rápida y decidir si continúan o no. Por el contrario, no se suelen producir muchos abandonos cuando se alcanzan las últimas pantallas.

sona que responde (medidos a través de la longitud de las respuestas abiertas, el número de respuestas ofrecidas en preguntas de respuesta múltiple, la variedad de las puntuaciones de las baterías de ítems con la misma escala) y en algunos casos hasta se pregunta al encuestado por el nivel de satisfacción con el cuestionario (Ganassali, 2008).

En cuarto lugar, puede ocurrir que el sesgo de los interesados (*stakeholders*) en obtener determinados resultados contribuya a la participación masiva de personas orientadas en una dirección, como de hecho ha ocurrido ante algunas consultas políticas. Medidas tales como la limitación a una respuesta por dirección de correo electrónico o dirección IP de un dispositivo, pueden ayudar a paliar este problema.

Por último, en quinto lugar, es bastante habitual tener que hacer frente al problema de las personas encuestadas no verificadas, o a personas de las que se desconoce su procedencia sociodemográfica, lo cual complica el trabajo con las cuotas preestablecidas.

En lo que sigue analizaremos algunos de los aspectos diferenciales de los cuestionarios *online*, que deben también ser tenidos en cuenta cuando se adopta este tipo de estrategia metodológica.

3. ALGUNOS ASPECTOS DIFERENCIALES DE LOS CUESTIONARIOS POR INTERNET

La estructura del cuestionario *online* ha sido objeto de distintas consideraciones. En sus orígenes, se dudaba de si resultaba mejor contar con una única página donde todas las preguntas aparecieran de forma consecutiva –lo cual permitiría al encuestado estimar su duración y características de entrada–, o si era preferible organizar las preguntas en distintas pantallas que fueran apareciendo consecutivamente a medida que se fueran cumplimentando. En los últimos años, se está imponiendo la organización de las preguntas en bloques que aparecen en distintas pantallas, aunque se suele insertar una banda de progreso en la parte inferior o superior de la pantalla, que va mostrando el porcentaje de preguntas presentadas sobre el total, para ofrecer a la persona encuestada información contextual sobre la actividad que se le requiere.

Aunque este aspecto no es exclusivo de los cuestionarios *online*, resulta esencial cuidar la redacción y el formato de las preguntas, considerando incluso aspectos concretos como el tratamiento de las cuestiones más sensibles (Coutts y Jann, 2008). Muchos manuales insisten en la necesidad de formular preguntas cortas, directas, sin tecnicismos ni dobles negaciones y circunscritas a un momento determinado, dado que todo ello influye necesariamente sobre la calidad de las respuestas obtenidas (entre otros, Baltinic *et al.*, 2002). Por otro lado, diversas investigaciones han demostrado que la forma en la que se presentan gráficamente las preguntas incide en la calidad de las respuestas: un mayor espacio en las preguntas abiertas arroja respuestas más extensas, se obtienen mayores tasas de respuesta con preguntas donde hay que marcar la respuesta en una casilla que cuando se trata de elegir una opción en una lista desplegable, etc. (Ganalassi, 2008; Kaczmirek, 2008).

Algunos autores ponen el énfasis en la interactividad (Ganalassi, 2008), considerada como el principal elemento diferencial de los cuestionarios *online*. La posibilidad de que el cuestionario se adapte en tiempo real al perfil de la persona que responde (lo cual supone flexibilizar su estructura y la redacción de las preguntas) estaba ya presente en los cuestionarios personales, pero no era descartable la posibilidad de que se produjeran errores humanos. Con la encuesta por Internet y otros sistemas *offline* digitalizados, el sistema va presentando a la persona que responde solamente las preguntas que se derivan de sus respuestas anteriores. Asimismo, pueden programarse instrucciones personalizadas o la posibilidad de cumplimentar el cuestionario en distintos idiomas.

Otro aspecto fundamental y clave para la viabilidad de una encuesta *online* es la duración del cuestionario y la consideración de otros factores que inciden sobre el tiempo como el grado de familiaridad con Internet y la velocidad de lectura en pantallas, así como el tipo de dispositivo desde el que se está cumplimentando la encuesta (por ejemplo, se estima que desde un teléfono móvil inteligente, se tarda entre un 20-30 por ciento más tiempo en contestar el mismo cuestionario que desde un ordenador³).

³ Ver <https://blog.questionmark.com/how-many-questions-should-you-have-in-a-web-survey>

Tradicionalmente, se venía pensando que el tiempo de cumplimentación no debería nunca pasar de los 20 minutos, pero distintas investigaciones han matizado esta recomendación. Así, en un interesante estudio (Revilla y Ochoa, 2017) dirigido a panelistas profesionales muy acostumbrados a contestar encuestas *online*, se determinó que la duración máxima media debía ser de 22,7 minutos (con un valor mediano de 20 minutos), pero al mismo tiempo se señalaba que la duración ideal debía estar en los 12,6 minutos de media (10 minutos de valor mediano). En ambos casos, un análisis de regresión múltiple determinó que las variables significativas que influían en la variación de la duración eran el agrado que suscitaba el cuestionario, la confianza en la anonimidad de las respuestas y la edad (a menor edad, mayor tiempo posible). En el caso de las encuestas cumplimentadas a través de los móviles, se recomienda que no duren más de 15 minutos, que no contengan más de dos preguntas abiertas, que las preguntas no tengan más de 15 atributos, que no haya que hacer *scrolling* horizontal, que no se utilice Adobe Flash Player (los Iphone no lo soportan) y que las escalas no excedan de cinco posiciones (Kantar Media, 2017).

No debe perderse de vista que los cuestionarios que se realizan a través de Internet permiten recopilar de forma automática metadatos o información sustancial sobre el propio proceso de cumplimentación: por ejemplo, la fecha, la hora de comienzo y finalización (y por tanto el tiempo transcurrido), el número de pantallas que han aparecido, en qué momento se produce el abandono, etc. Las modernas técnicas de neuromarketing que recopilan datos psicométricos, registrando los movimientos de la retina (*eye tracking*), las distintas cadencias al teclear o los movimientos realizados con el ratón, permitirán sin duda en un futuro no muy lejano afinar los análisis de carácter metodológico para mejorar la aplicación de los cuestionarios *online*. En todo caso, es muy recomendable seguir escrupulosamente las recomendaciones de la *Guía ESOMAR* (2012) para la investigación *online*, que contempla cuestiones éticas, regulatorias, tecnológicas y metodológicas.

Resulta obligado mencionar la tendencia creciente hacia la utilización de metodologías mixtas de recogida de información para asegurar una mayor tasa de respuesta y una reducción de costes, pero no debe perderse de vista que la combinación de los resultados obtenidos es

delicada y conviene asegurarse de que las metodologías son compatibles.

En la comparación entre las encuestas *online* y las encuestas personales asistidas por ordenador (CAPI) se ha puesto de manifiesto que, teniendo en cuenta la menor tendencia a la deseabilidad social y a evitar respuestas incorrectas marcando más el “no sabe” (“no estoy seguro” en las encuestas *online*) ambos resultados podrían ser equiparables utilizando formas de ponderación adecuadas (en base a aspectos demográficos o en función de las diferencias actitudinales o de comportamiento – “*propensity score weighting*”) (Duffy *et al.*, 2005).

En todo caso, la utilización de modelos mixtos debería estar condicionada a la existencia de lo que se denomina “validez de convergencia” entre ambos tipos de encuestas, o lo que es lo mismo, que lo que se mida no dependa del tipo de instrumento utilizado.

4. TIPOS DE MUESTREO APLICADOS A ENCUESTAS *ONLINE* Y ESTRATEGIAS PARA LA RECOGIDA DE LA INFORMACIÓN

Las posibilidades y limitaciones de las encuestas *online* han condicionado tanto las estrategias de recogida de información como los diseños muestrales, ambos aspectos estrechamente relacionados.

En términos generales, es mucho más difícil ejercer un control de la muestra en encuestas *online* que en encuestas telefónicas o personales (Sánchez-Carrión y Segovia, 2008). Un interesante estudio que comparaba dos modalidades de encuesta *online* entre sí (con muestreo por cuotas y muestreo bola de nieve) y también con una entrevista personal y otra telefónica, utilizando en todas ellas el mismo cuestionario, concluía que la mayor calidad de las muestras la ofrecía la entrevista personal, seguida de la telefónica. A distancia quedaban las dos modalidades de encuestas *online*, de las cuales la de bola de nieve era la que presentaba una peor calidad (Szolnokin y Hoffmann, 2013).

Como hemos comentado, las encuestas autorrellenadas que la modalidad *online* comparte con otros cuestionarios autocumplimen-

tados suelen presentar tasas de no respuesta total (o rechazo a responder a la encuesta) muy elevadas, dado que no se sienten comprometidas por un entrevistador (Groves, 1989; Díaz de Rada, 2012). No obstante, diversos estudios revelan que entre los cuestionarios autorrellenados, los *online* presentan tasas de respuestas aún más bajas que cuando se utilizan cuestionarios en papel (Nulty, 2008).

Quienes contestan suelen ser una pequeña minoría de quienes son invitados a responder y esto supone un sesgo, aparte del esfuerzo de reclutamiento, ya que a menudo esa minoría no presenta las características generales del colectivo objetivo. Por ejemplo, en una encuesta de satisfacción laboral, si no se ejerce presión para que respondan todos los empleados, tenderán a responder los más implicados con la temática, y tenderán a ser los más satisfechos y también los más descontentos.

Este condicionamiento hace que se tienda a trabajar más con encuestas *online* en muestreos no probabilísticos que en muestreos probabilísticos. Los muestreos probabilísticos se basan en la selección aleatoria de la muestra, y buscan que todas las variables del cuestionario estén distribuidas proporcionalmente a como se encuentran en la población, dentro de los márgenes de error aleatorio que la muestra asuma (normalmente entre +/- un 2 por ciento y un 5 por ciento), mientras que los muestreos no probabilísticos no buscan ni pueden afrontar esa clase de representación, sino otras diferentes. Su valor es menor en la investigación sociológica porque no se puede inferir que una determinada característica de la muestra esté presente en esa medida o siquiera aproximada en la población. Sin embargo, se aplican habitualmente en la investigación de mercados cuando se trata simplemente de recabar opiniones o valoraciones de consumidores promedio de un determinado producto, del que solo se desean determinar los puntos fuertes y débiles, sin intentar inferir en qué medida están presentes en la población.

Trabajando con encuestas *online*, hay varios tipos de muestras posibles. Sin pretensión de exhaustividad, pero tratando de reflejar lo más frecuente y representativo, presentaríamos los siguientes:

- a) En muestreos no probabilísticos
 - Muestras por bola de nieve

- Diseños por cuotas

b) En muestreos probabilísticos o que buscan similaridad distributiva

- Muestras mediante listados o bases de datos
- Muestras basadas en un panel
- Otras con énfasis en el control a *posteriori*

También merecen breve mención las malas prácticas o encuestas carentes de rigor. Como botón de muestra, cabe mencionar las encuestas que a menudo aparecen en periódicos digitales y numerosas webs, realizadas sin ningún control de quien responde y cuyo propósito es dar una idea aproximada de opiniones ante temas de actualidad o políticos. Ese tipo de encuestas pueden servir como entretenimiento, pero no tienen ningún valor científico y no serán consideradas aquí. Lamentablemente, a menudo se pretende que el sondeo sea un reflejo de la realidad, como ocurre a veces en la prensa digital.

4.1. Muestreo por bola de nieve

Los sondeos *online* más frecuentes, que existían mucho antes de las encuestas *online*, son los que utilizan como criterio de reclutamiento el método de bola de nieve, porque son los más simples y económicos, y más accesibles a quienes no son expertos en encuestas. Consisten en dirigir la encuesta en una primera fase a los contactos personales, y luego a partir de ahí procurar que los que han respondido faciliten a su vez otros contactos, y así sucesivamente. A veces los contactados no responden a la encuesta porque no cumplen los requisitos u otros motivos, pero contribuyen dándole difusión. La versión *online* de este método consiste en utilizar los contactos de las redes sociales (Whatsapp, Telegram, Facebook, Twitter, Instagram, LinkedIn, etc...), animando tanto a responder la encuesta como a darle difusión. Puede ser una buena opción cuando no se requiere que la muestra sea distributivamente análoga al universo y no se necesita recurrir a la inferencia estadística. Es decir, cuando los objetivos de la investigación permiten un muestreo no probabilístico, y funda-

mentalmente, cuando además nuestro universo utiliza las redes sociales.

El potencial atractivo de este método radica en que el número de contactos en cada paso aumenta exponencialmente. Suponiendo que la media de contactos sea de cien personas, (ronda ese valor como promedio) la invitación a la encuesta al final de la segunda etapa podría llegar a 10.000 contactos (100 x 100) y al final de la tercera etapa nada menos que a un millón (100 x 100 x 100). Pese al alto potencial de “viralización” de la invitación a la encuesta, la realidad es que la inmensa mayoría de las personas a las que llega dicha invitación no suelen responder ni reenviar la invitación a nadie. La bola de nieve *online* planteada de forma automática no suele ser eficaz si no se interviene. La experiencia demuestra que la implicación personal es insustituible. Se puede responder a la encuesta (y quizás reenviar después a otros) si te lo pide alguien que ejerce una influencia sobre ti. Y aún así, si este lo hace mediante un envío colectivo se tenderá a ignorar. Lo que mejor funciona es el contacto individualizado, de tú a tú; que alguien conocido te lo pida personalmente, que era el principio en el que se basaba el método de bola de nieve antes de Internet. En esencia, hay que hacer las cosas como antes, con la gran ventaja de que Internet nos puede hacer más fácil y asequible el contacto con las personas y nos resuelve (enviando el vínculo al cuestionario de forma telemática) con gran facilidad la distribución logística; no hay más que pinchar en el vínculo para empezar a responder, desde tu ordenador, móvil o tableta.

No obstante, está cada vez más extendida la utilización de técnicas basadas en el *marketing* digital para el reclutamiento de entrevistados en encuestas *online*, a través de *banners*, listas de distribución, *influencers*, códigos *UTM*⁴, etc., contribuyendo a mejorar la tasa de respuesta en los muestreos por bola de nieve. Cuando predominan estas técnicas, se puede sustituir la denominación ‘bola de nieve’ por *river sampling* (Trujillo Carmona, 2019).

Se trata de rastrear la circulación del movimiento en las redes para conocer mejor de dónde provienen las respuestas y diseñar estra-

⁴ Los códigos *UTM* (*Urchin Traffic Monitor*) son pequeños fragmentos de texto que se añaden al final de la *URL*. Su principal objetivo es ayudar a seguir el éxito de un sitio web y el tráfico que llega.

tegias y campañas para mejorar los resultados globales o de distintos colectivos. Con todo, los resultados obtenidos por estos reclutamientos activos se quedan lejos de la eficacia de los reclutamientos activos mediante la vinculación personal. Mientras en el primer caso la tasa de respuesta promedio oscila según un análisis entre el 5 por ciento y el 10 por ciento, en el segundo caso se sitúa entre un 50 por ciento y un 70 por ciento (Sánchez-Carrión y Segovia, 2008).

Hay que asumir que con el método de bola de nieve no es posible ejercer un control *a priori* sobre la muestra. No cabe por tanto ningún diseño de la muestra, porque no sabemos con qué muestra nos vamos a encontrar hasta que damos por terminada la recopilación de datos. Como mucho, se puede controlar *a posteriori* ejerciendo una ponderación, si conocemos la distribución del universo y si la muestra no está tan desequilibrada que la ponderación sea imposible o muy limitada. Por regla general, las ponderaciones no son una buena solución cuando la muestra se ha reclutado de cualquier manera y presenta importantes descompensaciones, que es lo que cabe esperar en estos casos. Por eso, debe ubicarse su uso dentro de los muestreos que no pretenden alcanzar la similaridad distributiva del universo. Estudios antecedentes no han hallado esta representatividad de estos sondeos ni siquiera tras intentar ajustes por ponderación (Szolnoki y Hoffmann, 2013).

4.2. Muestreos por cuotas, no probabilísticos

Otros tipos de muestreos no probabilísticos pueden establecer cuotas de entrevistas en algunas variables. Por ejemplo, en encuestas de consumo podría diseñarse una muestra en función de algunas características como la frecuencia de consumo de un determinado producto o marca, edad y sexo. Estas cuotas no estarían en función de la distribución proporcional de estas características en la población, sino en función de las necesidades de comparación de resultados entre las categorías de estas variables u otras. Los resultados se compararían entre sí y se testarían mediante pruebas de significación estadística las diferencias de porcentajes o de

medias en las diferentes respuestas. Muchas de estas distribuciones serían del tipo: 50 por ciento varones, 50 por ciento mujeres; un tercio grandes consumidores, un tercio consumidores intermedios, un tercio consumidores ocasiones; un tercio jóvenes, un tercio de edad intermedia, un tercio mayores.

Uno o varios captadores podrían ocuparse de seleccionar, mediante fuentes de información muy diversas, aquellas personas que cumplen los requisitos, invitándoles a participar en la encuesta *online*. Este tipo de muestreos sería análogo al de otras encuestas personales o telefónicas con idénticos propósitos, con la diferencia de la flexibilidad y versatilidad que ofrece la encuesta *online*, siempre y cuando quienes se ocupen de la captación tengan la opción de contactar digitalmente con las personas seleccionadas. En la medida que la población posee y maneja razonablemente bien un *smartphone* o un ordenador, esto es posible. Hasta hace tan solo unos años, esta modalidad de muestreo no era posible o no resultaba tan práctica o conveniente como otras formas de captación, pero en la actualidad empieza a ser una opción factible en un número creciente de investigaciones por encuesta, dada la elevada y veloz penetración de los usos digitales.

Utilizar a uno o varios captadores es el principio más elemental para poder controlar estas muestras, pudiendo ser un punto de partida para diseñar y reunir muestras más complejas y con más posibilidades de las que aquí se describen. No obstante, en este tipo de muestras a veces se combinan las cuotas con el método de bola de nieve. Cuando esto se hace, para no perder el control de las mismas, conviene rodar la bola de nieve al principio, para reducir la intervención de los captadores, pero en la fase final al menos, la mitad o más de la recogida de la información, deben acometerla ellos, si no queremos perder el control de las cuotas. Los captadores, además de las cuotas, controlan otras características de los respondentes: criterios generales que deben cumplir todos los que hacen la encuesta, como por ejemplo residir en ciertos territorios, no ser mayores o menores de determinada edad, o cumplir ciertas condiciones de idoneidad para participar en la encuesta. El grado de especificidad de todo esto, así como la duración del cuestionario, pueden desaconsejar recurrir parcialmente al método de bola de nieve para un reclutamiento exigente. Una

limitación del uso de la captación por cuotas es que si somos exigentes en relación al número de variables a controlar, sobre todo si se ejerce el control sobre las cuotas cruzadas entre sí, aumenta considerablemente la dificultad del reclutamiento. Para que sea factible y no excesivamente complejo, no conviene que intervengan más de tres o cuatro variables o una docena de categorías de respuesta en el conjunto de las diversas variables a considerar. Pasando de estos límites la captación se empieza a complicar.

4.3. Muestreos probabilísticos, con listados o bases de datos

Disponer de bases de datos o listados facilita mucho la labor de reclutamiento de la muestra, y, si dichos registros recogen a todo el universo (incluyendo teléfonos o direcciones de correo electrónico u otras formas de contacto digital) nos permitirán diseñar y quizás reclutar una muestra probabilística (Pavía *et al.*, 2011). Los listados y bases de datos también pueden ser utilizados en diseños no probabilísticos, por cuotas, liberándonos de la peor alternativa de recurrir al método de bola de nieve. Pero son de especial relevancia por representar el tipo de situaciones en principio más favorables para obtener una muestra probabilística en una encuesta *online*. También suele liberar de la necesidad de utilizar captadores, lo que agiliza y abarata el trabajo.

Cuando el universo no es grande y comprende unos pocos miles de individuos o menos, una alternativa es no muestrear y dirigir la encuesta a todos los individuos de la base de datos. Como lo más probable es que no respondan todos, obtendremos *de facto* una muestra. Si las características de los que no respondieron no difieren de los que cumplieron el cuestionario, entenderemos que la muestra obtenida ha funcionado como una muestra aleatoria. En caso contrario, trabajaríamos con un sesgo que habría que intentar equilibrar mediante ponderación, para que la muestra se comporte de la manera más parecida posible a una seleccionada completamente al azar.

Pero cuando el universo es grande y queremos ejercer un buen control sobre la muestra será necesario muestrear, seleccionando aleatoriamente a un conjunto representativo de

sujetos así como otro subconjunto también aleatorio para prever sustituciones. Las garantías de éxito de recoger finalmente una buena muestra dependen de la capacidad que tengamos de influir para que la mayoría de personas respondan. Aunque las tasas de respuesta a encuestas autorrellenadas es habitualmente muy baja, cuando existen listados o bases de datos lo más frecuente es estar en situaciones en las que sí es posible ejercer esa influencia. Por ejemplo, en estudios dirigidos a empleados de una empresa, o sus clientes o proveedores, estudiantes, miembros de una organización, etc. En estas ocasiones el vínculo con la organización que promueve la encuesta permitiría obtener tasas de respuesta relativamente elevadas, si el cuestionario no es muy largo y sabemos motivar adecuadamente. En cualquier caso, es importante comparar en lo posible las características de quienes responden con las de quienes no responden para asegurar la calidad de la muestra obtenida.

4.4. Muestreos por panel

Cuando el objetivo es entrevistar a la población general de un país o territorio extenso, o a un colectivo amplio aunque diferente de la población general, la encuesta *online* se encuentra mucho más limitada que las telefónicas o personales. Al problema de la elevadísima tasa de no respuesta se añade la dificultad práctica de contactarles digitalmente. Como respuesta a esta problemática, algunas empresas especializadas ofrecen muestras a partir de un panel. Inicialmente estos paneles se han reclutado mediante entrevistas personales en hogares, o a veces por teléfono, pero concertando con los panelistas el envío periódico de distintas encuestas *online*. El prerreclutamiento de panelistas a través de otros medios evita errores de cobertura y de no respuesta. Este prerreclutamiento sirve para seleccionar también a las personas interesadas en participar (Sánchez-Carrión y Segovia, 2008).

Es decir, la idea del panel no es destinarlo a un único estudio, salvo rara excepción, sino a muchos, a lo largo de años, por el elevado coste de montarlo y mantenerlo. El conjunto de panelistas es proporcionalmente representativo al del conjunto de la población, en las variables habituales que se utilizan en los diseños de muestras

de encuesta, si bien con la especificidad de que todos manejan y entienden los medios digitales para responder a dichas encuestas. A este sesgo hay que añadir el peligro de la profesionalización del panelista, que se habitúa a responder encuestas y podría hacerlo de forma diferente al del resto de la población, mucho menos expuesta a encuestas. La implicación del panelista suele estar incentivada económicamente, aunque las cuantías por responder suelen ser modestas. Con todo no siempre está garantizado que respondan siempre, si bien las tasas de respuesta de estos paneles no suelen presentar problemas (Sánchez-Carrión y Segovia, 2008).

Se trata en definitiva de un tipo especial de muestreo por cuotas, ya que la selección de los panelistas de la muestra se hace por cuotas de las principales variables de clasificación (sexo, edad, estudios, distribución geográfica, etc.). Y la muestra podrá ser análoga al universo contemplado en el panel (la población general mayor de 16 o 18 años, por lo común) o segmentada a partir de esta. Una diferencia operativa importante de la selección de una muestra por cuotas a través de un panel frente a otro tipo de muestreos de encuestas *online* por cuotas sin panel es la necesidad de rechazar participantes cuando una cuota se cubre, como principio para mantener el equilibrio de la muestra (Ochoa y Porcar, 2018).

Los muestreos por panel constituyen la forma habitual de obtener una muestra *online* dirigida a la población general o muy amplia de un territorio también amplio, cuando tratamos de obtener muestras distributivamente representativas del universo. Hay que entender estas muestras como mera aproximación razonable a dicho objetivo, dado que no suelen reunir la calidad de las muestras polietápicas de entrevistas personales o telefónicas, pero a cambio son mucho más baratas y rápidas, dicho sea a pesar de que este tipo de encuesta suele ser el más caro de las encuestas *online*.

4.5. Muestreos por cuotas buscando la similaridad distributiva del universo

Una modalidad alternativa de muestreo para encuestas *online*, pretendiendo cierta

aproximación razonable a una muestra probabilística, consiste en utilizar muestras con cuotas proporcionales al universo (o a veces mediante bola de nieve, con un cierto control de cuotas) y procurar ajustes por ponderación en las variables donde se aprecien desajustes. Se trata de muestras de diseño sencillo, con énfasis en el control *a posteriori*, utilizando captadores que se responsabilicen del control de las cuotas (o dinamizadores para mover la encuesta en redes). Estudios anteriores han afirmado que es posible encontrar la representatividad distributiva de estas muestras tras aplicar ajustes por ponderación (Szolnoki y Hoffmann, 2013).

Son factibles cuando de un lado hay información del universo suficiente como para evaluar y en su caso ajustar por ponderación los resultados de la muestra, y de otro lado, cuando quienes colaboren en la captación tengan posibilidades de reunir casos de todos los segmentos del perfil tipológico del público al que nos dirigimos. La clave del éxito dependerá de un lado de la centralidad de las variables controladas para que la muestra no presente desequilibrios insalvables ante una ponderación, y de otro lado, de los ajustes por ponderación que se puedan efectuar, en función de dichos desequilibrios, de la información disponible del universo y de la propia estrategia de ponderación adoptada en función de esos y otros factores. En encuestas dirigidas a la población general o a poblaciones muy amplias, las variables más propicias para el control de cuotas pueden ser: edad, estudios, ocupación (al menos si trabajan o no), sexo y ámbito territorial.

5. RESULTADOS DE LA EXPERIENCIA DE INVESTIGACIÓN *ONLINE*

Quienes firmamos este artículo acumulamos cierta experiencia en este último tipo de muestreos de encuestas *online*, en los que se pone el énfasis en la ponderación, con el propósito de aproximar una muestra no probabilística lo más posible al ideal de similaridad distributiva de los muestreos probabilísticos, habiendo alcanzado diversos niveles de éxito.

Los peores resultados (comparando datos del universo con el mejor ajuste posible de la muestra, tras ponderar) los hemos encontrado

en encuestas donde no se han aplicado captadores, sino dinamizadores que intervienen tanto en redes sociales como buscando fuera de las redes digitales. Los fuertes desequilibrios de la muestra han sido insalvables mediante ponderación, porque las cuotas teóricas de las variables fundamentales, en algunas categorías de respuesta no se han llegado a cubrir, o han reunido muy pocos casos. También hemos apreciado importantes problemas de concentración en determinados territorios de otras características importantes de la población, con ausencia de las mismas en otros territorios. Esta experiencia es también referida en contextos diferentes por Szolnoki y Hoffmann (2013), concluyendo estos autores que los muestreos por bola de nieve no pueden alcanzar la representatividad distributiva.

Mejores resultados se obtienen aplicando captadores, y tanto mejor cuando la población de referencia pertenece a un territorio homogéneo o más reducido que a un ámbito territorial más extenso o complejo. Por ejemplo, las peores experiencias provienen de muestras de ámbito nacional, mientras que las mejores se han obtenido a nivel de distrito, habiendo alcanzado resultados aceptables en encuestas en la Comunidad de Madrid o en la ciudad de Madrid, cuando se han utilizado captadores responsables del cumplimiento de cuotas. También es muy destacable que los resultados son mucho mejores si nos dirigimos a población joven que a población de más edad, por la familiaridad de los jóvenes con los medios digitales. De ambos aspectos no hemos encontrado referencias previas en la bibliografía.

Como ejemplo de un buen resultado exponemos a continuación una encuesta de 403 casos, realizada de octubre de 2014 a enero de 2015 mediante cuotas, dirigida a la población general residente en el distrito madrileño de Puente de Vallecas. La temática versaba sobre la situación social y condiciones de vida de la población del distrito. Fue realizada con el apoyo del tejido asociativo del distrito y de captadores voluntarios.

Se utilizó un cuestionario totalmente cerrado, anónimo y muy sencillo. Principalmente fue cumplimentado *online*, aunque se distribuyó también en papel para cubrir las cuotas de personas mayores y poco familiarizadas con ordenadores o *smartphones*. Se estima que

un 70 por ciento fueron rellenas *online* y el resto en papel.

Se controlaron por cuotas proporcionales a la población las variables edad, estudios y barrios. Se partió de unas cuotas por barrios (6) proporcionales a la población y en cada una se requería una proporción de cuotas de edad (3 intervalos) y estudios (3 intervalos), sin cruzar, también proporcionales a la población dentro de cada barrio. A quienes cumplían cuotas se les pedía que cumplimentaran el cuestionario preferentemente de forma *online*.

El cumplimiento de estas cuotas fue aproximado, no exacto. Los captadores incumplieron las cuotas cuando apremiamos para finalizar el campo, aunque permitieron obtener una muestra que podría ser corregida por ponderación. Tras la comparación de los resultados de la muestra con los datos del universo, tanto en las variables controladas *a priori* como en otras no controladas, se procedió a un equilibrage aplicando una ponderación por cuotas directas, mediante procedimiento iterativo. Las variables a equilibrar fueron: edad (3 intervalos), estudios (4), barrios (6), y voto en las pasadas elecciones europeas (7). Los resultados obtenidos tras la ponderación fueron muy satisfactorios, tanto por el buen ajuste de las variables ponderadas como por el comportamiento de aquellas variables no controladas en la muestra cuyos parámetros poblacionales resultaban conocidos. El comportamiento de dichas variables no controladas pero con valores conocidos es fundamental para determinar la calidad de los resultados obtenidos, en términos de precisión y fiabilidad, dado que permiten inferir con precisión las variables cuyos valores son desconocidos. Se obtuvieron los siguientes resultados en las variables no controladas:

Parámetros conocidos de variables no controladas (ni *a priori* ni en la ponderación): comparación entre muestra y universo:

- Viviendas con ascensor: 43 por ciento en la muestra y 45 por ciento en el *Censo de viviendas* de 2011.
- Población extranjera: 16,5 por ciento por ciento en la muestra y 15,4 por ciento en el *Padrón de habitantes* de 2013.

- Mujeres: 50,3 por ciento en la muestra y 52,6 por ciento en el *Padrón* de 2013.
- Viviendas compradas: 41 por ciento en la muestra y 43 por ciento en el *Censo de viviendas* de 2011.
- Viviendas pagando hipoteca: 27 por ciento en la muestra y 29 por ciento en el *Censo de viviendas* de 2011.
- Estado 'malo o deficiente' del edificio: 10 por ciento en la muestra y 11 por ciento en el *Censo de viviendas* de 2011.
- Media de tiempo que tardan a sus centros de trabajo: 36 minutos en la muestra y 36 minutos en estudio *¿Cuánto cuesta ir al trabajo?* de la Fundación La Caixa, 2008.

La diferencia promedio de estos siete parámetros, en valor absoluto, es de +/-1,47 por ciento entre la muestra y el universo, una diferencia realmente pequeña, muy por debajo del margen de error aleatorio para una muestra aleatoria del mismo tamaño (+/- 4,98 por ciento). Los buenos resultados obtenidos de la muestra ponderada (siempre muy por debajo de los límites de error aleatorio) hacen que esta muestra se comporte en la práctica como una muestra probabilística bien ejecutada, haciendo totalmente aceptable la inferencia poblacional de los datos obtenidos no conocidos, dentro del margen de error que le corresponde por su tamaño muestral, ya que esperamos que se comporten como aquellos que sí son conocidos.

Este ejemplo revela que, en determinadas condiciones, una encuesta *online* por cuotas, barata y recopilada con pocos recursos, mediante un cierto control de cuotas y posterior reequilibrage, puede alcanzar resultados próximos a los de una muestra reclutada mediante rigurosos métodos probabilísticos. De este modo, uno de los principales puntos débiles de las encuestas *online* presenta posibilidades de mejora. De avanzar por este camino, las encuestas *online* podrían acabar desplazando a las encuestas telefónicas también en los tipos de muestreo que buscan la similaridad distributiva con su universo, de la misma manera que las encuestas telefónicas desplazaron hace unas décadas a las personales en hogares.

6. CONCLUSIONES

En este artículo se ha puesto de relieve el fuerte desarrollo que están experimentando las encuestas *online*, junto con las ventajas relacionadas con dicha expansión. También se han presentado sus principales características metodológicas, tanto las referidas a sus posibilidades como a sus limitaciones y desafíos.

Entre las ventajas de su aplicación y principales causas de éxito deben considerarse fundamentalmente su menor coste, su facilidad de uso, la simplificación de la logística del trabajo de campo y la posibilidad de alcanzar segmentos de población difícilmente localizables a través de encuestas personales y telefónicas.

No obstante, su aplicación rigurosa no debe obviar algunos de sus inconvenientes, puntos débiles y principales desafíos, entre los que cabe mencionar la frecuente dificultad o imposibilidad de disponer de un adecuado marco muestral, la exclusión de colectivos no digitalizados, las bajas tasas de respuesta total, o las dificultades asociadas a la cumplimentación de los cuestionarios.

Se han destacado asimismo algunos de los aspectos diferenciales de las encuestas *online* y la posibilidad de abordar diseños metodológicos mixtos que aúnan distintos tipos de encuestas.

Algunas de estas cuestiones redundan en la dificultad o imposibilidad de diseñar muestras probabilísticas y plantean el desafío de cómo superar estas limitaciones a través del diseño de un muestreo por cuotas con ajustes *a posteriori* para obtener muestras parecidas a las que se obtendrían con muestreos probabilísticos, que requieren la similaridad distributiva del universo respecto de la muestra. El acceso cada vez mayor a los medios digitales de la mayoría de la población sin duda favorecerá que esta modalidad de encuesta llegue cada vez a públicos más amplios, con lo que aumentará no solo la aplicabilidad de la encuesta sino también las posibilidades de obtener muestras de mayor calidad.

El futuro de la investigación *online* está abierto. No es en absoluto descartable que los avances en inteligencia artificial, que ya se están

aplicando a la formación *online*, supongan una verdadera revolución en el caso de las encuestas: no resulta difícil imaginar cuestionarios *online* interactivos que, a partir del análisis de variables sociodemográficas de identificación y de los metadatos que se producen en el proceso de recopilación de la información, trabajen con algoritmos matemáticos que permitan adaptar las preguntas al perfil de la persona que responda. De hecho, Kaczmirek (2008) ya planteaba hace diez años el desarrollo de un algoritmo para minimizar las tasas de abandono del cuestionario con indicadores de progreso que se activaran en función de distintas variables consideradas.

En lo relativo al análisis de los datos recopilados, tampoco resulta muy lejano vislumbrar un escenario en el que las técnicas de *big data* permitan integrar las respuestas obtenidas a través de la encuesta *online* con información sobre los distintos sujetos encuestados que proceda de otras fuentes o bases de datos. La investigación social probablemente sea cada vez más un campo en el que abunden los enfoques transdisciplinarios, en el que los investigadores sociales seguramente habrán de velar para que no se pierda la perspectiva de la justicia social y de los enfoques éticos.

BIBLIOGRAFÍA

AIMC (Asociación para la Investigación de Medios de Comunicación) (2019). *Audiencia de Internet del Estudio General de Medios* (EGM). Disponible en: <http://Internet.aimc.es/index.html#/landing>

ANEIMO (Asociación Nacional de Estudios de Mercado y Opinión Pública) (2018). La investigación a análisis. *Nota de prensa X Encuentro 19* de octubre de 2018. Disponible en: <http://www.aneimo.com/wp-content/uploads/2018/10/NdP-Datos-X-Encuentro-DEFI-ALTA.pdf>

BATINIC, B. *et al.* (eds.) (2002). *Online Social Sciences*. Seattle-Toronto-Bern-Goettingen: Hogrefe & Huber Publishers.

COUTTS, E. y JANN, B. (2008). Sensitive Questions in *Online Surveys*: Experimental Results for the Randomized Response Technique

(RRT) and the Unmatched Count Technique (UCT). *Annual Meeting of the American Sociological Association*. Boston, 1-4 agosto.

DEUTSKENS, E., DE RUYTER, K. y WETZELS, M. (2006). An Assessment of Equivalence between *Online* and Mail Surveys in Service Research. *Journal of Service Research*, 8, pp. 346-355.

DEUTSKENS, E., DE RUYTER, K., WETZELS, M. y OOSTERVELD, P. (2004). Response Rate and Response Quality of Internet-Based Surveys: An Experimental Study. *Marketing Letters*, 15(1), pp. 21-36.

DÍAZ DE RADA, V. (2012). Ventajas e inconvenientes de encuestas por Internet. *Papers*, 97(1), pp. 193-223.

DUDA, M. D. y NOBILE, J. L. (2010). The Fallacy of *Online* Surveys: No Data Are Better than Bad Data. *Human Dimensions of Wildlife*, 15, pp. 55-64.

DUFFY, B., SMITH, K., TERHANINAN, G. y BREMER, J. (2005). Comparing data from *online* and face-to-face surveys. *International Journal of Market Research*, 47(6), pp. 615-639.

ESOMAR (European Society for Opinion and Marketing Research) (2012). *Guía para la investigación online*. World Research Codes and Guidelines. Disponible en: www.esomar.org

EVANS, J. y MATHUR, A. (2005). The value of *online* surveys. *Internet Research*, 15(2), pp. 195-219.

GANASSALI, S. (2008). The Influence of the Design of Web Survey Questionnaires on the Quality of Responses. *Survey Research Methods*, 2(1), pp. 21-32.

GROVES, R. (1989). *Survey Errors and Survey Costs*. Nueva York: Wiley.

KACZMIREK, L. (2008). *Human-Survey Interaction: Usability and Nonresponse in Online Surveys*. Tesis doctoral, Universidad de Mannheim. Disponible en: <https://d-nb.info/992375924/34>

KANTAR MEDIA (2017). *Lightspeed Mobile Case Study: Representative Data: Mobile Design Matters*. Disponible en: <http://www.lightspeedresearch.com/resource-center/case-studies/>

NULTY, D. (2008). The adequacy of response rates to *online* and paper surveys: what can be done?. *Assesment & Evaluation in Higher Education*, 33(3), pp. 301-314.

OCHOA, C. y PORCAR, J. M. (2018). Modeling the effect of quota sampling on *online* fieldwork efficiency: An analysis of the connection between uncertainty and sample usage. *International Journal of Market Research*, 60(5).

PAVÍA, J. M., RAUSELL, P., MARCO-SERRANO, F. y COLL, V. (2011). Encuestas electorales *online*: nuevos retos, viejos problemas. *Revista Española de Investigaciones Sociológicas*, 135, pp. 107-122.

REVILLA, M. y OCHOA, C. (2017). Ideal and maximum length for a web survey. *International Journal of Market Research*, 59(5), 9 págs.

SÁNCHEZ CARRIÓN, J. J. y SEGOVIA, J. M. (2008). La participación y la influencia del recordatorio en las encuestas panel *online* a estudiantes universitarios. *Empiria. Revista de Metodología de Ciencias Sociales*, 16, pp. 135-161.

SZOLNOKI, G. y HOFFMANN, D. (2013). *Online*, face-to-face and telephone surveys: comparing different sampling methods in wine consumer research. *Wine Economics and Policy*, 2, pp. 57-66.

TRUJILLO CARMONA, M. (2019). El futuro de las encuestas en la investigación Social. Working Paper. *XIII Congreso Nacional de Sociología*. Valencia, 3-6 de julio.

WE ARE SOCIAL (2018). *Informe de tendencias digitales y de redes sociales en el mundo*. Disponible en: <https://wearesocial.com/es/digital-2018-espana>

WRIGHT, K. B. (2005). Researching Internet-Based Populations: Advantages and Disadvantages of *Online* Survey Research, *Online Questionnaire Authoring Software Packages, and Web Survey Services*. *Journal of Computer-Mediated Communication* 10(3). Disponible en: <https://doi.org/10.1111/j.1083-6101.2005.tb00259.x>