

La calidad de la encuesta en la sociedad digital: Avances y retrocesos

Survey Quality in Digital Society: Advances and Setbacks

M.^a Ángeles Cea D'Ancona

Palabras clave

Calidad

- Encuesta
- Errores medición
- Errores no observación
- Muestras encuestas online

Key words

Quality

- Survey
- Measurement Errors
- Non-observation Errors
- Online Survey Samples

Resumen

La creciente digitalización de las sociedades desarrolladas contribuye a abaratar los costes de la investigación, aunque no siempre cumpliendo criterios de calidad. Este artículo ofrece una panorámica de avances en metodología de encuesta tras la revisión de recientes investigaciones empíricas publicadas en revistas especializadas. Especial énfasis se pone en la encuesta *online* y los métodos mixtos de encuesta, junto con los handicaps de la recopilación de datos mediante dispositivos móviles. Se detectan déficits de representatividad de las muestras por errores de cobertura, de no respuesta y la aplicación de muestreos no probabilísticos y paneles de voluntarios. También, déficits de representatividad de los datos por errores de medición, a cuya aminoración contribuyen recientes avances tecnológicos. Cada error se acompaña de actuaciones consensuadas para reducir su incidencia en la calidad de la encuesta.

Abstract

The increasing digitalization of developed societies contributes to reduce research costs, although it does not always meet quality criteria. This article offers an overview of advances in survey methodology after reviewing recent empirical research published in specialized journals. Special emphasis is placed on online survey and mixed survey methods, along with the handicaps of data collection using mobile devices. Deficits in the representativeness of the samples were detected due to coverage errors, non-response errors and the application of non-probabilistic sampling and volunteer panels. Also, there are deficits in the representativeness of the data due to measurement errors, to which recent technological advances contribute to their reduction. Each error is accompanied by consensual actions to reduce its impact on the quality of the survey.

Cómo citar

Cea D'Ancona, M.^a Ángeles (2025). «La calidad de la encuesta en la sociedad digital: Avances y retrocesos». *Revista Española de Investigaciones Sociológicas*, 191: 25-42. (doi: 10.5477/cis/reis.191.25-42)

La versión en inglés de este artículo puede consultarse en <http://reis.cis.es>

M.^a Ángeles Cea D'Ancona: Universidad Complutense de Madrid | maceada@ucm.es



INTRODUCCIÓN

Casi dos décadas después de la publicación del artículo «La senda tortuosa de la “calidad” de la encuesta» (*REIS*, 111), y más de la publicación de dos monografías comprehensivas de los errores de encuesta (Groves, 1989; Biemer y Lyberg, 2003), procede su actualización en un contexto de decrecimiento de la encuesta presencial y de expansión de la encuesta *online* y los métodos *mixtos* de encuesta, aunque en 2003 Biemer y Lyberg afirmaban que estos eran «la norma hoy en día» (p. 208). Principalmente, encuestas panel que comienzan siendo presenciales y después telefónicas o autocumplimentadas. El objetivo es reducir los costes económicos y temporales de la investigación aplicando métodos más baratos que la encuesta *cara-a-cara*, además de solventar errores de *no respuesta*, *cobertura* y *medición*, combinando marcos muestrales, procedimientos de selección de la muestra y administración del cuestionario. Una ilustración reciente la ofrece la Encuesta Social Europea (ESS), cuya Ronda 12 será *mixta* en 2025/2026: la mitad de la muestra se encuestará presencialmente y la otra mitad autocumplimentará el cuestionario *online* y por correo. La Ronda 13 (2027/2028) pasará a ser únicamente *online*, dependiendo del impacto de la Ronda 12 en sus series temporales.

Por otra parte, de acuerdo con ESOMAR (2023), el mundo digital se ha convertido en la primera fuente de recogida de información. En España, el 70 % de los estudios realizados en 2022 (y analizados por Insights+Analytics) fueron cuantitativos, siendo el 28 % de toma de información por medios electrónicos, el 31 % cuantitativos *online*/móvil, el 7 % por teléfono y apenas un 4 % presencial; de ellos, el 37 % fue de paneles. Unas proporciones similares a lo compilado a nivel mundial, con un 35 % de estudios cuantitativos *online*/móvil. Una expansión que no habría sido posible sin los

avances tecnológicos, que han facilitado el acceso a Internet y a dispositivos móviles incluso a la población de más edad y menor posición socioeconómica.

Este artículo desgana avances en metodología de encuesta y su traducción en mejoras de *calidad*, a partir de los resultados de investigaciones empíricas publicadas durante la última década en revistas científicas especializadas en encuesta. La mirada panorámica comienza por las potencialidades y límites de la encuesta *online* y los métodos *mixtos*, para después pasar al detalle de los diversos errores de encuesta dentro del marco teórico del error total de encuesta.

POTENCIALIDADES Y LÍMITES DE LA ENCUESTA ONLINE Y LOS MÉTODOS MIXTOS DE ENCUESTA EN LA INVESTIGACIÓN SOCIAL

La expansión de las encuestas web u *online* en gran parte se explica por su menor coste, que posibilita ampliar el tamaño y dispersión de la muestra, además de su seguimiento como estrategia para reducir la *no respuesta* de unidad y la aplicación de estudios panel. Acompaña la relativa rapidez de transmisión de datos, al almacenarse directamente en formato electrónico, al igual que las modalidades automatizadas de encuesta telefónica (CATI) y presencial (CAPI). Esto ayuda a reducir errores de codificación y acelera la preparación y análisis de los datos de encuesta. A estas potencialidades se añade la mayor flexibilidad del diseño del cuestionario (formatos de preguntas y respuestas), por el uso de *software* multimedia, y que la encuesta sea autocumplimentada. La autocumplimentación se asocia a la reducción del sesgo de *deseabilidad social*, al no darse la respuesta a un entrevistador/a en persona o por teléfono (Fricker *et al.*, 2005; Gooch y Vavreck,

2019). Pero su ausencia puede revertir negativamente en la obtención de respuestas de menor calidad, más precipitadas (mayor sesgo de *primacía*¹ y de no respuesta de ítem), porque no media una persona que motive a responder, aclare dudas y realice el seguimiento de la recogida de información (Cernat y Revilla, 2020; Heerwegh y Loosveldt, 2008).

Que la encuesta autocumplimentada *online* (CAWI: *Computer Assisted Web Interviewing*) pueda responderse mediante un dispositivo móvil (teléfono inteligente o tableta) añade otras posibilidades y límites. Permite integrar respuestas sobre actitudes y conductas con datos concretos de comportamiento recopilados de modo pasivo a través de sensores (ubicaciones GPS, acelerómetros, dispositivos para medir la actividad física, estrés...). Una integración de datos subjetivos con objetivos que mejora la medición de comportamientos, al proporcionar datos menos susceptibles a errores de recuerdo y deseabilidad social (Keusch y Conrad, 2022; Link *et al.* 2014; Struminskaya *et al.*, 2020). En su contra, que la aceptación de estos métodos de recopilar datos es baja (Wenz y Keusch, 2023). Al igual que sucede con la opción de tomar fotografías y realizar un seguimiento del uso del dispositivo móvil (como páginas web visitadas), que también pueden complementar (e incluso reemplazar) los datos recogidos mediante cuestionario. Su uso precisa que los participantes recuerden utilizarlo para cada evento solicitado. Requiere motivación y compromiso continuos. Además, pueden optar por informar solo de algunas actividades, generando exclusión diferencial de los eventos que se preguntan, junto con limitaciones de almacenamiento, que pueden provocar la pérdida de datos.

El principal hándicap de la encuesta *online* sigue siendo reducir la *no respuesta*

(Elevelt, Lugtig y Toepoel, 2019²; Jäckle *et al.*, 2019; Struminskaya *et al.*, 2021a), que compromete su calidad y posibilidades de inferencia. Este reto se suma a errores de *cobertura* (tratados en el siguiente apartado) y los límites de cumplimentar la encuesta mediante dispositivo móvil: el cuestionario ha de tener un formato que facilite la respuesta, cuando se pulsa una pantalla táctil pequeña, y que se autorice el uso de sensores instalados en el dispositivo.

Mientras que responder a las preguntas del cuestionario o tomar fotografías permite controlar la información que se aporta, en otras actividades (como la ubicación por GPS) el único control es poder desconectar la recopilación de datos por deseo de privacidad. Estudios sobre la voluntad de realizar tareas adicionales en un dispositivo móvil como parte de una encuesta concluyen que la predisposición es mayor para tareas donde pueda controlarse el contenido que se transmite (como fotografías) frente a aquellas que recopilan datos automáticamente (caso de la ubicación GPS) (Revilla, Couper y Ochoa, 2019; Revilla *et al.*, 2016³; Wenz, Jäckle y Couper, 2019; Wenz y Keusch, 2023⁴). Asimismo, se observa que las personas que utilizan su dispositivo con mayor intensidad (medido por la frecuencia de descarga de aplicaciones y el número de aplicacio-

² En su Estudio sobre el Uso del tiempo, el 43 % de los miembros del panel respondieron positivamente a la invitación para participar en la versión para teléfonos inteligentes, y apenas el 29 % completaron todas las etapas del estudio, siendo su perfil sociodemográfico diferente de quienes no participaron en alguna de las tareas, como registrar datos de GPS.

³ Su estudio muestra que la disposición a utilizar el GPS varía por países: desde el 30 % de los encuestados en México al 17 % en Portugal; en España, el 24 %.

⁴ Guiados por el Modelo de Aceptación de la Tecnología, examinaron la predisposición a descargar una aplicación en un *smartphone* en 1876 miembros del NORC AmeriSpeak Panel. Constataron que la predisposición aumentaba en estudios donde podían controlar la recopilación de datos: poder desactivarlo temporalmente o revisar los datos antes de enviarlos.

¹ La aleatorización de las opciones de respuesta convierte el sesgo en aleatorio y no sistemático.

nes utilizadas) están más predispuestas a participar en tareas de recopilar datos móviles que aquellas preocupadas por la privacidad y seguridad de los datos que aportan. La participación también se ve afectada por la organización que patrocine el estudio y su duración, como en otros métodos de encuesta, favoreciendo a los que duren menos tiempo y los patrocinados por universidades (Struminskaya *et al.*, 2021b). Que la persona pueda elegir el modo (voz, texto, vídeo) de responder a la encuesta igualmente favorece la participación, al aumentar la satisfacción con la encuesta (Conrad *et al.*, 2017).

Respecto a los métodos *mixtos* de encuesta (autocumplimentada *online* con presencial o telefónica), tras su expansión igualmente se hallan razones económicas y de incrementar la tasa de respuesta en estudios transversales y *panel*. Favorece abarcar un conjunto más diverso de población, reduciendo errores de *cobertura* y *no respuesta*, que merman la representatividad del estudio (Cornesse y Bosnjak, 2018; Jäckle, Lynn y Burton, 2015; Lugtig *et al.*, 2011). Algunas encuestas ofrecen la opción de elegir el modo preferido para ser encuestado desde el principio, mientras que en otras se asigna dependiendo de las propensiones de respuesta de cada grupo de población durante el trabajo de campo. Este último enfoque tiene la ventaja de aplicar el de mayor probabilidad de respuesta en esa población (Cornesse y Bosnjak, 2018). No obstante, la tasa de respuesta de las encuestas *online* sigue jugando en su contra, incluso cuando se pregunta a profesionales con pleno acceso a Internet y elevado nivel educativo (Cea D'Ancona y Valles, 2021)⁵.

⁵ En su encuesta a 7989 docentes e investigadores de universidades públicas y privadas españolas aleatoriamente seleccionados para el proyecto MEDIM II (CSO2016-75946-R), 1667 acabaron cumplimentándola después de seis recordatorios.

AVANCES EN EL CUMPLIMIENTO DE CRITERIOS DE CALIDAD EN EL MARCO DEL ERROR TOTAL DE ENCUESTA

Los modelos de regresión logística y discriminante obtenidos en la Encuesta de *Percepción social de las encuestas* (III), realizada por el Centro de Investigaciones Sociológicas (CIS) en 2017 (Cea D'Ancona, 2022), muestran que la *confianza* en la encuesta depende de la utilidad que se la atribuya. Una utilidad que se conecta a la representatividad de la *muestra* y la *validez* de los datos que aporte, además de su consideración de beneficiosa para las personas. De sus resultados se colige que el grado de cumplimiento de requisitos de *calidad* puede determinar la participación en una encuesta, en función de la fiabilidad que se atribuya a los datos que aporta. Pero ¿qué determina la *calidad* de la encuesta? Aunque exista consenso en considerar que una baja tasa de respuesta disminuye su calidad, que la tasa sea alta no es sinónimo de calidad (Eckman y Koch, 2019), al depender de errores varios.

Cuando se evalúa la *calidad* de la encuesta, el marco teórico de referencia es el *error total* de la encuesta que, como indica Lyberg (2012), permite optimizar la encuesta, minimizando el tamaño acumulado de todas las fuentes de error, dadas las limitaciones presupuestarias. Comprende distintas fuentes de error que contribuyen a que las estimaciones de la encuesta se desvíen de los valores reales (Groves, 1989; Groves y Lyberg, 2010; Lyberg y Stukel, 2017). Incluye errores de *no observación*, que inciden en la selección de la *muestra* que se analiza: errores de *cobertura* (no se cubre toda la población de estudio), de *muestreo* (la muestra no representa a la población), y *no respuesta* (de unidad o de ítem). Aun siendo estos los errores más analizados, habiendo fórmulas especí-

ficas que cuantifican su incidencia (Groves, 1989), el análisis completo de la *calidad* también abarca errores de *medición*, observación o respuesta. Estos últimos atienden a la representatividad de la información que la encuesta proporciona. Afecta el método de encuesta aplicado, especialmente cuando se abordan temas susceptibles al sesgo de *deseabilidad social* (Cea D'Ancona, 2017; Heerwegh y Loosveldt, 2008; Kreuter, Presser y Tourangeau, 2008; Zhang *et al.*, 2017). También, el diseño del cuestionario, la mediación del entrevistador/a (cuando haya), la actitud de la persona encuestada y el tratamiento de los datos de encuesta (edición, codificación de preguntas abiertas, grabación, ponderación, imputación, tabulación, modelado estadístico...). Por ello, los avances sobre calidad de la encuesta comprenden ambas agrupaciones de errores.

Incidencia de los errores de *no observación* en la representatividad de la muestra

A la creciente aplicación de encuestas *online* acompaña el debate sobre la representatividad de la *muestra* que acaba cumplimentándolas. No solo respecto a errores de *cobertura*, por no disponer de marcos muestrales de usuarios de Internet que posibiliten aplicar muestreos probabilísticos en encuestas a la población general (aun aplicando *web-push*). También son objeto de debate la representatividad de muestras no probabilísticas y las bajas tasas de respuesta.

Existen errores de *cobertura* cuando ciertas unidades de la población de interés no tienen oportunidad de ser encuestadas, al no estar incluidas en el marco muestral (Groves *et al.*, 2009). Afecta la proporción de la población no cubierta y sus diferencias con quienes sí están; sobre todo si están relacionadas con el tema de la encuesta. Aunque se esté detectando una

disminución significativa en la sobrerrepresentación de personas de alto nivel educativo en la población que accede a Internet (Sterret *et al.*, 2017), que se necesite un cierto nivel de competencia para cumplimentar cuestionarios *online* puede afectar negativamente a la participación de personas de menor nivel educativo e interés en el tema de la encuesta: la «brecha digital» en el acceso y uso diferencial de las nuevas tecnologías. Por ello, el error de *cobertura* sigue siendo la fuente que más atenúa la representatividad de encuestas *online* dirigidas a la población general, aunque igualmente existan en otras encuestas (como las telefónicas con muestreo exclusivo de teléfonos fijos o móviles). Este error se incrementa en encuestas *online* mediante dispositivo móvil. Implica disponer del dispositivo, de la capacidad de utilizarlo para la tarea solicitada, y la voluntad de dar consentimiento para compartir datos (Antoun *et al.*, 2019; Couper *et al.*, 2018; Keusch *et al.*, 2023; Keusch *et al.*, 2019; Wenz, Jäckle y Couper, 2019). Su reducción supondría proporcionar acceso a dispositivos móviles, conexión a internet móvil, y acompañamiento en el proceso de autocumplimentar la encuesta.

Que la encuesta *online* se cumplimente con paneles de voluntarios añade el debate sobre la representatividad de muestras seleccionadas mediante métodos *no probabilísticos*. Especialmente, cuando las personas se reclutan a sí mismas respondiendo a anuncios de encuestas; una forma habitual en encuestas no probabilísticas (Callegaro *et al.*, 2014; Cornesse *et al.*, 2020). Anuncios que suelen atraer a personas de un perfil sociodemográfico, valores y hábitos específicos, que pueden estar participando simultáneamente en varios paneles *online* (Tourangeau, Conrad y Couper, 2013). Se distinguen por su mayor conocimiento político y preferencia de partidos y políticas de centroizquierda (Karp y Luehiste, 2015; Valentino *et al.*, 2020), y la menor presen-

cia de mayores de sesenta y cinco años (Loosveldt y Sonck, 2008). Ello deteriora la representatividad de la *muestra* y la obtención de estimaciones sesgadas (Bethlehem, 2010; Chang y Krosnick, 2009; Cornesse y Bosnjak, 2018; Wang *et al.*, 2015). Además, se observa que las personas que participan activamente en varios *paneles* pueden incluso proporcionar datos erróneos para aumentar sus recompensas económicas (Toepoel, Das y Soest, 2008; Cornesse y Bosnjak, 2018).

Mientras el muestreo *probabilístico* posibilita estimar la precisión de las estimaciones muestrales, con intervalos de confianza y márgenes de error (Kish, 1965), el *no probabilístico* (en su mayoría muestras de *conveniencia*) no calibra la concurrencia de errores de cada etapa del diseño muestral. Únicamente se comprueba lo cerca que se halla la muestra final de la población de estudio en características concretas. Al igual que sucede con las muestras *no probabilísticas* por *cuotas*, diseñadas para garantizar que la muestra coincida con la población en parámetros demográficos clave. En la medida en que esto sea cierto, las inferencias hechas a partir de muestras por *cuotas* serán precisas (Cornesse *et al.*, 2020).

Pero incluso las muestras probabilísticas pueden ser inexactas por variaciones en la probabilidad de que ciertos grupos de la población acaben participando en la encuesta, habiendo una *no respuesta* sistemática (no aleatoria). Los ajustes estadísticos que se utilizan para reducir sesgos sistemáticos en muestras probabilísticas igualmente se aplican en muestras no probabilísticas, en ajustes globales y por resultados específicos. Entre ellos destaca la ponderación del puntaje de propensión (*propensity score weighting*), que se aplica concluida la recogida de datos de encuesta y tomando como referencia los datos de la población o de una muestra probabilística grande. Generalmente, se parte de un modelo de regresión logística utilizando varia-

bles demográficas, de comportamiento y actitud, medidas en ambos conjuntos de datos para predecir la probabilidad de que una unidad particular pertenezca a la muestra *no probabilística*. Esta se pondera utilizando la inversa de la probabilidad predicha derivada de estos modelos de propensión (Lee, 2006; Valliant y Dever, 2011). En cambio, el *emparejamiento de muestras* intenta componer una muestra no probabilística equilibrada seleccionando unidades de un marco muy grande (como la lista de miembros de un panel de participación voluntaria), basándose en una serie de características que coincidan con las correspondientes a las unidades de la muestra probabilística de referencia (Bethlehem, 2016). El procedimiento de comparación se basa en una métrica de distancia (como la *euclídea*) para identificar la coincidencia más cercana entre pares de unidades, basándose en el conjunto de características comunes. Que el emparejamiento se haga antes de comenzar la encuesta intenta reducir las diferencias entre muestra no probabilística y población en variables clave. A diferencia de la *ponderación de propensión*, el emparejamiento (*matching*) de muestras no es una técnica de ponderación explícita, sino un método que intenta equilibrar la muestra no probabilística. Si bien, en ambos casos no hay garantía de que los sesgos en muestras no probabilísticas se eliminen por completo (Cornesse *et al.*, 2020; Little *et al.*, 2020).

Tras revisar la evidencia empírica disponible, Cornesse *et al.* (2020) insisten en la recomendación de continuar confiando en encuestas de muestreo probabilístico. Al igual que después harían Lavrakas *et al.* (2022), tras comparar paneles en línea administrando el mismo cuestionario en ocho muestras nacionales independientes. Asimismo, recomiendan una mayor transparencia por parte de las empresas demoscópicas. Disponer de informes que describan la metodología utilizada para re-

copilar y manipular los datos se considera de suma importancia para poder deducir si las personas encuestadas realmente representan a la población de la que provienen.

Respecto a los errores de *no respuesta* (de unidad y de ítem), se reitera que la tasa de respuesta solo está débilmente asociada a dicho error (Groves *et al.*, 2008; Groves y Peytcheva, 2008). Tasas de respuesta relativamente bajas pueden reflejar con precisión la población, si el conjunto de las personas que contesta la encuesta varía aleatoriamente de quienes no responden (Bethlehem, Cobben y Schouten, 2011; Cornesse y Bosnjak, 2018). La incidencia de la *no respuesta* en la calidad de la encuesta depende de los perfiles de quienes responden, su conexión con el tema abordado, el interés que suscite en la población a encuestar (Groves, Presser y Dipko, 2004; Keusch, 2013) y su sensibilidad (Couper *et al.*, 2010; Tourangeau y Yan, 2007). Las encuestas que abordan conductas especialmente estigmatizadas suelen ser menos respondidas por quienes más participan en dichas conductas, menoscabando su representatividad (Plutzer, 2019).

De las encuestas telefónicas se reitera que sobrerrepresentan a las personas de más *edad*, que en cambio están subrepresentadas en las encuestas *online*, al igual que las personas de menor posición socioeconómica (Bech y Kristensen, 2009; Couper, 2000; Roster *et al.*, 2004). Por otra parte, las más implicadas con la comunidad suelen ser más participativas porque perciben la encuesta como bien público y su participación como un comportamiento prosocial, que beneficia a la sociedad en su conjunto (Beller y Geyer, 2021; Groves, Singer y Corning, 2000). Concuera con la conclusión de que los valores altruistas predicen la participación en la encuesta (Groves, Cialdini y Couper, 1992). A ella también contribuye la estructura de la encuesta (tipo y formato de las preguntas del cuestionario), junto con las garantías de privacidad y confidencialidad que se den.

Si la encuesta es *online*, en la probabilidad de participar se suma la capacidad física (visión, destreza para responder) y familiaridad con dispositivos digitales, además del tipo de datos que se quiera recopilar porque el requisito de descarga de una aplicación reduce la voluntad de participar en la encuesta (Jäckle *et al.*, 2019; Wenz, Jäckle y Couper, 2019). Para evitarlo se recomienda proporcionar instrucciones adicionales o capturas de pantalla sobre cómo acceder a la tienda de aplicaciones, descargarla e instalarla en el dispositivo. Cuando la persona no esté lo suficientemente familiarizada o utilice el ordenador o dispositivo con menor intensidad, se aconseja que un entrevistador/a le ofrezca asistencia mediante la creación de una línea directa de soporte. Y, para que la seguridad no preocupe, que la carta de invitación informe de las garantías de confidencialidad y resalte la importancia de participar en la encuesta. También, que se envíen recordatorios adicionales a poblaciones sobreencuestadas y estudios *panel* (Struminskaya *et al.*, 2021b; Mol, 2017). Los recordatorios por SMS muestran mayor efectividad para aumentar la tasa de respuesta, al llamar más la atención y ser eficaces para establecer legitimidad (Andreadis, 2020; Kocar, 2022).

Respecto al cuestionario, la duración de la encuesta, la dificultad de las preguntas, el contenido de la primera pregunta y el uso de la barra de progreso (en encuestas *online*) igualmente se relacionan con la tasa de respuesta (Liu y Wronski, 2018). También parece afectar la época del año (mejor septiembre e invierno) y el día de la semana (lunes, seguido del martes), en contra del sábado y el domingo, cuando es menos probable que se cumplimente una encuesta *online* (Fang *et al.*, 2021). Se pospone al lunes, por obligaciones familiares o domésticas y la necesidad de desconectar de actividades que sean cognitivamente exigentes.

Por las repercusiones de la *no respuesta* en la calidad de la encuesta, agrandando

otros errores de representatividad de la muestra, se recomienda reducir su incidencia mediante actuaciones diversas. Entre ellas, ofrecer incentivos en encuestas *online* (Becker, Möser y Glauser 2019; Göritz, 2006), como en otros métodos, y la opción de revisar y eliminar los datos que no deseen transmitirse al investigador (Wenz y Keusch, 2023). También, que la carta de invitación adjunte el enlace a una tienda de aplicaciones (Lawes *et al.*, 2022) para facilitar la descarga con seguridad.

Cuando los datos ya han sido recogidos, se aplican ajustes estadísticos para reducir la incidencia negativa de la *no respuesta*, como en los otros errores de no observación. Entre ellos, *ponderaciones* que corrijan las diferencias sociodemográficas entre muestra final y población. Su efectividad dependerá de lo relacionadas que estén las variables elegidas con el tema de la encuesta, la propensión a responderla y la calidad de los datos disponibles: datos estadísticos de la población (censo, padrón de habitantes...), administrativos (si es posible vincular registros) y de fuentes comerciales que contengan características de vecindarios y unidades de vivienda (West *et al.*, 2015).

Una fuente de información auxiliar son las *observaciones* del entrevistador/a sobre las características de la población encuestada, cuando median en la recogida de datos. En comparación con las características disponibles a nivel de área, las que proporcione sobre la vivienda puede aportar información de interés para la encuesta y los ajustes de ponderación. En su contra, que estas observaciones pueden variar mucho entre observadores y carecer de la calidad necesaria. Su aplicación requeriría más formación de los entrevistadores/as y que acompañen sus respuestas con fotografías, que se revisarían grupalmente para reducir la variación entre los entrevistadores/as del estudio (Ren *et al.*, 2022). Por otra parte, ha de considerarse que las observaciones

suelen captar variables de clasificación observables, que no siempre son claves en la encuesta. Por lo que su utilidad para reducir el error de *no respuesta* dependerá de lo relacionadas que estén con el tema de la encuesta y que no constituyan juicios de valor. En caso de observaciones virtuales a través de *Google Street View*, se destaca su susceptibilidad a problemas de cobertura (menor en áreas no urbanas) y al desfase temporal con la fecha de toma de las imágenes (Vercruyssen y Loosveldt, 2017).

Repercusión de los errores de *medición* en la representatividad de la información

Los errores de observación o *medición* constituyen desviaciones de las respuestas de los valores reales (Groves, 1989; Couper, 2000). Su tamaño se ve afectado por las decisiones que se tomen al diseñar la encuesta, desde la elección del método hasta la formulación exacta de preguntas y respuestas, con efectos en los resultados de la encuesta y la obtención de conclusiones erróneas (Saris y Revilla, 2016). En un estudio reciente, Poses *et al.* (2021) cuantificaron la calidad de medición promedio en 0,65 para sesenta y siete preguntas de la Encuesta Social Europea en hasta cuarenta y un grupos lingüísticos de países. El 65 % de la varianza observada provino de los conceptos de interés latentes, mientras que el 35 % se debió a errores de medición. Con anterioridad, DeCastellarnau y Revilla (2017) encontraron estimaciones de la calidad de la medición entre 0,60 y 0,89 para las preguntas de la quinta ola del panel en línea de *Ciudadanos Noruegos*.

Respecto al proceso pregunta-respuesta, Tourangeau, Rips y Rasinski (2000) sugieren que la calidad de la respuesta depende de la minuciosidad de cuatro pasos cognitivos: comprensión de la pregunta, recuperación de información relevante de la memoria, formulación de un juicio y se-

lección de una respuesta. Los efectos sesgados, que llevan a respuestas no reales, suelen denominarse *efectos de respuesta*. Se habla de respuestas aleatorias, desatendidas o de esfuerzo insuficiente para responder (Maniaci y Rogge, 2014). Independientemente del contenido de la pregunta, incluyen el sesgo de *aquiescencia* (o tendencia al acuerdo indistintamente de lo que se pregunte), de *primacía* (seleccionar la primera opción de respuesta razonable) y *recencia* (elegir la última), al no dedicar esfuerzo al proceso de respuesta. Mientras que el sesgo de *recencia* está más presente en la encuesta telefónica, el de *primacía* se da más en las autocumplimentadas (Christian, Dillman y Smyth, 2007).

A ellos se suman los errores debidos al orden de la pregunta en el cuestionario y su contenido (preguntas referidas al pasado y vulnerables al sesgo de *deseabilidad social*). Afecta el interés en el tema de la encuesta (Anduiza y Galais, 2016) y el nivel educativo de la persona encuestada. En general, los errores de *medición* son más frecuentes en personas de menor nivel educativo. La excepción, el sesgo de *deseabilidad social*, más habitual en aquellas de mayor nivel educativo, que más perciben la intencionalidad de la pregunta, habiendo una respuesta diferencial en función del método de encuesta aplicado, a favor de las autocumplimentadas (Cea D'Ancona, 2017; Chang y Krosnick, 2009; Heerwegh y Loosveldt, 2008; Kreuter, Presser y Tourangeau, 2008; Zhang *et al.*, 2017).

Dada su complejidad y los diferentes factores que intervienen en la respuesta, aunque se haya estudiado cómo minimizarlos y diseñar mejor las encuestas (Callegaro, Manfreda y Vehovar, 2015; Couper, 2008; Tourangeau, Conrad y Couper, 2013), los errores de *medición* siguen siendo difíciles de controlar. Aunque las encuestas *online* sean rentables, rápidas y fáciles de implementar, la calidad de los datos (en términos de *medición*) se ve

perjudicada cuando las preguntas se responden al azar o con baja motivación para interpretar correctamente su contenido y cumplir las instrucciones de la encuesta. Lo que motiva que se cuestione la *calidad* de sus mediciones, al carecer del control de las entrevistas presenciales y su mayor vulnerabilidad al sesgo de *aquiescencia* (Fricker *et al.*, 2005; Zhang y Conrad, 2014), junto con errores propiciados por escribir usando las yemas de los dedos en un pequeño teclado virtual (cuando se responde mediante dispositivo móvil). Lo que lleva a considerar la longitud de las respuestas a preguntas abiertas como un indicador de satisfacción (Mavletova y Couper, 2013).

Como encuesta autocumplimentada, se destaca el potencial de la encuesta *online* para elegir cuándo responder al cuestionario y poder verificar información relevante antes de cumplimentarlo. Que cree menos presión (que la encuesta telefónica) para dar una respuesta rápida propicia respuestas más precisas a preguntas de conocimiento y referidas al pasado (Braunsberger, Wybenga y Gates, 2007; Fricker *et al.*, 2005). Al igual que favorece el informar de opiniones o comportamientos socialmente indeseables, a diferencia de las encuestas telefónicas, más vulnerables a la *deseabilidad social* (Chang y Krosnick, 2009; Christian, Dillman y Smyth, 2007; Kreuter, Presser y Tourangeau, 2008). En cambio, la encuesta presencial favorece una mejor relación encuestador-encuestado y la validación de la legitimidad de la encuesta (Jäckle, Roberts y Lynn, 2010). De ahí que sea menos vulnerable al sesgo de *deseabilidad social* que la telefónica (Hope *et al.*, 2022).

Entre las actuaciones para reducir respuestas descuidadas (menos habituales en mujeres y personas de mayor nivel educativo) se recomienda utilizar ítems que verifiquen si se presta atención al responder las preguntas del cuestionario (Berinsky, Margolis y Sances, 2014). Especialmente en encuestas *online*, más dependientes del

diseño del cuestionario, al no mediar un entrevistador/a que aclare las preguntas y motive a responderlas. Que incluyan estímulos visuales favorece la motivación y participación en la encuesta porque resultan más divertidas o amenas (Bărbulescu y Cernat, 2012; Liu *et al.*, 2015; Mavletova, 2015). Otras mejoras de diseño específicas para dispositivos móviles son incluir herramientas de entrada fáciles de utilizar y evitar formatos que lo dificulte (controles deslizantes, cuadros desplegables que se conviertan en selectores), además de aplicar *Responsive Web Design* para adaptar el cuestionario a diferentes tamaños de pantalla (Antoun, Couper y Conrad, 2017).

En las encuestas presenciales y telefónicas, la actuación del entrevistador/a puede añadir errores de *medición*. Aunque contribuya a reducir la dificultad de la tarea, al disminuir las exigencias cognitivas de la pregunta (ofreciendo cualquier aclaración sobre preguntas-respuestas), pueden introducir errores al formular preguntas y anotar respuestas (West y Blom, 2017). Al igual que se observa que las (des)coincidencias sociodemográficas entre la persona que entrevista y la que responde pueden afectar a la *no respuesta* de unidad e ítem en encuestas presenciales (Bittman, 2020; Durrant *et al.*, 2010). El principal marco teórico que lo explica es la teoría de la *vinculación o conexión* (Groves, Cialdini y Couper, 1992). Sugiere que se prefiere interactuar con personas que agradan por sus características sociodemográficas, actitudes o creencias. Se observa que el emparejamiento por género, edad, nivel educativo y color de piel aumenta la cooperación y participación en la encuesta (Blanchard, 2022; Durrant *et al.*, 2010; Vercruyssen, Wuyts y Loosveldt, 2017). En cambio, la teoría de la *distancia social* defiende que demasiada distancia (en términos sociodemográficos) entre entrevistador-encuestado producirá respuestas sesgadas (Dohrenwend, Colombotos y Dohrenwend, 1968). En en-

cuestas *panel* se observa que mantener el mismo entrevistador/a propicia la confianza de la persona encuestada y la sinceridad de su respuesta (Kühne, 2018).

Las observaciones que anote tras la entrevista pueden utilizarse como indicadores de calidad de las respuestas, incluido el grado de comprensión y cooperación de quien responde al cuestionario (como en la Encuesta Social Europea o las realizadas por el CIS). Ayudan a identificar posibles fallos en la calidad de los datos. Si bien, dichas observaciones generalmente se centran en errores de *no respuesta*, de utilidad para ajustar la no respuesta de unidad y el desgaste del panel (West, Kreuter y Trappmann, 2014). No obstante, están sujetas a efectos de varianza del entrevistador/a y errores de *medición* (Sinibaldi, Durrant y Kreuter, 2013), como antes se expuso. Igualmente, ha de considerarse la desviación que puedan introducir los entrevistadores en la selección de casos (unidades muestrales). Como suelen ser evaluados por las tasas de respuesta obtenidas, seleccionar hogares o personas con mayor probabilidad de cumplimentar la encuesta les hace más productivos, sobre todo cuando cobran por cuestionarios completados. Las medidas de control de *calidad* comúnmente utilizadas, como las comprobaciones (reentrevistas telefónicas o visitas breves para verificar que fueron encuestados), las grabaciones de audio y las marcas de tiempo, no necesariamente detectan desviaciones del protocolo, llevando a una situación peligrosa. Tasas de respuesta artificialmente altas, porque los casos difíciles de contactar no quedan registrados como no encuestados, indican que han manipulado la selección y que los datos pueden no representar a la población (Eckman y Koch, 2019). Para evitar dicha manipulación en la selección de las unidades muestrales se recomienda utilizar métodos de muestreo que minimicen su capacidad de selección, mejorar su formación y supervisión, y que no se sientan

presionados para lograr altas tasas de respuesta. También, que se apliquen controles de calidad adicionales a quienes completen muchas entrevistas en el primer contacto e incluso verificar su comportamiento a través de dispositivos GPS.

Por último, la duración de la entrevista también puede afectar a la *calidad* de la respuesta (Olson y Peytchev, 2007; Roberts *et al.*, 2019; Vandenplas, Beullens y Loosveldt, 2019). Mediciones de la duración, su ritmo (minutos por pregunta) y velocidad (preguntas por minuto), calculadas a partir de parados, sirven como indicadores del desempeño del entrevistador/a. Quienes más se desvían del protocolo de entrevista estandarizada y más aceleran la entrevista son quienes más contribuyen a este componente del error de *medición* (Olson, Smyth y Kirchner, 2020; Vandenplas, Beullens y Loosveldt, 2019; Wuyts y Loosveldt, 2022).

En suma, la varianza del entrevistador/a es un componente clave del error de *medición* y cuantificable. Abarca todas las desviaciones de la respuesta media general que resulta de la combinación individual de sus características físicas, estilo de entrevista y cumplimentación del cuestionario (como escribir la respuesta literal a preguntas abiertas, marcar correctamente las respuestas a preguntas cerradas o no saltarse alguna pregunta). Si bien, su efecto en la respuesta puede ser aleatorio (diferentes errores en cada entrevista) o sistemático (en la totalidad de las realizadas). En este último caso tendría mayor impacto en la *calidad* de la encuesta. Su reducción precisa ampliar el número de entrevistadores/as para favorecer que una mala actuación repercuta en un menor número de cuestionarios y que el error sea aleatorio, ampliándose la varianza del entrevistador/a; además de intensificar su formación y supervisión. A este respecto Wuyts y Loosveldt (2022) recomiendan las grabaciones de audio para eliminar o vol-

ver a formar a los entrevistadores/as con las peores prácticas de entrevista al principio del trabajo de campo. También, utilizar los datos de «rastros» de su actuación, como las pulsaciones de teclas, que registran todas las entradas realizadas desde el teclado, ratón o pantalla táctil. Al igual que los datos de tiempo de las entrevistas, su recopilación es gratuita y pueden utilizarse para señalar prácticas sospechosas. A ello se añade la valoración de circunstancias donde sus características específicas pueden afectar a la respuesta. Especialmente, cuando el tema de la encuesta está directamente relacionado con alguna de sus características visibles y la persona encuestada oculta su respuesta por considerar ofensiva o embarazosa exponérsela, como ya indicaron Fowler y Mangione (1990).

CONCLUSIONES

La encuesta se mantiene como estrategia metodológica prevalente para obtener un gran volumen de información que permita describir y comprender la formación de la opinión pública, los cambios a lo largo del tiempo y los vínculos entre actitudes y comportamientos de la población. Pero, para alcanzar sus objetivos, se precisa que aporte datos creíbles para quienes la financian, utilizan y analizan; que cumpla mínimos criterios de *calidad* que justifiquen su elevado coste.

En la última década, ha habido avances, pero también retrocesos en su cumplimiento, por el deseo de reducir su coste económico y el tiempo preciso para disponer de la información que aporte. Sin duda, los avances informáticos están contribuyendo a la digitalización de la encuesta y su abaratamiento, acompañado de mejoras de *calidad*, al reducir errores debidos a la administración del cuestionario, la grabación de las respuestas y la actuación del entrevistador/a (al grabarse datos de la en-

trevista). Si bien, dichos avances aún no son la panacea. Siguen detectándose notorias lagunas en errores de *no observación y medición*, que no siempre solventa la aplicación de encuestas *mixtas*, por incompatibilidades de marcos muestrales, procedimientos de selección de la muestra (probabilísticos y no), diseño del cuestionario y comparabilidad de las respuestas; en especial, en temas subjetivos y vulnerables al sesgo de *deseabilidad social*.

Tampoco lo es que la encuesta sea cumplimentada mediante dispositivo móvil. Mejoras en la conectividad, duración de baterías, interfaces móviles (ingreso de texto más fácil), el diseño del cuestionario y la recopilación de datos objetivos pueden incrementar su dominio en la investigación social, aun no siendo panacea. Continúa enfrentándose a grandes hándicaps, a lo que no ayuda que no medie un entrevistador/a en la recogida de información.

De la revisión de la investigación empírica realizada a lo largo del presente artículo se desprende el debate de qué ha de priorizarse: la disponibilidad de datos en un plazo relativamente corto a un mínimo coste o que la encuesta sea de *calidad*, aunque con un mayor coste económico y temporal. De lo leído el lector podrá extraer sus propias conclusiones.

BIBLIOGRAFÍA

- Andreadis, Ioannis (2020). «Text Message (SMS) Pre-Notifications, Invitations and Reminders for Web Surveys». *Survey Methods: Insights from the Field*, 8. doi: 10.13094/SMIF-2020-00019
- Anduiza, Eva y Galais, Carol (2016). «Answering without Reading: IMCS and Strong Satisficing in Online Surveys». *International Journal of Public Opinion Research*, 29(3): 497-519. doi:10.1093/ijpor/edw007
- Antoun, Christopher; Couper, Mick P. y Conrad, Frederick G. (2017). «Effects of Mobile Versus PC Web on Survey Response Quality: A Crossover Experiment in a Probability Web Panel». *Public Opinion Quarterly*, 81(1): 280-306. doi:10.1093/poq/nfw088
- Antoun, Christopher; Conrad, Frederick G.; Couper, Mick P. y West, Brady T. (2019). «Simultaneous Estimation of Multiple Sources of Error in a Smartphone-Based Survey». *Journal of Survey Statistics and Methodology*, 7(1): 93-117. doi:10.1093/jssam/smy002
- Bărbulescu, Marinică y Cernat, Alexandru (2012). «The Impact of Pictures on Best-Worst Scaling in Web Surveys». *International Review of Social Research*, 2(3): 79-93. doi:10.1515/irsr-2012-0028
- Bech, Mickael y Kristensen, Morten Bo (2009). «Differential Response Rates in Postal and Web-Based Surveys in Older Respondents». *Survey Research Methods*, 3(1): 1-6. doi:10.18148/srm/2009.v3i1.592
- Becker, Rolf; Möser, Sara y Glauser, David (2019). «Cash vs. Vouchers vs. Gifts in Web Surveys of a Mature Panel Study». *Social Science Research*, 81: 221-234. doi: https://doi.org/10.1016/j.ssresearch.2019.02.008
- Beller, Johannes y Geyer, Siegfried (2021). «Personal Values Strongly Predict Study Dropout». *Survey Research Methods*, 15(3): 269-280. doi:10.18148/srm/2021.v15i3.7801
- Berinsky, Adam J.; Margolis, Michele F. y Sances, Michael W. (2014). «Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-Administered Surveys». *American Journal of Political Science*, 58(3): 739-753. doi: 10.1111/ajps.12081
- Bethlehem, Jelke (2010). «Selection Bias in Web Surveys». *International Statistical Review*, 78(2): 161-188. doi: 10.1111/j.1751-5823.2010.00112.x
- Bethlehem, Jelke (2016). «Solving the Nonresponse Problem with Sample Matching?». *Social Science Computer Review*, 34: 59-77. doi:10.1177/0894439315573926
- Bethlehem, Jelke; Cobben, Fannie y Schouten, Barry (2011). *Handbook of Nonresponse in Household Surveys*. Hoboken: John Wiley & Sons.
- Biemer, Paul P. y Lyberg, Lars E. (2003). *Introduction to Survey Quality*. Hoboken: John Wiley & Sons.
- Bittmann, Felix (2020). «The More Similar, the Better? How (Mis)Match Between Respondents and Interviewers Affects Item Nonresponse and Data Quality in Survey Situations». *Survey Research Methods*, 14(3): 301-323. doi:10.18148/srm/2020.v14i3.7621

- Blanchard, Maxime (2022). «Skin Tones and Polarized Politics: How Skin Color Differences Between Interviewers and Respondents Influence Survey Answers in Bolivia». *International Journal of Public Opinion Research*, 34(1). doi:10.1093/ijpor/edac007
- Braunsberger, Karin; Wybenga, Hans y Gates, Roger (2007). «A Comparison of Reliability Between Telephone and Web-based Surveys». *Journal of Business Research*, 60(7): 758-764. doi:10.1016/j.jbusres.2007.02.015
- Callegaro, Mario; Baker, Reg; Bethlehem, Jelke; Göritz, Anja S.; Krosnick, John A. y Lavrakas, Paul J. (2014). *Online Panel Research: A Data Quality Perspective*. UK: John Wiley and Sons.
- Callegaro, Mario; Manfreda, Katja L. y Vehovar, Vasja (2015). *Web Survey Methodology*. London: Sage.
- Cea D'Ancona, M.^a Ángeles (2017). «Measuring Multiple Discrimination Through Survey Methodology». *Social Science Research*, 67: 239-251. doi:10.1016/j.ssresearch.2017.04.006
- Cea D'Ancona, M.^a Ángeles (2022). «Calidad, Confianza y Participación en Encuestas». *Papers* 107(4): 1-27. doi:10.5565/rev/papers.3074.
- Cea D'Ancona, M.^a Ángeles y Valles Martínez, Miguel S. (2021). «Multiple Discrimination: From Perceptions and Experiences to Proposals for Anti-discrimination Policies». *Social & Legal Studies*, 30(6): 937-958. doi: 10.1177/0964663920983534
- Cernat, Alexandru y Revilla, Melanie (2020). «Moving from Face-to-Face to a Web Panel: Impacts on Measurement Quality». *Journal of Survey Statistics and Methodology*, 9(4): 1-19. doi: 10.1093/jssam/smaa007
- Chang, Linchiat y Krosnick, Jon A. (2009). «National Surveys via RDD Telephone Interviewing Versus the Internet: Comparing Sample Representativeness and Response Quality». *Public Opinion Quarterly*, 73(4): 641-678. doi: 10.1093/poq/nfp075
- Christian, Leah M.; Dillman, Don A. y Smyth, Jolene D. (2007). The Effects of Mode and Format on Answers to Scalar Questions in Telephone and Web Surveys. En: J.M. Lepkowski; C. Tucker; M. Bryck et al. (eds.). *Advances in Telephone Survey Methodology*. Hoboken: Wiley & Sons.
- Conrad, Frederick G.; Schober, Michael F.; Antoun, Christopher; Yan, H. Yanna; Hupp, Andrew L.; Johnston, Michael; Ehlen, Patrick; Vickers, Lucas y Zhang, Chan (2017). «Respondent Mode Choice in a Smartphone Survey». *Public Opinion Quarterly*, 81(1): 307-337. doi: 10.1093/poq/nfw097
- Cornesse, Carina y Bosnjak, Michael (2018). «Is There an Association Between Survey Characteristics and Representativeness? A Meta-analysis». *Survey Research Methods*, 12(1): 1-13. doi: 10.18148/srm/2018.v12i1.7205
- Cornesse, Carina; Blom, Annelies G.; Dutwin, David; Krosnick, Jon A.; De Leeuw, Edith D.; Legleye, Stéphane; Pasek, Josh; Pennay, Darren; Phillips, Benjamin; Sakshaug, Joseph W.; Struminskaya, Bella y Wenz, Alexander (2020). «A Review of Conceptual Approaches and Empirical Evidence on Probability and Nonprobability Sample Survey Research». *Journal of Survey Statistics and Methodology*, 8(1): 4-36. doi: 10.1093/jssam/smz041
- Couper, Mick P. (2000). «Web Surveys: A Review of Issues and Approaches». *Public Opinion Quarterly*, 64(4): 464-494. doi: 10.1086/318641
- Couper, Mick P. (2008). *Designing Effective Web Surveys*. New York: Cambridge University Press.
- Couper, Mick P.; Singer, Eleanor; Conrad, Frederick G. y Groves, Robert M. (2010). «Experimental Studies of Disclosure Risk, Disclosure Harm, Topic Sensitivity, and Survey Participation». *Journal of Official Statistics*, 26: 287-300.
- Couper, Mick P.; Gremel, Garret; Axinn, William; Guyer, Heidi; Wagner, James y Wes, Brady T. (2018). «New Options for National Population Surveys: The Implications of Internet and Smartphone Coverage». *Social Science Research*, 73: 221-235. doi: 10.1016/j.ssresearch.2018.03.008
- DeCastellarnau, Anna y Revilla, Melanie (2017). «Two Approaches to Evaluate Measurement Quality in Online Surveys». *Survey Research Methods*, 11(4): 415-433. doi: 10.18148/srm/2017.v11i4.7226
- Dohrenwend, Barbara S.; Colombotos, John y Dohrenwend, Bruce (1968). «Social Distance and Interviewer Effects». *Public Opinion Quarterly*, 32(3): 410-422. doi: 10.1086/267624
- Durrant, Gabriele B.; Groves, Robert M.; Staetsky, Laura y Steele, Fiona (2010). «Effects of Interviewer Attitudes and Behaviors on Refusal in Household Surveys». *Public Opinion Quarterly*, 74(1): 1-36. doi: 10.1093/poq/nfp098
- Eckman, Stephanie y Koch, Achim (2019) «Interviewer Involvement in Sample Selection Shapes the Relationship Between Response Rates and Data Quality». *Public Opinion Quarterly*, 83(2): 313-337. doi: 10.1093/poq/nfz012

- Elevelt, Anne; Lugtig, Peter y Toepoel, Vera (2019). «Doing a Time Use Survey on Smartphones Only: What Factors Predict Nonresponse at Different Stages of the Survey Process?». *Survey Research Methods*, 13(2): 195-213. doi: 10.18148/srm/2019.v13i2.7385
- ESOMAR (2023). «Global Market Research Report». *Nota de prensa*. Disponible en: <https://www.ia-espana.org/wp-content/uploads/2022/10/Ndp-datos-sector-2021.pdf>, acceso 19 de febrero 2024.
- Fang, Qixiang; Burger, Joep; Meijers, Ralph y Berkel, Kees van (2021). «The Role of Time, Weather and Google Trends in Understanding and Predicting Web Survey Response». *Survey Research Methods*, 15(1): 1-25. doi: 10.18148/srm/2021.v15i1.7633
- Fowler, Floyd J. y Mangione, Thomas W. (1990). *Standardized Survey Interviewing: Minimizing Interviewer-Related Error*. London: Sage.
- Fricker, Scott; Galesic, Mirta; Tourangeau, Roger y Yan, Ting (2005). «An Experimental Comparison of Web and Telephone Surveys». *Public Opinion Quarterly*, 69(3): 370-392. doi: 10.1093/poq/nfi027
- Gooch, Andrew y Vavreck, Lynn (2019). «How Face-to-Face Interviews and Cognitive Skill Affect Item Non-Response». *Political Science Research and Methods*, 7(1): 143-162. doi: 10.1017/psrm.2016.20
- Göriz, Anja. S. (2006). «Incentives in Web Studies». *International Journal of Internet Science*, 1(1): 58-70.
- Groves, Robert M. (1989). *Survey Errors and Survey Costs*. Hoboken: John Wiley and Sons.
- Groves, Robert M.; Cialdini, Robert B. y Couper, Mick P. (1992). «Understanding the Decision to Participate in a Survey». *Public Opinion Quarterly*, 56(4): 475-495. doi: 10.1086/269338
- Groves, Robert M.; Singer, Eleanor y Corning, Amy (2000). «Leverage-Saliency Theory of Survey Participation». *Public Opinion Quarterly*, 64(3): 299-308. doi: 10.1093/poq/nfh00210.1086/317990
- Groves, Robert M.; Presser, Stanley y Dipko, Sarah (2004). «The Role of Topic Interest in Survey Participation Decisions». *Public Opinion Quarterly*, 68(1): 2-31. doi: 10.1093/poq/nfh002
- Groves, Robert M. y Peytcheva, Emilia (2008). «The Impact of Nonresponse Rates on Nonresponse Bias». *Public Opinion Quarterly*, 72(2): 167-189. doi: 10.1093/poq/nfn011
- Groves, Robert M.; Brick, Michael, Couper, Mick P.; Kalsbeek, William; Harris-Kojetin, Brian; Kreuter, Frauke; Pennell, Beth-Ellen; Raghunathan, Trivellore; Schouten, Barry; Smith, Tom; Tourangeau, Roger; Bowers, Ashley; Jans, Matthew; Kennedy, Courtney; Levenstein, Rachel; Olson, Kristen; Peytcheva, Emilia; Ziniel, Sonja y Wagner, James (2008). «Issues Facing the Field: Alternative Practical Measures of Representativeness of Survey Respondent Pools». *Survey Practice*, 1(3): 1-6. doi: 10.29115/SP-2008-0013
- Groves, Robert M.; Fowler, Floyd J.; Couper, Mick P.; Lepkowski, James L.; Singer, Eleanor y Tourangeau, Roger (2009). *Survey Methodology*. New York: John Wiley & Sons.
- Groves, Robert M. y Lyberg, Lars (2010). «Total Survey Error: Past, Present, and Future». *Public Opinion Quarterly*, 74(5): 849-879. doi: 10.1093/poq/nfq065
- Heerwegh, Dirk y Loosveldt, Geert (2008). «Face-to-Face Versus Web Surveying in a High-Internet-Coverage Population». *Public Opinion Quarterly*, 72(5): 836-846. doi: 10.1093/poq/nfn045
- Hope, Steven; Campanelli, Pamela; Nicolaas, Gerry; Lynn, Peter y Jäckle, Annette (2022). «The Role of the Interviewer in Producing Mode Effects». *Survey Research Methods*, 16(2): 207-226. doi: 10.18148/srm/2022.v16i2.7771
- Jäckle, Annette; Roberts, Caroline y Lynn, Peter (2010). «Assessing the Effect of Data Collection Mode on Measurement». *International Statistical Review*, 78(1): 3-20. doi: 10.1111/j.1751-5823.2010.00102.x
- Jäckle, Annette; Lynn, Peter y Burton, Jonathan (2015). «Going Online with a Face-to-Face Household Panel». *Survey Research Methods*, 9(1): 57-70. doi: 10.18148/srm/2015.v9i1.5475
- Jäckle, Annette; Burton, Jonathan; Couper, Mick P. y Lessof, Carli (2019). «Participation in a Mobile App Survey to Collect Expenditure Data as Part of a Large-Scale Probability Household Panel». *Survey Research Methods*, 13(1): 23-44. doi: 10.18148/srm/2019.v1i1.7297
- Karp, Jeffrey A. y Luehiste, Maarja (2015). «Explaining Political Engagement with Online Panels». *Public Opinion Quarterly*, 80(3): 666-693. doi: 10.1093/poq/nfw014
- Keusch, Florian (2013). «The Role of Topic Interest and Topic Salience in Online Panel Web Surveys». *International Journal of Market Research*, 55(1): 59-80. doi:10.2501/IJMR-2013-007

- Keusch, Florian; Struminskaya, Bella; Antoun, Christopher; Couper, Mick P. y Kreuter, Frauke (2019). «Willingness to Participate in Passive Mobile Data Collection». *Public Opinion Quarterly*, 83(1): 210-235. doi: 10.1093/poq/nfz007
- Keusch, Florian y Conrad, Frederick G. (2022). «Using Smartphones to Capture and Combine Self-Reports and Passively Measured Behavior in Social Research». *Journal of Survey Statistics and Methodology*, 10(4): 863-885. doi: 10.1093/jssam/smab035
- Keusch, Florian; Bähr, Sebastian; Haas, Georg-Christoph; Kreuter, Frauke y Trappmann, Mark (2023). «Coverage Error in Data Collection Combining Mobile Surveys with Passive Measurement Using Apps». *Sociological Methods & Research*, 52(2): 841-878. doi: 10.1177/0049124120914924
- Kish, Leslie (1965). *Survey Sampling*. New York: John Wiley & Sons.
- Kocar, Sebastian (2022). «Survey Response in RDD-Sampling SMS-Invitation Web-Push Study». *Survey Research Methods*, 16(3): 283-299. doi: 10.18148/srm/2022.v16i3.7846
- Kreuter, Frauke; Presser, Stanley y Tourangeau, Roger (2008). «Social Desirability Bias in CATI, IVR, and Web Surveys». *Public Opinion Quarterly*, 72(5): 847-865. doi: 10.1093/poq/nfn063
- Kühne, Simon (2018). «From Strangers to Acquaintances? Interviewer Continuity and Socially Desirable Responses in Panel Surveys». *Survey Research Methods*, 12(2): 121-146. doi: 10.18148/srm/2018.v12i2.7299
- Lavrakas, Paul J.; Pennay, Darren; Neiger, Dina y Phillips, Benjamin (2022). «Comparing Probability-Based Surveys and Nonprobability Online Panel Surveys in Australia». *Survey Research Methods*, 16(2): 241-266. doi: 10.18148/srm/2022.v16i2.7907
- Lawes, Mario; Hetschko, Clemens; Sakshaug, Joseph W. y Griebemer, Stephan. (2022). «Contact Modes and Participation in App-Based Smartphone Surveys». *Social Science Computer Review*, 40(5): 1076-1092. doi: 10.1177/0894439321993832
- Lee, Sunghee (2006). «Propensity Score Adjustments as a Weighting Scheme for Volunteer Panel Web Surveys». *Journal of Official Statistics*, 22: 329-349.
- Link, Michael W.; Murphy, Joe; Schober, Michael F.; Buskirk, Trent D.; Hunter Childs, Jennifer y Langer Tesfaye, Casey (2014). «Mobile Technologies for Conducting, Augmenting and Potentially Replacing Surveys». *Public Opinion Quarterly*, 78(4): 779-787. doi: 10.1093/poq/nfu054
- Little, Roderick J. A.; West, Brady T.; Boonstra, Phillip S. y Hu, Jingwei (2020). «Measures of the Degree of Departure from Ignorable Sample Selection». *Journal of Survey Statistics and Methodology*, 8(5): 932-964. doi: 10.1093/jssam/smz023
- Liu, Mingnan; Kuriakose, Noble; Cohen, Jon y Cho, Sarah (2015). «Impact of Web Survey Invitation Design on Survey Participation, Respondents, and Survey Responses». *Social Science Computer Review*, 34(5): 631-644. doi: 10.1177/0894439315605606
- Liu, Mingnan y Wronski, Laura (2018). «Examining Completion Rates in Web Surveys via Over 25,000 Real-World Surveys». *Social Science Computer Review*, 36(1): 116-124. doi: 10.1177/0894439317695581
- Loosveldt, Geert y Sonck, Nathalie (2008). «An Evaluation of the Weighting Procedures for an Online Access Panel Survey». *Survey Research Methods*, 2(2): 93-105. doi: 10.18148/srm/2008.v2i2.82
- Lugtig, Peter; Lensvelt-Mulders, Gerty J. L. M.; Frerichs, Remco y Greven, Assyn (2011). «Estimating Nonresponse Bias and Mode Effects in a Mixed-Mode Survey». *International Journal of Market Research*, 53(5): 669-686. doi: 10.2501/IJMR-53-5-669-686
- Lyberg, Lars E. (2012). «Survey Quality». *Survey Methodology*, 38(2): 107-130.
- Lyberg, Lars E. y Stukel, Diana M. (2017). The Roots and Evolution of the Total Survey Error Concept. En: P. P. Biemer; E. de Leeuw; S. Eckman; B. Edwards; F. Kreuter; L. E. Lyberg; N. C. Tucker y B. T. West (eds.). *Total Survey Error in Practice*. New York: Wiley.
- Maniaci, Michael R. y Rogge, Ronald D. (2014). «Caring About Carelessness: Participant Inattention and its Effects on Research». *Journal of Research in Personality*, 48: 61-83. doi: 10.1016/j.jrp.2013.09.008
- Mavletova, Aigul (2015). «Web Surveys Among Children and Adolescents: Is There a Gamification Effect?». *Social Science Computer Review*, 33(3): 372-398. doi: 10.1177/0894439314545316
- Mavletova, Aigul y Couper, Mick P. (2013). «Sensitive Topics in PC Web and Mobile Web Surveys».

- Survey Research Methods*, 7(3): 191-205. doi: 10.18148/srm/2013.v7i3.5458
- Mol, Christof van (2017). «Improving Web Survey Efficiency». *International Journal of Social Research Methodology*, 20(4): 317-327. doi: 10.1080/13645579.2016.1185255
- Olson, Kristen y Peytchev, Andy (2007). «Effect of Interviewer Experience on Interview Pace and Interviewer Attitudes». *Public Opinion Quarterly*, 71(2): 273-286. doi: 10.1093/poq/nfm007
- Olson, Kristen; Smyth, Jolene D. y Kirchner, Antje (2020). «The Effect of Question Characteristics on Question Reading Behaviors in Telephone Surveys». *Journal of Survey Statistics and Methodology*, 8(4): 636-666. doi: 10.1093/jssam/smz031
- Poses, Carlos; Revilla, Melanie; Asensio, Marc; Schwarz, Hannah y Weber, Wiebke (2021). «Measurement Quality of 67 Common Social Sciences Questions Across Countries/Languages». *Survey Research Methods*, 15(3): 235-256. doi: 10.18148/srm/2021.v15i3.7816
- Plutzer, Eric (2019). «Privacy, Sensitive Questions, and Informed Consent: Their Impacts on Total Survey Error, and the Future of Survey Research». *Public Opinion Quarterly*, 83(1): 169-184. doi: 10.1093/poq/nfz017
- Ren, Weijia; Krenzke, Tom; West, Brady T. y Cantor, David (2022). «An Evaluation of the Quality of Interviewer and Virtual Observations and Their Value for Potential Nonresponse Bias Reduction». *Survey Research Methods*, 16(1): 97-131. doi: 10.18148/srm/2022.v16i1.7767
- Revilla, Melanie; Toninelli, Daniele; Ochoa, Carlos y Loewe, Germán (2016). «Do Online Access Panels Need to Allow and Adapt Surveys to Mobile Devices?». *Internet Research*, 26(5): 1209-1227. doi: 10.1108/IntR-02-2015-0032
- Revilla, Melanie; Couper, Mick. P. y Ochoa, Carlos (2019). «Willingness of Online Panelists to Perform Additional Tasks». *Methods, Data, Analyses*, 13(2): 223-252. doi: 10.12758/mda.2018.01
- Roberts, Caroline; Gilbert, Emily; Allum, Nick y Eisner, Léila (2019). «Research Synthesis: Satisficing in Surveys». *Public Opinion Quarterly*, 83(3): 598-626. doi: 10.1093/poq/nfz035
- Roster, Catherine A.; Rogers, Robert D.; Albaum, Gerald y Klein, Darin (2004). «A Comparison of Response Characteristics from Web and Telephone Surveys». *International Journal of Market Research*, 46(3): 359-374. doi: 10.1177/147078530404600301
- Saris, Willem E. y Revilla, Melanie (2016). «Correction for Measurement Errors in Survey Research». *Social Indicators Research*, 127(3): 1005-1020. doi: /10.1007/s11205-015-1002-x
- Sinibaldi, Jennifer; Durrant, Gabriele B. y Kreuter, Frauke (2013). «Evaluating the Measurement Error of Interviewer Observed Paradata». *Public Opinion Quarterly*, 77(1): 173-193. doi: 10.1093/poq/nfs062
- Sterrett, David; Malato, Dan; Benz, Jennifer; Tompson, Trevor y English, Ned (2017). «Assessing Changes in Coverage Bias of Web Surveys in the United States». *Public Opinion Quarterly*, 81(1): 338-356. doi: 10.1093/poq/nfx002
- Struminskaya, Bella; Lugtig, Peter; Keusch, Florian y Höhne, Jan Karem (2020). «Augmenting Surveys with Data from Sensors and Apps». *Social Science Computer Review*, 0(0): 1-13. doi: 10.1177/0894439320979951
- Struminskaya, Bella; Lugtig, Peter; Toepoel, Vera; Schouten, Barry; Giesen, Deirdre y Dolmans, Ralph (2021a). «Sharing Data Collected with Smartphone Sensors». *Public Opinion Quarterly*, 85(1): 423-462. doi: 10.1093/poq/nfab025
- Struminskaya, Bella; Toepoel, Vera; Lugtig, Peter; Haan, Marieke; Luiten, Annemieke y Schouten, Barry (2021b). «Understanding Willingness to Share Smartphone-Sensor Data». *Public Opinion Quarterly*, 84(1): 725-759. doi: 10.1093/poq/nfab025
- Toepoel, Vera; Das, Marcel y Van Soest, Arthur (2008). «Effects of Design in Web Surveys: Comparing Trained and Fresh Respondent». *Public Opinion Quarterly*, 72(5): 985-1007. doi: 10.1093/poq/nfn060
- Tourangeau, Roger; Rips, Lance y Rasinski, Kenneth (2000). *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Tourangeau, Roger y Yan, Ting (2007). «Sensitive Questions in Surveys». *Psychological Bulletin*, 133(5): 859-883. doi: 10.1037/0033-2909.133.5.859
- Tourangeau, Roger; Conrad, Fredrick G. y Couper, Mick P. (2013). *The Science of Web Surveys*. New York: Oxford University Press.
- Valentino, Nicholas A.; Zhirkov, Kirill; Hillygus, D. Sunshine y Guay, Brian (2020). «The Consequences of Personality Biases in Online Panels for Measuring Public Opinion». *Public Opinion Quarterly*, 84(2): 446-468. doi: 10.1093/poq/nfaa026
- Valliant, Richard y Dever, Jill A. (2011). «Estimating Propensity Adjustments for Volunteer Web Surveys». *Sociological Methods and Research*, 40(1): 105-137. doi: 10.1177/0049124110392533

- Vandenplas, Caroline; Beullens, Koen y Loosveldt, Geert (2019). «Linking Interview Speed and Interviewer Effects on Target Variables in Face-to-Face Surveys». *Survey Research Methods*, 13(3): 249-265. doi: 10.18148/srm/2019.v13i3.7321
- Vercruyssen, Anina y Loosveldt, Geert (2017). «Using Google Maps and Google Street View to Validate Interviewer Observations and Predict Non-response». *Survey Research Methods*, 11(3): 345-360. doi: 10.18148/srm/2017.v11i3.6301
- Vercruyssen, Anina; Wuyts, Celine y Loosveldt, Geert (2017). «The Effect of Sociodemographic (Mis)match Between Interviewers and Respondents on Unit and Item Nonresponse in Belgium». *Social Science Research*, 67: 229-238. doi: 10.1016/j.ssresearch.2017.02.007
- Wang, Wei; Rothschild, David; Goel, Sharad y Gelman, Andrew (2015). «Forecasting Elections with Non-Representative Polls». *International Journal of Forecasting*, 31(3): 980-991. doi: 10.1016/j.ijforecast.2014.06.001
- Wenz, Alexander; Jäckle, Annette y Couper, Mick P. (2019). «Willingness to Use Mobile Technologies for Data Collection in a Probability Household Panel». *Survey Research Methods*, 13(1): 1-22. doi: 10.18148/srm/2019.v1i1.7298
- Wenz, Alexander y Keusch, Florian (2023). «Increasing the Acceptance of Smartphone-Based Data Collection». *Public Opinion Quarterly*, 87(2): 357-388. doi: 10.1093/poq/nfad019
- West, Brady T. y Blom, Annelies G. (2017). «Explaining Interviewer Effects». *Journal of Survey Statistics and Methodology*, 5(2): 175-211. doi: 10.1093/jssam/smw024
- West, Brady T.; Kreuter, Frauke y Trappmann, Mark (2014). «Is the Collection of Interviewer Observations Worthwhile in an Economic Panel Survey?». *Journal of Survey Statistics and Methodology*, 2(2): 159-181. doi: 10.1093/jssam/smu002
- West, Brady T.; Wagner, James; Hubbard, Frost y Gu, Haoyu (2015). «The Utility of Alternative Commercial Data Sources for Survey Operations and Estimation». *Journal of Survey Statistics and Methodology*, 3(2): 240-264. doi: 10.1093/jssam/smv004
- Wuyts, Celine y Loosveldt, Geert (2022). «Interviewer Performance in Slices or by Traces». *Survey Research Methods*, 16(2): 147-163. doi: 10.18148/srm/2022.v16i2.7672
- Zhang, Chan y Conrad, Frederick (2014). «Speeding in Web Surveys». *Survey Research Methods*, 8(2): 127-135. doi: 10.18148/srm/2014.v8i2.5453
- Zhang, XiaoChi; Kuchinke, Lars; Woud, Marcella L.; Velten, Julia y Margraf, Jürgen (2017). «Survey Method Matters». *Computers in Human Behavior*, 71: 172-180. doi: 10.1016/j.chb.2017.02.006

RECEPCIÓN: 05/04/2024

REVISIÓN: 18/10/2024

ACEPTACIÓN: 02/12/2024

