

AÑO 30 NO. ESPECIAL 13, 2025
ENERO-JUNIO



AÑO 30 NO. ESPECIAL 13, 2025

ENERO-JUNIO



Revista Venezolana de Gerencia



UNIVERSIDAD DEL ZULIA (LUZ)
Facultad de Ciencias Económicas y Sociales
Centro de Estudios de la Empresa

ISSN 1315-9984

Esta obra está bajo una licencia de Creative Commons
Reconocimiento-NoComercial-CompartirIgual 3.0 Unported.
http://creativecommons.org/licenses/by-nc-sa/3.0/deed.es_ES



Eficiencia en el tiempo de respuesta y calidad de los datos en encuestas en línea

Izaguirre Olmedo, Jorge Andrés*
Jordán Correa, Dennys Patricia**
Palacios Sarmiento, Tania Yolanda***

Resumen

El uso de encuestas digitales se intensificó durante el confinamiento por la pandemia del SARS-CoV-2, facilitando la recolección de información en la investigación científica. Sin embargo, la incorporación de un mayor número de preguntas ha extendido el tiempo de respuesta, sin considerar su incidencia en la calidad de los datos obtenidos. Este estudio examina la relación entre el tiempo de respuesta y la calidad de los datos en encuestas digitales aplicadas a investigaciones sobre liderazgo. A partir de un diseño cuantitativo no experimental, se analizó una muestra de 224 participantes que completaron un cuestionario adaptado del MLQ con 32 ítems. La regresión lineal múltiple, con género y edad como variables de control, evidenció una relación positiva entre el tiempo de respuesta y la calidad de los datos, aunque al superar los 20 minutos esta tendencia se revierte. También se observó que la edad y el género influyen significativamente en la relación analizada. Se concluye que un mayor tiempo de respuesta mejora la calidad de los datos en encuestas breves, por lo que se recomienda continuar investigando este efecto en cuestionarios más extensos.

Palabras clave: Tiempo de respuesta; calidad de los datos; regresión lineal múltiple; encuestas digitales.

Recibido: 26.02.25

Aceptado: 14.05.25

* Magister en Finanzas y Proyectos Corporativos; Estudiante de Doctorado, Escuela de posgrados, Universidad San Ignacio de Loyola, Lima – Perú; Email: jorge.izaguirre@epg.usil.pe; ORCID: <https://orcid.org/0000-0001-5178-8641>.

** Magister en Educación; Licenciada en Comunicación Social; docente, Universidad Internacional del Ecuador, Guayaquil – Ecuador; Email: dejordanco@uide.edu.ec; ORCID: <https://orcid.org/0000-0001-6962-0855>

*** Doctorando en Proyectos; Magister en negocios internacionales y comercio exterior; Docente, Universidad Internacional del Ecuador - Ecuador; Email: tapalaciossa@uide.edu.ec; ORCID: <https://orcid.org/0000-0001-5111-6319>.

Efficiency in Response Time and Data Quality in Online Surveys.

Abstract

The use of digital surveys increased significantly during the SARS-CoV-2 pandemic lockdown, facilitating data collection in scientific research. However, the inclusion of a larger number of questions has extended response time without considering its impact on data quality. This study examines the relationship between response time and data quality in online surveys applied to leadership research. Using a non-experimental quantitative design, a sample of 224 participants completed an adapted MLQ questionnaire with 32 items. Multiple linear regression, controlling for gender and age, revealed a positive relationship between response time and data quality, although this trend reverses when the response time exceeds 20 minutes. Additionally, age and gender significantly influence the analyzed relationship. The study concludes that longer response times improve data quality in short surveys, suggesting the need for further research on this effect in more extensive questionnaires.

Keywords: Response time; data quality; multiple linear regression; online surveys.

1. Introducción

Con el avance tecnológico, la interacción de los usuarios con las encuestas en línea ha experimentado cambios significativos. Según Décieux & Sischka (2024), debido a las diferencias entre los dispositivos móviles y las computadoras, como en su manejo, se han detectado impactos del tipo de dispositivo en diversos indicadores de calidad de los datos en encuestas en línea. No obstante, los hallazgos sobre estos efectos siguen siendo parciales e inconclusos, ya que gran parte de las investigaciones previas se fundamentan en diseños poco optimizados o no distinguen entre distintos tipos de dispositivos móviles, como smartphones y tabletas.

Los cuestionarios digitales de mayor extensión pueden resultar

monótonos y generar fatiga en los participantes, aumentando la probabilidad de abandono antes de completarlos. Este fenómeno no solo compromete la calidad de los datos recolectados, sino que también introduce un sesgo en los resultados, ya que los encuestados que finalizan la encuesta pueden no ser representativos de la población general. A medida que el tiempo de respuesta se prolonga, la calidad de los datos tiende a disminuir debido al agotamiento del usuario (Nur et al., 2024; Meitinger et al., 2021).

Es fundamental considerar que la claridad en la formulación de las preguntas impacta directamente en la calidad de los datos obtenidos. Para garantizar respuestas precisas y relevantes, es esencial adaptar el lenguaje del cuestionario al nivel de conocimiento y experiencia de los

encuestados, evitando ambigüedades o interpretaciones erróneas que puedan comprometer la validez de los resultados (Bauer et al., 2023).

La obtención de datos de alta calidad representa un desafío constante. En encuestas presenciales y telefónicas, las tasas de respuesta han disminuido en los últimos años. Cernat et al. (2022) recomiendan la adopción de estrategias mixtas que combinen encuestas autoadministradas con entrevistas dirigidas por un entrevistador, ya que este enfoque mejora tanto la tasa de participación como la calidad de los datos recopilados. La combinación de métodos, como la recolección de datos a través de encuestas web junto con entrevistas telefónicas o presenciales, permite captar una muestra más diversa y adaptarse a las preferencias individuales de los encuestados. Además, los cuestionarios autoadministrados brindan mayor flexibilidad, ya que los participantes pueden responder a su propio ritmo, favoreciendo la calidad de sus respuestas.

Para mejorar la calidad de los datos en encuestas digitales, De Rada (2022) propone cinco estrategias clave: garantizar un acceso cuidadoso a los encuestados, establecer un contacto frecuente con la muestra, monitorear el tiempo del trabajo de campo, notificar a los participantes sobre la encuesta y ofrecer incentivos. Estas estrategias pueden incrementar la tasa de respuesta y mejorar la representatividad de la muestra. En particular, las recompensas se han identificado como un estímulo efectivo para motivar a los participantes a completar los cuestionarios, lo que contribuye a obtener datos de mayor calidad. Además, el contacto constante con los encuestados a través de recordatorios y actualizaciones sobre la

importancia de su participación fomenta una mayor implicación en el proceso de respuesta.

Este estudio tiene como objetivo examinar la relación entre el tiempo de respuesta y la calidad de los datos en encuestas digitales aplicadas a investigaciones sobre liderazgo. Se busca determinar el impacto del tiempo de respuesta en la calidad de los datos en encuestas breves. Para ello, se emplea un cuestionario de liderazgo con 32 ítems, y se plantea la hipótesis de que un mayor tiempo de respuesta puede generar fatiga en los encuestados, afectando negativamente la calidad de sus respuestas.

2. Aspectos teóricos sobre la calidad de respuesta

Peytchev & Peytcheva (2017) sostienen que la evidencia empírica sobre la relación entre la longitud del cuestionario y el error de medición es limitada. Esta incertidumbre ha permitido el diseño de encuestas más extensas de lo óptimo, aunque los autores enfatizan la necesidad de estrategias que reduzcan su duración sin afectar los objetivos de la investigación. Wind & Lugu (2024) recomiendan el uso de modelos no paramétricos, frente a modelos paramétricos, para el estudio de calidad de respuesta.

Los cuestionarios digitales han incrementado su complejidad, y mientras algunos estudios proponen descomponerlos por ítems, otros consideran que esta práctica podría afectar la coherencia de las respuestas (Vehovar et al., 2023). Andreadis & Kartsounidou (2020) sugieren un enfoque alternativo mediante la división de encuestas en segmentos secuenciales, lo que permite presentar

nuevas preguntas conforme se completan las anteriores, obteniendo respuestas más precisas. No obstante, tanto en cuestionarios extensos como en formatos segmentados, se han identificado problemas como la omisión de respuestas, la aceleración en el llenado y patrones de respuesta en línea recta.

Chauliac et al. (2023) destacan que la atención de los encuestados puede disminuir durante la realización de cuestionarios digitales, ya que completar esta tarea de manera adecuada implica varios procesos cognitivos. Para proporcionar respuestas precisas, los participantes deben interpretar correctamente las preguntas, recuperar información relevante de su memoria y evaluar las distintas opciones de respuesta. Este esfuerzo cognitivo puede generar fatiga, especialmente cuando el tiempo de respuesta se extiende debido a la complejidad del cuestionario o la ambigüedad en la redacción de los ítems (Kunz & Gummer, 2025).

Dado que responder encuestas es un proceso cognitivo que requiere concentración, la multitarea durante la realización de cuestionarios en línea puede afectar la calidad de los datos obtenidos. La división de la atención entre múltiples actividades reduce la capacidad de reflexión sobre las preguntas, lo que puede derivar en respuestas apresuradas o inexactas. Décieux (2024) señala que la multitarea es un factor crítico que impacta negativamente en la calidad de los datos, ya que la distracción impide que los encuestados mantengan una evaluación profunda de sus respuestas.

En relación con la calidad de las respuestas, Ganassali (2008) recomienda el uso de una redacción clara y concisa para reducir la tasa de

abandono de las encuestas. Estrategias como la inclusión de preguntas dinámicas, opciones de respuesta visuales o retroalimentación en tiempo real pueden mantener el interés de los encuestados y fortalecer su conexión con el propósito de la investigación, lo que se traduce en una mayor tasa de finalización con datos más precisos y confiables (Meitinger, Behr, & Braun, 2021).

La longitud del cuestionario y su incidencia en la respuesta descuidada (response carelessness - RC) han sido objeto de estudio en diversas investigaciones. Gibson & Bowling (2020) analizaron cómo los incentivos pueden mitigar los efectos negativos de cuestionarios extensos sobre la calidad de los datos. Sus hallazgos indican que (a) existe evidencia sobre la influencia del tiempo de respuesta en la calidad de los datos, (b) las manipulaciones en el diseño del cuestionario pueden generar efectos significativos y (c) aún se requieren más estudios sobre el rol moderador de estas estrategias. En consecuencia, sugieren profundizar en las causas y métodos de prevención de la RC para mejorar la precisión de los datos recopilados.

Gummer et al. (2021) destacan la utilidad de las encuestas adaptativas y reactivas para diversificar la composición de la muestra en estudios digitales. Sin embargo, advierten que la variabilidad en la calidad de las respuestas puede estar influida por el error de medición, el sesgo en la muestra y las diferencias en la interpretación de las preguntas. Además, la falta de control sobre el contexto en el que se responden las encuestas dificulta la evaluación de la fiabilidad de los datos obtenidos. En este sentido, la precisión de la información recopilada sigue siendo un desafío en

estudios que dependen de encuestas en línea (Ito & Todoroki, 2021).

El análisis de la verificación y control de la calidad de los datos en encuestas digitales requiere considerar distintos indicadores. Korytnikova (2021) propone una clasificación de los parados en tres categorías: etapa de investigación, método de registro y software utilizado. Esta clasificación facilita la estandarización y universalización del proceso de control de calidad. La selección de parados adecuados permite identificar factores que pueden afectar la calidad de la información, como la coherencia de las respuestas, la completitud de los cuestionarios y la representatividad de la muestra. Además, la implementación de procesos de limpieza de datos y la definición de criterios claros para corregir anomalías minimizan los errores en los resultados.

El fraude en encuestas digitales es otro aspecto relevante en la recopilación de datos de calidad, especialmente cuando se ofrecen incentivos a los participantes. Lawlor et al. (2021) desarrollaron un método sistemático para detectar respuestas fraudulentas en encuestas digitales. Su propuesta incluye el uso de preguntas de control para identificar inconsistencias en las respuestas. La aplicación de herramientas especializadas en la validación de datos contribuye a garantizar la integridad y fiabilidad de la información obtenida en estudios basados en encuestas en línea.

En el diseño de cuestionarios, el uso de la respuesta forzada (AF) se considera una estrategia efectiva para asegurar la completitud de los datos. Sin embargo, este método puede generar

incomodidad en los encuestados, ya que impide omitir preguntas irrelevantes o sensibles. Si bien la respuesta forzada reduce la incidencia de respuestas ambiguas o incompletas, su aplicación debe evaluarse cuidadosamente para evitar que afecte la disposición de los participantes a responder con precisión (Sischka et al., 2022). La calidad de los datos también puede mejorarse mediante técnicas de depuración y limpieza de respuestas. Wang et al. (2024) destacan que la implementación de estos procedimientos es fundamental en entornos de medición para garantizar que los datos recopilados sean fiables y representativos de la población objetivo.

En este contexto, se reconoce que la duración de un cuestionario influye en la calidad de los datos, pero aún persisten interrogantes sobre la magnitud de este efecto y su aplicabilidad en diferentes tipos de encuestas. Aunque estudios previos han analizado los factores que determinan la calidad de las respuestas, sigue sin definirse el tamaño óptimo de un cuestionario digital que maximice la calidad de los datos sin generar agotamiento en los participantes.

Garbarski et al. (2025) indican que el formato de selección categórica es preferido y se asocia con una menor falta de respuesta en el ítem y tiempos de respuesta más cortos. Schneider et al. (2024) estudiaron la calidad de respuesta en personas de edades avanzadas y verificaron que las respuestas de menor calidad entre los encuestados con habilidades cognitivas más bajas introducen errores aleatorios y sistemáticos en las mediciones de encuestas, lo que reduce la fiabilidad, validez y reproducibilidad de los resultados.

3. Enfoque del estudio

La investigación se enmarca dentro del enfoque básico-deductivo y adopta un diseño cuantitativo con un alcance causal. Se emplea el método de encuesta como técnica principal de recolección de datos, lo que justifica su carácter cuantitativo. Además, la relación entre el tiempo de respuesta y la calidad de los datos permite establecer un análisis de causalidad, dado que el tiempo de llenado del cuestionario se considera la variable independiente que puede influir en la calidad de las respuestas proporcionadas.

El diseño del estudio es no experimental y la población objetivo está compuesta por personas mayores de 18 años en Ecuador y Perú. Se aplicó un muestreo no probabilístico por conveniencia, obteniendo un total de 224 observaciones. El instrumento de medición corresponde a una versión adaptada del MLQ de Avolio & Bass, con 32 preguntas. Esta adaptación se realizó con el objetivo de garantizar que la medición de la calidad de los datos no se vea afectada por problemas de comprensión o dificultades en la interpretación de las preguntas. Dado que las variables clave en este estudio son el tiempo de respuesta y la calidad de los datos, la eliminación de ciertos reactivos del cuestionario original del MLQ no altera la validez de los resultados.

El muestreo utilizado conlleva una limitación inherente, ya que la recolección de datos se llevó a cabo mediante la difusión de un enlace de encuesta en redes sociales y grupos de afinidad, lo que introduce un sesgo de selección. Debido a esto, el estudio no tiene un alcance inferencial. Sin embargo, el interés de la investigación no radica en conocer la opinión o características

individuales de los encuestados, sino en analizar el comportamiento asociado al llenado de cuestionarios y su impacto en la calidad de los datos.

Dado el enfoque cuantitativo del estudio y la necesidad de registrar con precisión la variable independiente, el cuestionario fue implementado en la plataforma Microsoft Forms, que permite registrar automáticamente la hora de inicio y finalización del cuestionario para calcular el tiempo de respuesta. Las preguntas se estructuraron en formato GRID, siguiendo la metodología utilizada en estudios previos. Para optimizar la tasa de respuesta, los reactivos fueron organizados en bloques de 10 preguntas, en concordancia con los hallazgos de Andreadis & Kartsounidou (2020). No se implementó la respuesta forzada, permitiendo a los encuestados abandonar el cuestionario en cualquier momento.

El procesamiento de los datos se realizó utilizando el software estadístico SPSS. El análisis incluyó los siguientes pasos: i) asignación de etiquetas y codificación de variables; ii) cálculo de la variable independiente con base en los datos registrados automáticamente por la plataforma; iii) construcción de las dimensiones de la variable dependiente, calidad de los datos, según lo establecido en la Tabla 1; iv) construcción de variable ficticia que recoja tiempos de respuesta superiores a 20 minutos, dado que podrían indicar distracción o multitarea, conforme a lo señalado por Décieux (2024); y v) aplicación de análisis estadístico para examinar la relación entre las variables.

El análisis de fiabilidad de los datos mostró un Alfa de Cronbach de 0.824, lo que supera el umbral crítico de 0.75, garantizando la consistencia interna del instrumento. El Cuadro 1 presenta el cómputo detallado de la variable calidad de los datos.

Cuadro 1
Cómputo de la variable de calidad de los datos – CD

Variable	Código	Descripción / Cómputo
Abandono	aband	Toma el valor de 1 si el encuestado no finalizó el cuestionario.
Perdidos	lost	Toma el valor de 1 si existe algún ítem sin contestar.
Desviación estándar 0	ddesvest	Toma el valor de 1 si la desviación estándar es igual a 0.
Respuesta en extremos	respext	Corresponde al porcentaje de respuestas 1, 3 o 5 en el cuestionario.
Falta de atención reportada	lackatt_rep	Toma valores entre 0 y 1 como porcentaje de atención. Si el encuestado responde que no está prestando atención en las preguntas de verificación, la variable toma valor de 1.
Falta de atención observada	lackatt_obs	Toma el valor de 1 si se observan contradicciones en las respuestas acorde a las preguntas de control ingresadas.
Calidad de los datos	CD	Toma valores de 0 a 6.

La ecuación del cálculo de calidad de los datos se presenta a continuación:

$$CD = 6 - aband - lost - ddesvest - respext - lackatt_{rep} - lackatt_{obs} \quad (\text{Ecuación 1})$$

Para el análisis se desarrolló un modelo de mínimos cuadrados ordinarios que utilizó como variable independiente el tiempo de respuesta del cuestionario y como variable dependiente a la calidad de los datos; controlando efectos de género, edad y esfuerzo requerido para el llenado de la encuesta (carga).

La carga es considerada una de las dimensiones de calidad de respuesta que, en este modelo, se utiliza como variable explicativa. Una mayor sensación de carga en el llenado de la encuesta incidirá en la calidad de respuesta. El modelo econométrico se presenta en la siguiente ecuación.

$$CD_i = \beta_1 \cdot tiempo_i + \beta_2 \cdot edad_i + \beta_3 \cdot femenino_i + \beta_4 \cdot carga_i + \beta_5 \cdot tiempoexc_i + \varepsilon_i \quad (\text{Ecuación 2})$$

Donde CD corresponde a la calidad de datos, tiempo es el tiempo de respuesta en minutos, edad corresponde a la edad en años del encuestado, femenino es una variable dummy que toma el valor de 1 cuando el encuestado reporta género femenino, carga es tomado directo de la encuesta en el reactivo final referente al esfuerzo

de llenado percibido por el encuestado. La variable tiempoexc es una variable ficticia que toma el valor de 1 cuando el encuestado tardó más de 20 minutos. ε representa el error estocástico del modelo. El modelo teórico considera que la regresión se obtenga sin intercepto debido a que, si el tiempo de respuesta es 0, implica que no se desarrolló la

encuesta y, por tanto, no hay calidad de datos que obtener.

Se debe reconocer, como alcance de la investigación, que los resultados obtenidos no pueden generalizarse a la población, así como tampoco son representativos de otros instrumentos de medición. Los resultados pueden sugerir la incidencia del tiempo de respuesta sobre calidad de datos para los participantes del estudio y al respecto del cuestionario utilizado.

3. Relación tiempo de llenado de la encuesta y calidad de los datos

Como se indicó en la sección de Método, se estimó un modelo de regresión lineal múltiple que tomó como variable dependiente a la calidad de los

datos, la cual fue tratada en términos de logaritmos naturales para interpretar los resultados como cuasielasticidades (Caria et al., 2024). La variable independiente de principal interés es el tiempo de respuesta medido en minutos. Otras variables independientes incluidas fueron el nivel de dificultad o carga, la edad y una variable dummy para encuestados de género femenino.

La tabla 1 presenta la descripción de las variables. El test Kolmogorov-Smirnov muestra que los datos no se distribuyen de manera normal, por lo que se procedió a calcular la correlación de Spearman (Staliunaite et al., 2024). La cuarta columna muestra la correlación de las variables independientes con la variable dependiente. Se observa que el tiempo presenta una correlación positiva – leve y significativa.

Tabla 1
Descripción de las variables

Variables	Media	K-S test*	Correlación con V. Dep.
Calidad de los datos	4.96	0.000	1.000
Tiempo de respuesta (minutos)	6.72	0.000	0.161**
El encuestado se identifica con género femenino	0.48	0.000	0.91
Nivel de dificultad o carga del cuestionario reportado	1.94	0.000	-0.036
Edad	33.13	0.002	0.001

* Hipótesis nula: La variable se distribuye normal.

** Correlación de Spearman significativa al 1%.

El análisis de varianza reporta que el modelo es significativo ($p < 0.01$). El R cuadrado ajustado presenta una explicación del modelo en 0.932. Es importante recordar que el modelo no

incluye un intercepto o constante debido al modelo teórico: si no existe tiempo de respuesta, no existe encuesta y por tanto no hay calidad de datos. La tabla 2 presenta los resultados del modelo.

Tabla 2
Regresión lineal – Logaritmo natural de calidad de datos

Variables predictoras	B	Std. Error	Beta est.	t	Sig.
Tiempo de respuesta (minutos)	.014	.005	.128	3.018	.003

Cont... Tabla 2

El encuestado se identifica con género femenino	.265	.056	.117	4.719	.000
Nivel de dificultad o carga del cuestionario reportado	.135	.020	.195	6.679	.000
Edad	.032	.001	.684	21.536	.000
El tiempo supera los 20 minutos	-.79	.294	-.102	-2.672	.008

La variable independiente principal es el tiempo de respuesta de la encuesta, que se ha dividido en dos partes: tiempo en minutos y variable ficticia que toma el valor de 1 si el tiempo supera los 20 minutos. Como se observa en la tabla 3, ambas variables se muestran significativas ($p < 0.01$).

Al respecto del tiempo en minutos, los resultados sugieren que por cada minuto adicional que toma el encuestado en llenar la encuesta, la calidad de datos se incrementa en 1.4%. En contraste, aquellas encuestas en las que el encuestado tardó más de 20 minutos en el llenado presentan una calidad de respuesta 78.60% menor que aquellas encuestas en las que el encuestado se tomó menos de 20 minutos. Lo anterior coincide con los resultados de investigaciones previas respecto del efecto existente del tiempo de respuesta sobre la calidad de datos en encuestas digitales (Nur et al., 2024; Goerres & Höhne, 2023).

Andreadis & Kartsounidou (2020) señalaron que cuestionarios extensos implican menor calidad de data, lo que sería contrario a la evidencia de este estudio respecto al tiempo de respuesta en minutos; sin embargo, sería congruente con respecto a la evidencia de la dummy de tiempo de 20 minutos.

La variable ficticia de género femenino se muestra significativa. En promedio, la calidad de datos en encuestas digitales llenadas por personas de género femenino es 26.5%

mayor que la de las encuestas llenadas por personas de otro género. La edad también mostró un impacto positivo y significativo sobre la calidad de la data: por cada año adicional que tiene el encuestado, la calidad de la data se incrementa en 3.2%.

Finalmente, la dificultad percibida o carga del cuestionario también mostró un impacto positivo y significativo. Esta variable era ordinal en una escala de Likert de 5 dimensiones y salía directamente del cuestionario. En este sentido, por cada nivel adicional de carga reportado por el encuestado, la calidad de la data se incrementó en 13.5%, lo cual es consistente con las evidencias de Nun et al. (2023) y Vehovar et al. (2023). Este resultado tiene sentido a juicio de los investigadores, puesto que el nivel de carga está relacionado con la concentración que tiene el encuestado al momento de llenar la encuesta. A su vez, esto contribuye a proporcionar respuestas más reflexivas y precisas, destacando la importancia de considerar la percepción de dificultad en el diseño de cuestionarios para optimizar la calidad de los datos obtenidos.

4. Conclusiones

El modelo de regresión lineal estimado proporcionó una comprensión valiosa de los factores que influyen en la calidad de datos de los encuestados. Al tratar la calidad en términos de logaritmos, se logró interpretar los

resultados como cuasielasticidades, lo que facilitó una evaluación más precisa de la relación entre el tiempo de respuesta y la calidad de la data. La inclusión de variables como el nivel de dificultad, la edad y el género femenino en el análisis permitió identificar matices importantes que pueden afectar esta relación. Estos hallazgos enriquecieron la comprensión del proceso de respuesta en encuestas y sugieren áreas de mejora para optimizar la calidad de la recolección de datos en futuros estudios.

El tiempo de respuesta de la encuesta fue la principal variable independiente de este estudio, lo que permite revelar hallazgos significativos que impactan en la calidad de los datos. La división del tiempo en minutos y la inclusión de una variable ficticia para tiempos superiores a 20 minutos proporcionaron una comprensión más matizada de esta relación. Estos resultados enfatizan la importancia de optimizar el tiempo que los encuestados dedican a las encuestas para garantizar la validez y la calidad de los datos recolectados.

Los hallazgos permiten concluir también que las encuestas digitales completadas por féminas presentan una mayor calidad de data. Este comportamiento puede atribuirse a diversos factores: mayor compromiso y atención al detalle durante el proceso de llenado, más disposición a proporcionar respuestas reflexivas y precisas. Así, se corrobora la importancia de considerar el género como una variable relevante en la investigación, sugiriendo que las diferencias en la calidad de respuesta podrían influir en los resultados y en la interpretación de los datos. Cabe recalcar que la edad del encuestado también se asocia positivamente con la calidad de los datos, lo que indica

que, a medida que aumenta la edad, mejora la precisión y la fiabilidad de las respuestas. Este patrón sugiere que los encuestados mayores pueden tener una mayor capacidad para reflexionar sobre las preguntas y proporcionar respuestas más confiables; además, la familiaridad con el formato de las encuestas digitales y una mejor comprensión de los temas tratados contribuyen a esta tendencia. Estos resultados subrayan la importancia de considerar tanto el género como la edad en el diseño y análisis de encuestas digitales, ya que estos factores pueden influir significativamente en la calidad de la información recolectada, contribuyendo a una mejor comprensión y aplicación de los datos en la investigación.

La relación entre el tiempo de respuesta y la calidad de los datos responde a dinámicas complejas enmarcadas en teorías sobre la extensión del cuestionario, la fatiga cognitiva y la carga de los encuestados. Si bien los resultados respaldan la idea de que un mayor tiempo de llenado contribuye a respuestas más reflexivas y detalladas, la disminución de la calidad después de los veinte minutos sugiere la presencia de un umbral crítico en la atención y el compromiso del encuestado. Este hallazgo concuerda con estudios sobre fatiga cognitiva, que indican que el esfuerzo sostenido puede generar desmotivación y respuestas menos precisas. Además, la identificación de género y edad como variables significativas implica que ciertos grupos pueden ser más resilientes ante cuestionarios extensos, lo que desafía enfoques homogéneos en el diseño de encuestas digitales. En este sentido, la optimización del tiempo de respuesta no solo debe considerar la extensión del cuestionario,

sino también las características de la población encuestada, lo que abre nuevas líneas de investigación sobre estrategias adaptativas para mejorar la precisión y representatividad de los datos recolectados.

Referencias

- Andreadis, I., & Kartsounidou, E. (2020). The impact of splitting a long online questionnaire on data quality. *Survey Research Methods*, 14(1), 31–42. https://doi.org/10.18148/srm/2020_v14i1.7294
- Avolio B., & Bass, B. (2004). *Multifactor Leadership Questionnaire. Instrument (Leader and Rater Form) and Scoring Guide (Form 5X-Short)*. English and Spanish versions. Mind Garden, Inc.
- Bauer, I., Kunz, T., & Gummer, T. (2025). Plain language in web questionnaires: effects on data quality and questionnaire evaluation. *International Journal of Social Research Methodology*, 28(1), 57–69. <https://doi.org/10.1080/13645579.2023.2294880>
- Caria, A., Delogu, M., Meleddu, M., & Sotgiu, G. (2024). People inflows as a pandemic trigger: Evidence from a quasi-experimental study. *Economics and Human Biology*, 52(101341), 101341. <https://doi.org/10.1016/j.ehb.2023.101341>
- Cernat, A., Sakshaug, J., Christmann, P., & Gummer, T. (2024). The impact of survey mode design and questionnaire length on measurement quality. *Sociological Methods & Research*, 53(4), 1873–1904. <https://doi.org/10.1177/00491241221140139>
- Chauliac, M., Willems, J., Gijbels, D., & Donche, V. (2023). The prevalence of careless response behaviour and its consequences on data quality in self-report questionnaires on student learning. *Frontiers in education*, 8. <https://doi.org/10.3389/educ.2023.1197324>
- Cornesse, C., & Blom, A. G. (2023). Response quality in nonprobability and probability-based online panels. *Sociological Methods & Research*, 52(2), 879–908. <https://doi.org/10.1177/0049124120914940>
- Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. (5th Edition). SAGE Publications.
- De Rada, V. D. (2022). Strategies to improve response rates to online surveys. *Papers: Revista de sociología*, 107(4), e3073. <https://doi.org/10.5565/rev/papers.3073>
- Décieux, J. P. (2024). Sequential on-device multitasking within online surveys: A data quality and response behavior perspective. *Sociological Methods & Research*, 53(3), 1384–1411. <https://doi.org/10.1177/00491241221082593>
- Décieux, J. P., & Sischka, P. E. (2024). Comparing data quality and response behavior between smartphone, tablet, and computer devices in responsive design online surveys. *SAGE Open*, 14(2). <https://doi.org/10.1177/21582440241252116>
- Ganassali, S. (2008) The Influence of the Design of Web Survey Questionnaires on the Quality of Responses. *Survey Research Methods*, 2(1), 21–32. https://doi.org/10.18148/srm/2008_v2i1.598
- Garbarski, D., Dykema, J., Yonker, J. A., Bae, R. E., & Rosenfeld, R. A. (2025). Improving the Measurement of Gender in Surveys: Effects of

- Categorical Versus Open-Ended Response Formats on Measurement and Data Quality Among College Students. *Journal of Survey Statistics and Methodology*, 13(1), 18–38. <https://doi.org/10.1093/jssam/smae043>
- Gibson, A. M. & Bowling, N. A. (2020). The Effects of Questionnaire Length and Behavioral Consequences on Careless Responding. *European Journal of Psychological Assessment*, 36(2), 410–420. <https://doi.org/10.1027/1015-5759/a000526>
- Goerres, A., & Höhne, J. K. (2023). Evaluating the response effort and data quality of established political solidarity measures: a pre-registered experimental test in an online survey of the German adult resident population in 2021. *Quality & Quantity*, 57(6), 1–17. <https://doi.org/10.1007/s11135-022-01594-4>
- Gummer, T., Bach, R., Daikeler, J., & Eckman, S. (2021). The relationship between response probabilities and data quality in grid questions. *Survey Research Methods*, 15(1), 65–77. <https://doi.org/10.18148/srm/2021.v15i1.7727>
- Hernández, R., Fernández, C., y Baptista, M. (2014). *Metodología de la investigación*. McGraw Hill Education.
- Ito, D., & Todoroki, M. (2021). Evaluating the quality of online survey data collected in 2018 in the USA: Univariate, bivariate, and multivariate analyses. *International Journal of Japanese Sociology*, 30(1), 140–162. <https://doi.org/10.1111/ijjs.12117>
- Korytnikova, N. V. (2021). Paradata as indicators of online survey data quality: Classification experience. *Sotsiologicheskie issledovaniia*, 3, 111–120. <https://doi.org/10.31857/s013216250010298-0>
- Kunz, T., & Gummer, T. (2025). Effects of objective and perceived burden on response quality in web surveys. *International Journal of Social Research Methodology*, 28(4), 385–395. <https://doi.org/10.1080/13645579.2024.2393795>
- Lawlor, J., Thomas, C., Guhin, A. T., Kenyon, K., Lerner, M. D., UCAS Consortium, & Drahota, A. (2021). Suspicious and fraudulent online survey participation: Introducing the REAL framework. *Methodological Innovations*, 14(3), 2059799121110504. <https://doi.org/10.1177/205979912111050467>
- Meitinger, K., Behr, D., & Braun, M. (2019). Using apples and oranges to judge quality? Selection of appropriate cross-national indicators of response quality in open-ended questions. *Social Science Computer Review*, 089443931985984. <https://doi.org/10.1177/0894439319859848>
- Nur, A. A., Leibbrand, C., Curran, S. R., Votruba-Drzal, E., & Gibson-Davis, C. (2024). Managing and minimizing online survey questionnaire fraud: Lessons from the triple C project. *International Journal of Social Research Methodology*, 27(5), 613–619. <https://doi.org/10.1080/13645579.2023.2229651>
- Peytchev, A. & Peytcheva, E. (2017). Reduction of measurement error due to survey length: Evaluation of the split questionnaire design approach. *Survey Research Methods*, 11(4), 361–368. <https://doi.org/10.18148/srm/2017.v11i4.7145>
- Schneider, S., Lee, P.-J., Hernandez, R., Junghaenel, D. U., Stone, A. A., Meijer, E., Jin, H., Kapteyn, A., Orriens, B., & Zelinski, E. M. (2024). Cognitive Functioning and the Quality of Survey Responses: An Individual Participant Data Meta-Analysis of

- 10 Epidemiological Studies of Aging. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, 79(5). <https://doi.org/10.1093/geronb/gbae030>
- Sischka, P. E., Décieux, J. P., Mergener, A., Neufang, K. M., & Schmidt, A. F. (2022). The impact of forced answering and reactance on answering behavior in online surveys. *Social Science Computer Review*, 40(2), 405–425. <https://doi.org/10.1177/0894439320907067>
- Staliunaite, I.R., Valvoda, J. & Satoh, K. (2024) Comparative Study of Explainability Methods for Legal Outcome Prediction. NLLP 2024 - Natural Legal Language Processing Workshop 2024, Proceedings of the Workshop, pp. 243–258.
- Vehovar, V., Couper, M. P., & Čehovin, G. (2023). Alternative layouts for grid questions in PC and mobile web surveys: An experimental evaluation using response quality indicators and survey estimates. *Social Science Computer Review*, 41(6), 2122–2144. <https://doi.org/10.1177/08944393221132644>
- Wang, Y., Chen, X., & Zhou, X. (2024). A new method for identifying low-quality data in perceived usability crowdsourcing tests: Differences in questionnaire scores. *International Journal of Human-Computer Interaction*, 40(22), 7297–7313. <https://doi.org/10.1080/10447318.2023.2263694>
- Wind, S. A., & Lugu, B. (2024). Combining nonparametric and parametric item response theory to explore data quality: Illustrations and a simulation study. *Applied Measurement in Education*, 37(2), 109–131. <https://doi.org/10.1080/08957347.2024.2345592>