

Effect of the representation of information on the statistical literacy level of high school students in fake news

Efecto de la representación de la información en el nivel de alfabetización estadística de alumnos de Bachillerato en noticias falseadas

DAVID MOLINA MUÑOZ

Dpto. Didáctica de la Matemática
Facultad Ciencias de la Educación
Universidad de Granada (España)
Campus de Cartuja, s/n, 18071, Granada
dmolinam@ugr.es
<https://orcid.org/0000-0002-7139-9173>

ANA ALCALÁ NAVARRETE

Dpto. Didáctica de la Matemática
Facultad Ciencias de la Educación
Universidad de Granada (España)
Campus de Cartuja, s/n, 18071, Granada
anaalcala@correo.ugr.es

JOSÉ MIGUEL CONTRERAS GARCÍA

Dpto. Didáctica de la Matemática
Facultad Ciencias de la Educación
Universidad de Granada (España)
Campus de Cartuja, s/n, 18071, Granada
jmcontreras@ugr.es
<https://orcid.org/0000-0001-6821-0563>

ELENA MOLINA PORTILLO

Dpto. Didáctica de la Matemática
Facultad Ciencias de la Educación
Universidad de Granada (España)
Campus de Cartuja, s/n, 18071, Granada
elemo@ugr.es
<https://orcid.org/0000-0002-9955-3080>

DOI: <https://doi.org/10.24197/st.2.2022.165-185>

RECIBIDO: 12/01/2022

ACEPTADO: 15/06/2022

Resumen: En la actualidad, el impacto de las noticias falseadas ha aumentado de forma exponencial debido, en parte, a la facilidad con

Abstract: Currently, the impact of fake news has increased exponentially due, in part, to the ease with which it can be disseminated through

que pueden difundirse por medio de las redes sociales. Esto convierte a los jóvenes, usuarios mayoritarios de las redes sociales, en el grupo más expuesto a este tipo de noticias, las cuales suelen incluir información numérica representada de distintas formas. En este trabajo se pretende estudiar el efecto de la representación (frecuencia, gráfica, porcentaje o probabilidad) en que se presenta una determinada información sesgada en el nivel de alfabetización estadística de estudiantes de Bachillerato. Los resultados muestran un bajo nivel de alfabetización estadística de los estudiantes independientemente de la forma en que venga dada la información, con una ligera mejoría en el caso en que la información se presente en forma de frecuencia.

Palabras clave: alfabetización estadística; Bachillerato; noticias falseadas; representación; información.

social networks. This makes young people, the majority users of social networks, the group most exposed to this type of news, which usually includes numerical information represented in different ways. The aim of this work is to study the effect of the representation (frequency, graph, percentage or probability) in which certain biased information is presented on the level of statistical literacy of Baccalaureate students. The results show a low level of statistical literacy of the students regardless of the form in which the information is given, with a slight improvement in the case in which the information is presented in the form of frequency.

Keywords: statistical literacy; Baccalaureate; fake news; representation; information.

1. INTRODUCCIÓN

La presencia, cada vez más frecuente, de grandes volúmenes de información en distintos contextos del día a día ha convertido el tratamiento de los datos y la adecuada interpretación de los mismos en habilidades imprescindibles para los ciudadanos del presente (Biehler et al., 2018). En efecto, la comprensión de datos relativos a temas de marcada relevancia social como pueden ser la migración, el calentamiento global, la salud o la pobreza es fundamental para una participación plena en la sociedad actual (Engel, 2017).

Esta omnipresencia de los datos plantea nuevos retos sobre las necesidades formativas de los individuos. Uno de los más importantes se refiere a la detección de las conocidas como fake news, término que suele traducirse al castellano como bulos, noticias falsas, o, incluso, tal y como sugieren autores como Rodríguez Pérez (2019), desinformación. Allcott y Gentzkow (2017) definen las fake news como noticias intencionalmente falsas que pueden confundir a quienes las leen. En la actualidad, el alcance de las fake news es tal que más de la mitad de los españoles mayores de 15 años aseguran encontrar información sesgada o falsa cada día o casi cada día (Comisión Europea, 2018).

Las fake news han encontrado en las redes sociales un aliado que permite su rápida propagación a lo largo y ancho de, prácticamente, todo el globo. De hecho, solo el 27% de los españoles confían en las noticias e información que se difunde a través de las redes sociales, siendo este porcentaje del 74% para la radio y del 57% para la televisión (Comisión Europea, 2018). Pese a ello, las redes sociales son la principal fuente de información de determinados grupos de población, como el de adolescentes y jóvenes (Fernández-García, 2017; Gottfried y Shearer, 2017), lo que implica una especial exposición de estos a las noticias falsas. A pesar de esta sobreexposición a las fake news, casi 1 de cada 4 adolescentes (22%) afirma no sentirse seguro a la hora de detectarlas (Comisión Europea, 2018). Se hace necesaria, por tanto, una educación mediática para dotar a los jóvenes de las herramientas necesarias para desarrollar una actitud crítica frente a los mensajes difundidos por los medios de comunicación. El conjunto de las habilidades que se requieren en este proceso recibe el nombre de alfabetización mediática (Bulger y Davison, 2018; Pérez Tornero, 2008). La alfabetización mediática está estrechamente relacionada con otros tipos de alfabetización como, por ejemplo, la estadística (Watson, 1997). En los casos en los que las noticias vienen acompañadas de datos numéricos o representaciones gráficas, la información que contienen se suele considerar como cierta por el simple hecho de venir avalada por elementos estadísticos (Hofstadter, 2008). No obstante, en ocasiones, el individuo es víctima del abuso de los datos, debido a que acepta de forma incuestionable su uso para explicar cualquier fundamentación aprovechando la inalterable objetividad de los mismos (Bauman, 2005). Es en estas situaciones donde se pone de manifiesto la importancia de fomentar la alfabetización estadística de los individuos, entendida

como la habilidad para interpretar, evaluar de forma crítica y comunicar información estadística (Gal, 2002). Esta definición revela uno de los principales puntos en común que tienen la alfabetización mediática y la alfabetización estadística: el papel central que desempeña el pensamiento crítico en ambas. Tan estrecha es la relación que algunos autores, incluso, consideran la alfabetización estadística como una parte o un aspecto de la alfabetización mediática (Poljičak Sušec et al., 2014).

Los creadores de fake news utilizan distintas estrategias para dotarlas de credibilidad. Algunas de ellas pasan por la elección de aspectos como la representación de los datos o del contenido y la ubicación de las imágenes, ilustraciones u otros elementos visuales, especialmente los de naturaleza cuantitativa o estadística (Mehta y Guzmán, 2018). Igualmente, se elige con cuidado el formato de distribución de la noticia, siendo el texto y el texto con imagen los formatos más frecuentemente usados (Gutiérrez-Coba et al., 2020). Cuando los textos de las fake news incorporan valores numéricos, estos pueden presentarse de diversas formas. La elección del modo de representación de los datos en las noticias falseadas suele atender más a criterios sensacionalistas que a criterios periodísticos o informativos (Fernández-García, 2017). De hecho, los medios de comunicación aprovechan las posibles carencias o dificultades de los lectores en la interpretación de frecuencias (Álvarez Alfonso et al., 2020; Gea et al., 2017), porcentajes (Bassan Cincinatus y Sheffét, 2016; Knapp, 2009) o probabilidades (Scheufele y Krause, 2019) en busca del formato numérico que mayor impacto genere. Además de las representaciones numéricas, en las noticias suelen aparecer elementos de carácter visual. Las representaciones gráficas resultan más atractivas que el texto (Añaños y Astals, 2013) y la información que recogen es más fácil de entender e interpretar que la dada en formato numérico (Meerman Scott, 2015), por lo que son frecuentemente utilizadas por los medios de comunicación a la hora de transmitir información. Pero, en ocasiones, los elementos que componen los gráficos son alterados con el objeto de engañar, confundir o condicionar a aquellos que los tienen que interpretar (Cairo, 2015).

Varios estudios han comparado el efecto de distintos formatos de presentación de la información en el nivel de comprensión de la misma. Sin embargo, no existe un consenso claro sobre cuál es la mejor forma de representar los datos para facilitar la comprensión de la información que se quiere transmitir. Así, Sinayev et al. (2015) concluyeron que la información dada a través de porcentajes se comprende mejor que la que viene dada mediante frecuencias. Por el contrario, autores como Gigerenzer y Edwards (2003) recomiendan el uso de frecuencias en lugar de porcentajes para transmitir información numérica. Prasad y Ojha (2012), por su parte, obtuvieron que la tasa de comprensión de una determinada información numérica es mayor cuando esta viene expresada en formato texto que cuando se utiliza un gráfico para representarla. La información que se presenta como estímulo en la mayoría de estos trabajos está relacionada con

el campo de la salud y los grupos de población estudiados dejan fuera a los adolescentes. El presente trabajo viene a cubrir este hueco, puesto que tiene por objetivo analizar la influencia que tienen diferentes tipos de representaciones de un mismo dato en la interpretación de noticias sesgadas en alumnos de Bachillerato.

2. MARCO TEÓRICO

Durante las últimas décadas se viene realizando una promoción y fomento de la alfabetización estadística desde distintas instituciones (Organización para la Cooperación y el Desarrollo Económico (OCDE), International Association for the Evaluation of Educational Achievement (IEA), entre otras) así como desde la investigación en distintas áreas de conocimiento, entre ellas de la educación estadística (Alsina et al., 2020; Engel, 2017; Gal, 2002; Rodríguez-Muñiz et al., 2020). Dicho impulso se orienta a destacar el valor de la estadística para ser ciudadanos activos en la sociedad, favoreciendo su inclusión en las normativas curriculares españolas, entre muchas otras, y el refuerzo del proceso de su enseñanza y aprendizaje desde la escuela.

Asimismo, en relación a la alfabetización estadística que se pretende evaluar, mucho del conocimiento o información que se concluye de datos recogidos o disponibles para ser analizados, procede de fuentes heterogéneas, se transmite por diferentes medios de comunicación y aparece presentada mediante diversas formas de representación tales como frecuencias, porcentajes, probabilidad y gráficos.

En consecuencia, tanto en el currículo de Educación Primaria como en el de Educación Secundaria y en la mayor parte de itinerarios de Bachillerato aparecen elementos curriculares relacionados con tales formas de representación, y la transnumeración entre ellas, enfatizando en la interpretación, postura crítica y justificación de la información que mediante estas se presenta.

Como indica Rumsey (2002), el término alfabetización es demasiado amplio, destacando en la conceptualización del constructo el modelo de Gal (2002). En él se distinguen dos componentes básicos interrelacionados: una primera relacionada con la capacidad de interpretar y evaluar críticamente la información estadística y, la segunda, relacionada con la capacidad para discutir o comunicar opiniones respecto a dicha información. Este modelo se concreta en 7 facetas incluyendo destrezas estadísticas y matemáticas, habilidades básicas de lectura, conocimiento del contexto, capacidad crítica hacia los conocimientos estadísticos, junto a las disposiciones creencias y actitudes hacia el uso de la estadística.

Posteriormente se han realizado nuevos modelos (Nicholson et al., 2018; Watson y Callingham, 2003) que contemplan nuevos matices al respecto, pero en concordancia con las directrices de las definiciones previas.

3. MATERIAL Y MÉTODOS

La muestra seleccionada para el estudio consta de 64 individuos, 30 chicos y 34 chicas, estudiantes de primer curso de Bachillerato de las modalidades de Ciencias y Tecnología y de Ciencias Sociales y Humanidades de dos centros públicos de la región de Andalucía (España). La muestra se seleccionó siguiendo un muestreo de voluntarios. El estudio se llevó a cabo durante el curso académico 2019-2020.

Con base en el modelo de Gal, como instrumentos de recogida de datos se han elaborado cuatro cuestionarios (a los que nos referiremos como A, B, C y D). En cada uno de ellos aparecen las mismas cuatro noticias con información falsa o fake news, con diferencia en las formas de representación de la información numérica contenida en ellas. En concreto, para cada noticia se ha representado la misma información a través de frecuencias absolutas, de probabilidades, de porcentajes y de un gráfico, correspondiéndose una de estas representaciones a la originalmente incluida en la noticia. En el Anexo I se incluyen las cuatro representaciones que se han obtenido para cada una de las cuatro noticias empleadas en el estudio.

En cada cuestionario se ha incluido una representación de la información distinta para cada noticia teniendo en cuenta que cada formato apareciera una única vez por cuestionario. En concreto, la distribución es la que se muestra en la Tabla 1.

Tabla 1

Distribución por cuestionarios de las representaciones de cada noticia

Cuestionario	Noticia 1	Noticia 2	Noticia 3	Noticia 4
A	Frecuencia	Probabilidad	Porcentaje	Gráfico
B	Porcentaje	Frecuencia	Gráfico	Probabilidad
C	Probabilidad	Gráfico	Frecuencia	Porcentaje
D	Gráfico	Porcentaje	Probabilidad	Frecuencia

En todos los modelos de cuestionarios utilizados para cada noticia, tras la representación de la información numérica correspondiente, se incluyeron 7 ítems, los cuales suponen una adaptación de los diseñados por Contreras et al. (2017) para la evaluación crítica de representaciones gráficas. Dichos ítems son los que siguen:

- 1) Haz un resumen de la noticia indicando los datos más relevantes y justificando su relevancia.

- 2) ¿Sabrías indicar si los datos son relevantes en el contexto de la noticia? Justifica la respuesta.
- 3) ¿Es fiable la fuente de procedencia de los datos? Justifica la respuesta.
- 4) ¿Qué opinión te merece la noticia?
- 5) ¿Los datos son apropiados para explicar la información? Justifica la respuesta. En caso contrario responde cómo los representarías.
- 6) ¿Sabrías explicar cómo se han calculado esos valores?
- 7) ¿Crees que es correcta la información?

Las respuestas de los estudiantes a estos ítems permiten determinar sus habilidades en ámbitos fundamentales de la alfabetización estadística como, por ejemplo, la interpretación contextualizada de la noticia o el conocimiento de los elementos matemáticos que aparecen en ella (Gal, 2002). Se presta especial interés a la evaluación de la capacidad de los estudiantes para cuestionar los datos que se les presentan, para analizarlos desde un punto de vista crítico y para justificar y argumentar las respuestas que proporcionan. A su vez, la comparación entre las respuestas recogidas en los distintos cuestionarios, posibilita la comprobación del efecto del formato de representación en la interpretación de las noticias.

Al tratarse de ítems de respuesta abierta, cada una de estas respuestas se ha clasificado como correcta, parcialmente correcta o incorrecta atendiendo a los siguientes criterios: se considera que una respuesta a un ítem es correcta cuando está justificada y, además, se refleja en ella el carácter falso o engañoso de la noticia; se considera que una respuesta a un ítem es parcialmente correcta cuando está justificada pero no se identifica la noticia como fraudulenta o cuando, por el contrario, aun señalando la noticia como falsa, la respuesta no viene acompañada de una justificación; en caso de que una respuesta no esté justificada ni se indique en ella el carácter falso o engañoso de la noticia, se considera incorrecta.

Siguiendo una metodología similar a la empleada por Kahne y Bowyer (2017), cada estudiante cumplimentó uno de los cuatro cuestionarios, que le fue asignado de forma aleatoria. En total, los cuestionarios A, B, C y D fueron respondidos por 17, 14, 16 y 17, alumnos, respectivamente.

La información proporcionada por los estudiantes se ha analizado a través de técnicas descriptivas (porcentajes, representados mediante gráficos de barras). En el análisis se han considerado únicamente las respuestas válidas.

4. RESULTADOS

A continuación se muestran los resultados del análisis por separado de las respuestas de los 7 ítems a cada noticia en función del formato de representación de la información utilizado.

En términos generales, como indica la Figura 1, el número de respuestas correctas en los ítems relativos a la noticia 1, que trata sobre la mutilación genital femenina, es muy bajo independientemente del tipo de representación considerado.

Se aprecia que los alumnos tienen más dificultades para resumir y sintetizar los datos de esta noticia cuando vienen representados mediante un gráfico (obteniéndose un 41,2% de respuestas incorrectas y un 47,1% de respuestas parcialmente correctas en el ítem 1) y que los resumen ligeramente mejor cuando vienen dados a través de frecuencias y porcentajes (en cuyo caso la tasa de respuestas incorrectas es del 11,7% y del 7,7% y la de respuestas parcialmente correctas asciende hasta el 82,4% y el 92,3%, respectivamente, en el mismo ítem). La representación en forma de probabilidad de los datos numéricos dificulta considerablemente la identificación de los datos más relevantes de la noticia (ítem 2) así como la expresión de una opinión crítica (ítem 4) a la mayoría de los estudiantes, con más del 60% de respuestas erróneas en ambos casos. Esto puede ser debido a que no están tan familiarizados con la probabilidad como con otros conceptos matemáticos. Además, son muy pocos los casos en los que se ha tenido en cuenta el dato de la población mundial para la verificación de la información de la noticia. En este mismo sentido, la probabilidad es la forma de representación de los datos que genera menos respuestas incorrectas en el ítem 5, lo que quiere decir que los alumnos identifican la probabilidad como un formato poco apropiado para explicar la información que se quiere transmitir. Sin embargo, esto podría deberse a que presentan dificultades a la hora de comprender este concepto matemático y, por ello, lo ven poco apropiado para explicar la información.

Según la Figura 2, en la noticia 2 que informa sobre muertes por arma de fuego de niños en Estados Unidos, y al igual que lo que ocurría en el caso anterior, la tasa de respuestas correctas es muy discreta, independientemente de la representación de la información considerada.

El uso de las frecuencias para representar los datos hace que los estudiantes resuman la noticia de forma algo más precisa en comparación con el resto de representaciones. Pero, en el extremo opuesto, la frecuencia destaca, junto a la gráfica, como la representación que tiene asociados los peores resultados en el ítem 4, en el que se pedía a los estudiantes que expusieran su opinión, debidamente justificada, sobre la noticia. De hecho, todas las respuestas dadas por los estudiantes a los cuales se les presentó la información de la noticia en forma de frecuencia o en forma de gráfico fueron incorrectas, al carecer de una argumentación crítica. En esta misma línea, la frecuencia es el formato más engañoso, en el sentido que es el que los estudiantes consideran más apropiado para representar los datos, según las respuestas al ítem 5. Esto quiere decir que, cuando los datos vienen dados en forma de frecuencia, la noticia, aun siendo falsa, adquiere un nivel de veracidad, superior al que proporcionan las otras representaciones estudiadas. Además, la frecuencia y la gráfica son las representaciones que, en mayor grado, consiguen hacer creer a los estudiantes que la información dada es correcta, como indican los porcentajes de

respuestas incorrectas al ítem 7, que superan el 80% para las dos representaciones. Aunque los alumnos comprenden estos sistemas de representación con facilidad, no los interpretan en el contexto dado de forma correcta, ya que no perciben el significado de la duplicación anual de las muertes ni sus implicaciones en un periodo de tiempo de varios años.

Por otro lado, es destacable la dificultad que tienen los alumnos al determinar la relevancia de los datos. Tanto es así que ningún estudiante ha respondido de forma correcta al ítem 2, al no tener en cuenta la evolución temporal del dato que se incluye en la noticia a la hora de determinar la validez de la información. Se observan también errores importantes en las explicaciones que proporcionan los estudiantes sobre el procedimiento de cálculo de los datos, siendo la probabilidad el formato con más respuestas equivocadas (76,5%).

En línea con los resultados de las noticias anteriores, las respuestas de los alumnos a los ítems referidos a la noticia 3, referente al número de muertes por violencia de género en España por comunidades autónomas, son predominantemente incorrectas, como se puede apreciar en la Figura 3.

La representación gráfica es la que más problemas presenta a los estudiantes de cara a la síntesis de la noticia (ítem 1). Las respuestas de todos los estudiantes que resumieron la noticia en base al gráfico fueron incorrectas, pues en ningún caso se tuvo en cuenta el número de habitantes de cada región al determinar la relevancia de la información de la noticia. Por este mismo motivo, los estudiantes no perciben lo inapropiado del gráfico para representar los datos de la noticia (ítem 5), ya que consideran que los habitantes de las regiones a las que corresponden colores más oscuros son más susceptibles de convertirse en una víctima de violencia de género. Al igual que con el gráfico, apenas se señala la probabilidad como una representación poco conveniente para la información de la noticia, a pesar de tratarse, en este caso, de valores muy pequeños que pueden no reflejar adecuadamente el alcance de los datos. Precisamente la pequeña magnitud de las probabilidades y de los porcentajes hace que los estudiantes sean incapaces, en gran parte de los casos, de determinar la relevancia de los datos en el contexto de la noticia cuando se enfrentan a estos tipos de representaciones, como reflejan las respuestas al ítem 2. Además, son estas dos representaciones las que más dificultan a los estudiantes la construcción de una opinión apropiadamente argumentada, tal y como indican las respuestas al ítem 4 (con una tasa del 62,5% y del 71,4% de respuestas incorrectas cuando la información se da en forma de probabilidades y de porcentajes, respectivamente). Los valores tan bajos de porcentajes y probabilidades que se proporcionan hacen que los alumnos no sean conscientes de la trascendencia de la noticia. No tienen en cuenta el contexto, no analizan la información expuesta ni verifican la veracidad de la misma por lo que difícilmente pueden involucrarse directamente con el significado social de la noticia. Por otra parte, se observa que la probabilidad, por encima del resto de representaciones, hace que los estudiantes consideren que la información recogida en la noticia es adecuada (ítem 7). En la

mayoría de casos, además, la respuesta que se da no viene acompañada de ningún tipo de justificación, de lo que se deduce una escasa comprensión de dicho sistema de representación y que en estos casos los alumnos utilizan la intuición.

Finalmente, cabe destacar que, aunque la tasa de respuestas incorrectas ante el ítem que pide a los estudiantes explicar cómo se han calculado los valores que se incluyen en las noticias es muy parecida (ítem 6), independientemente de la forma en que se haya representado la información, un 25% de los estudiantes a los que se presentó la información en forma de frecuencias describieron el procedimiento de cálculo de estas frecuencias de forma acertada.

Al igual que se viene observando en las noticias anteriores, la Figura 4 muestra que la mayoría de las respuestas dadas por los estudiantes a los ítems referidos a la noticia 4, que muestra la información sobre violaciones en Estados Unidos, son incorrectas o, a lo sumo parcialmente correctas, independientemente del sistema de representación de los datos. Se aprecia, también, que ninguno de los formatos de representación de la información destaca uniformemente por encima del resto en todos los ítems cuanto a tasa de respuestas incorrectas se refiere.

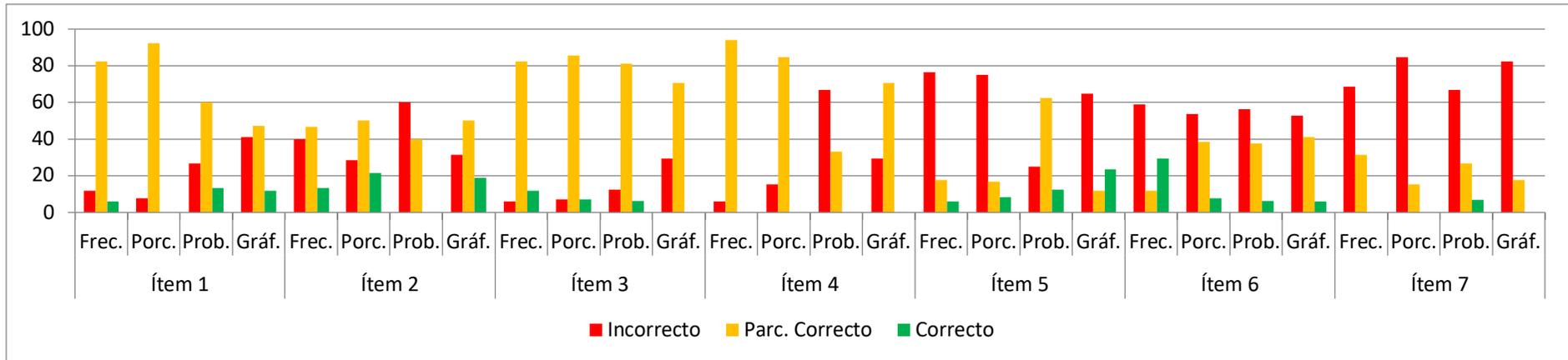


Figura 1. Respuestas incorrectas, parcialmente correctas y correctas a los ítems de la noticia 1, en función del tipo de representación.

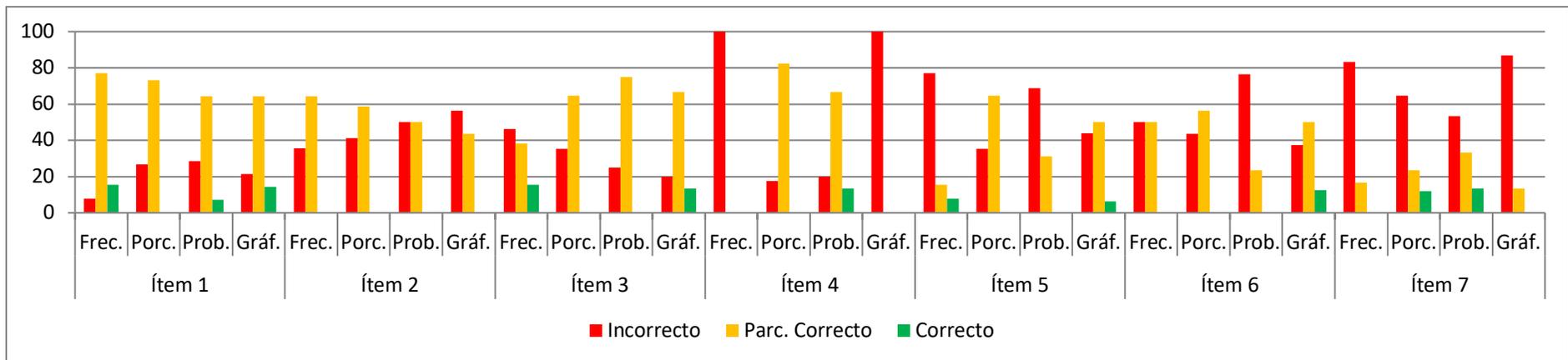


Figura 2. Respuestas incorrectas, parcialmente correctas y correctas a los ítems de la noticia 2, en función del tipo de representación.

Cuando la información de la noticia viene expresada en términos probabilísticos, la mitad de los estudiantes proporciona una respuesta incorrecta en el ítem 1. No son capaces de justificar la relevancia del dato dado por la probabilidad al considerar que es un valor muy pequeño o, incluso, insignificante. En este caso, la otra mitad de los estudiantes solo alcanza a proporcionar una respuesta parcialmente correcta. El formato probabilístico también confunde de forma importante a los participantes a la hora de determinar si la información de la noticia es correcta o no (ítem 7), ya que más del 90% de los participantes no puede detectar el carácter erróneo de la noticia cuando el dato que la acompaña es una probabilidad.

Por otro lado, el diagrama de sectores es la representación que mejores resultados proporciona en el ítem 2, con un tercio de respuestas correctas. Este gráfico ayuda a los estudiantes a darse cuenta de lo desmedido del dato de la noticia en el contexto dado. En el extremo opuesto, cuando el mismo dato aparece en forma de porcentaje, la mitad de los estudiantes no perciben dicha desproporción. De hecho, en las justificaciones que proporcionan en estos casos, los estudiantes no tienen en cuenta, ni siquiera de forma estimada, el tamaño de la población de mujeres de Estados Unidos, lo que dificulta descubrir lo exagerado del dato y, en consecuencia, lo sesgado de la noticia. En la misma línea, las opiniones críticas de los alumnos sobre la noticia son más imprecisas, en el sentido de que carecen de una justificación sólida, cuando el dato numérico viene dado como un porcentaje (ítem 4). No obstante, según las respuestas al ítem 5, los estudiantes identifican el porcentaje como el formato menos adecuado para la representación de la información de la noticia, alegando que simboliza una parte de un total que no acaba de concretarse numéricamente.

Para finalizar, notar la dificultad de los estudiantes para explicar el cálculo de los datos, en especial de los expresados en forma de porcentaje y de probabilidad, con más de dos tercios de respuestas incorrectas en cada caso.

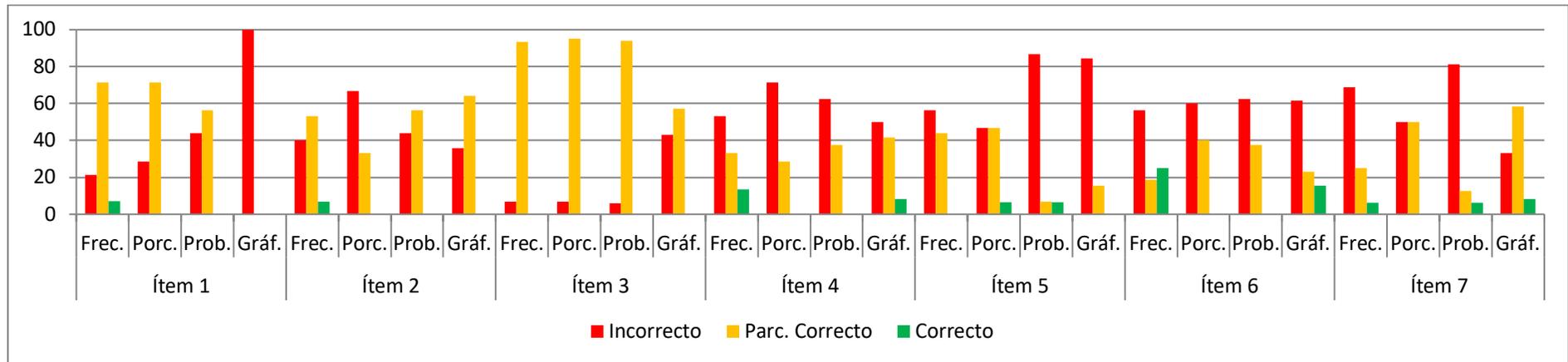


Figura 3. Respuestas incorrectas, parcialmente correctas y correctas a los ítems de la noticia 3, en función del tipo de representación.

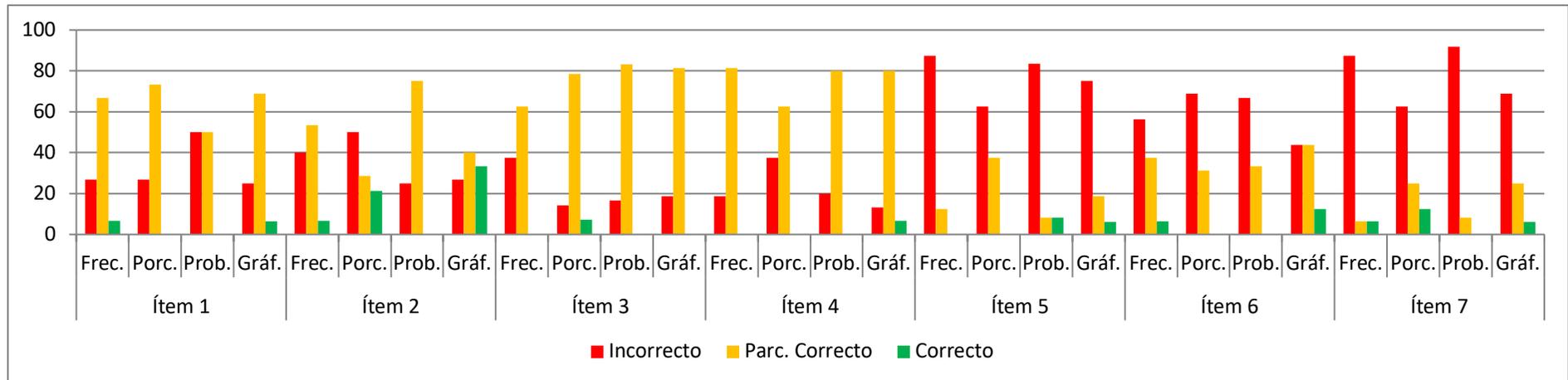


Figura 4. Respuestas incorrectas, parcialmente correctas y correctas a los ítems de la noticia 4, en función del tipo de representación.

5. DISCUSIÓN Y CONCLUSIONES

A través del presente estudio se pretende indagar en el efecto que tiene el formato en el que viene representada la información estadística de diversas fake news en el grado de interpretación de las mismas y, consecuentemente, en el nivel de alfabetización estadística, de estudiantes de Bachillerato.

De forma global, y dada la baja tasa de respuestas correctas dadas por los participantes, podemos concluir un bajo nivel de alfabetización estadística de los estudiantes de la muestra independiente del formato de representación considerado. Los alumnos conocen los conceptos de frecuencia, porcentaje, probabilidad y gráfico estadístico, pero presentan dificultades en su comprensión cuando se utilizan en contextos sociales. En muchos casos, los estudiantes no han sido capaces de acompañar sus respuestas con argumentaciones o justificaciones válidas y, lo que es más preocupante aún, no han identificado el carácter sesgado de la noticia.

Se han detectado ciertas diferencias en función de la representación empleada en cada noticia. En general, los estudiantes proporcionan respuestas más certeras cuando la información que incorpora la noticia viene dada como una frecuencia, hecho que puede atribuirse a la sencillez del formato y a la familiaridad que los estudiantes tienen con esta forma de representación de la información en comparación con el resto. En el otro extremo, en línea con lo expuesto por Engel (2017), la probabilidad resulta ser el formato que peores resultados obtiene. Los alumnos tienen problemas al interpretar probabilidades ya que las consideran siempre como valores pequeños, por lo que no perciben su verdadera magnitud en el contexto social de la noticia. Pero, como señala Engel (2017), muchas veces, un número aparentemente pequeño puede referirse a una cantidad elevada en un determinado contexto. Algo similar ocurre cuando los valores de los porcentajes son pequeños, como los que aparecen en la noticia 3.

A la hora de resumir la noticia, la mayoría de los alumnos se limita a transcribirla. Muy pocos alumnos han empleado un lenguaje propio ni han justificado la relevancia de los datos. Pese a que, en muchos casos, identifican que los datos son relevantes a nivel social (indicando, por ejemplo, que son muchas las niñas sometidas a ablación o que el número de niños abatidos por un arma de fuego es muy alto) no llegan a comprobar si esos datos son realmente ciertos teniendo en cuenta la población a la que se refiere la noticia. En este sentido, y tal y como ya apuntó Gal (2019), podemos concluir que sus destrezas lingüísticas están condicionadas en función del dato matemático utilizado en cada noticia.

Un error común cuando se trata de determinar la fiabilidad de la procedencia de datos es confundir la fuente de los datos con el medio en el que se publica la noticia. Muchos de los individuos de la muestra piensan que la información es fiable por el mero hecho de estar publicada un medio de comunicación. En este caso es del medio del que se fían y no de la fuente en sí. Esta confianza en el medio de comunicación hace que los alumnos fallen al considerar que la información de la

noticia es correcta, independientemente del tipo de representación de la información que se utilice.

En las escasas ocasiones en que los estudiantes detectan que la representación usada no es la más conveniente para explicar la noticia, apenas proponen alternativas más apropiadas para ello. Además, presentan dificultades para explicar cómo se ha obtenido los datos, sobre todo los dados en forma de probabilidad y porcentaje, puesto que no se preguntan por el tamaño de la población a la que se refiere la noticia.

Se ha podido observar, también, una alta tendencia al uso de las intuiciones por parte de los estudiantes, respondiendo a cada pregunta en función de estas y de las opiniones generadas por las mismas en el contexto social correspondiente.

Por otro lado, durante la realización de este estudio se han encontrado ciertas limitaciones que restringen el alcance de los hallazgos obtenidos. En primer lugar, el tamaño de la muestra es ciertamente reducido y, además, los individuos que la componen han sido seleccionados mediante un esquema muestral no probabilístico. En consecuencia, los resultados observados no pueden extrapolarse a una población general de estudiantes de Bachillerato. No obstante, a partir de los resultados de este trabajo se pueden vislumbrar las dificultades relacionadas con la interpretación de datos estadísticos dados en diferentes formatos que tienen los estudiantes de esta etapa al interpretar noticias de medios de comunicación y, a su vez, las necesidades formativas de los mismos.

En conclusión, los participantes del estudio tienen importantes carencias en la detección de noticias falsas o imprecisas, lo que es síntoma de un bajo nivel de alfabetización estadística. Para evitarlo éstos deben formarse para ser capaces de ver el conjunto de datos como un todo, describir su forma y usar sus características para comparar la información intrínseca en los datos. De esta manera podrán analizar si los datos proporcionan la información necesaria para responder a las cuestiones por la que fueron formulados. Es aquí donde la alfabetización estadística es fundamental para que los estudiantes puedan dar sentido y tomar decisiones sensatas sobre la gran cantidad de datos que encuentran en su vida diaria (Chick y Pierce, 2012). Si, pese a ello, estas carencias no son corregidas, estos individuos tendrán dificultades importantes para ser ciudadanos críticos ante un mundo en la que la proliferación de información sesgada es un hecho.

REFERENCIAS BIBLIOGRÁFICAS.

Allcott, H. y Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236.
<https://doi.org/10.1257/jep.31.2.211>

Alsina, Á., Vásquez, C. A., Muñiz-Rodríguez, L. y Rodríguez Muñiz, L. (2020).

- ¿Cómo promover la alfabetización estadística y de datos en contexto? Estrategias y recursos a partir de la COVID-19 para Educación Secundaria. *Épsilon - Revista de Educación Matemática*, 104, 99–128.
- Álvarez Alfonso, I., Guerrero Gutiérrez, Y. A. y López, Y. D. T. (2020). Taxonomía de errores y dificultades en la construcción e interpretación de tablas de frecuencia. *Zetetike*, 28, 1–22.
<https://doi.org/10.20396/zet.v28i0.8656553>
- Añaños, E. y Astals, A. (2013). ¿Imagen o texto? El poder de captar la atención visual de los elementos gráficos analizado con el Eye tracker. *Grafica*, 1(2), 87–98.
- Bassan Cincinatus, R. y Sheffet, M. (2016). “With percentages the 100 is always in the denominator”: From the field to pre-service teachers. *International Journal of Research in Education and Science*, 2(1), 143–155.
<https://doi.org/10.21890/ijres.75271>
- Bauman, Z. (2005). *Modernity and ambivalence*. Polity Press.
- Biehler, R., Frischmeier, D., Reading, C. y Shaughnessy, J. M. (2018). Reasoning About Data. In *International Handbook of Research in Statistics Education* (pp. 139–192). Springer, Cham. https://doi.org/10.1007/978-3-319-66195-7_5
- Bulger, M. y Davison, P. (2018). The Promises, Challenges, and Futures of Media Literacy. *Journal of Media Literacy Education*, 10(1), 1–21.
<https://doi.org/10.23860/jmle-2018-10-1-1>
- Cairo, A. (2015). Graphics lies, misleading visuals: Reflections on the challenges and pitfalls of evidence-driven visual communication. In *New Challenges for Data Design* (pp. 103–116). Springer, London. https://doi.org/10.1007/978-1-4471-6596-5_5
- Chick, H. L. y Pierce, R. (2012). Teaching for statistical literacy: utilising affordances in real-word data. *International Journal of Science and Mathematics Education*, 10(2), 339–362. <https://doi.org/10.1007/S10763-011-9303-2>
- Comisión Europea (2018). Fake news and disinformation online. *Recuperado de* <https://europa.eu/eurobarometer/api/deliverable/download/file?deliverableId=65673>

- Contreras, J. M., Molina-Portillo, E., Godino, J. D. y Batanero, C. (2017). Construcción de un cuestionario para evaluar la interpretación crítica de gráficos estadísticos por futuros profesores. In J. M. Muñoz, A. Arnal-Bailera, P. Beltrán-Pellicer, M. L. Callejo y J. Carrillo (Eds.), *Investigación en Educación Matemática XXI* (pp. 207–216).
- Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal*, 16(1), 44–49.
- Fernández-García, N. (2017). Fake news: una oportunidad para la alfabetización mediática. *Nueva Sociedad*, 269, 66–77.
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1–25. <https://doi.org/10.1111/j.1751-5823.2002.tb00336.x>
- Gal, I. (2019). Understanding statistical literacy: About knowledge of contexts and models. *Comprensión de la cultura estadística (alfabetización estadística): Sobre el conocimiento de contextos y modelos*. In J. M. Contreras, M. M. Gea, M. M. López-Martínez y E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística*.
- Gea, M. M., Arteaga, P. y Cañadas, G. R. (2017). Interpretación de gráficos estadísticos por futuros profesores de Educación Secundaria. *Avances de Investigación en Educación Matemática*, 12, 19–37. <https://doi.org/10.35763/aiem.v1i12.189>
- Gigerenzer, G. y Edwards, A. (2003). Simple tools for understanding risks: from innumeracy to insight. *BMJ*, 327, 741–744. <https://doi.org/10.1136/BMJ.327.7417.741>
- Gottfried, J. y Shearer, E. (2017). *Americans' Online News Use is Closing in on TV News Use*. Pew Research Center. <http://www.pewresearch.org/fact-tank/2017/09/07/americans-online-news-use-vs-tv-news-use/>.
- Gutiérrez-Coba, L. M., Coba-Gutiérrez, P. y Gómez-Díaz, J. A. (2020). Noticias falsas y desinformación sobre el Covid-19: análisis comparativo de seis países iberoamericanos. *Revista Latina de Comunicación Social*, 78, 237–264. <https://doi.org/10.4185/RLCS-2020-1476>
- Hofstadter, D. (2008). *Metamagical themas: Questing for the essence of mind and pattern*. Basic Books.

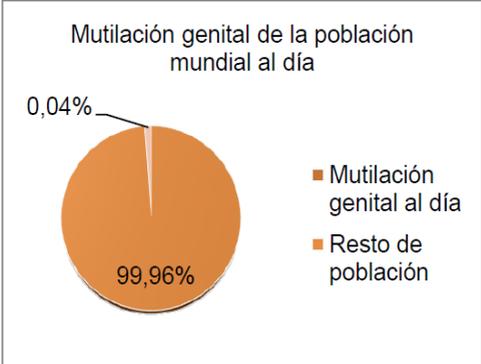
- Kahne, J. y Bowyer, B. (2017). Educating for Democracy in a Partisan Age: Confronting the Challenges of Motivated Reasoning and Misinformation. *American Educational Research Journal*, 54(1), 3–34. <https://doi.org/10.3102/0002831216679817>
- Knapp, T. R. (2009). *Percentages: The most useful statistics ever invented*.
- Meerman Scott, D. (2015). An image is worth a thousand words. In *The new rules of marketing and PR: how to use social media, online video, mobile applications, blogs, news releases, and viral marketing to reach buyers directly* (pp. 310 – 323). John Wiley y Sons, Inc.
- Mehta, R. y Guzmán, L. (2018). Fake or visual trickery? Understanding the quantitative visual rethoric in the news. *Journal of Media Literacy Education*, 10(2), 104–122.
- Nicholson, J., Gal, I. y Ridgway, J. (2018). *Understanding Civic Statistics: A Conceptual Framework and its Educational Applications. A product of the ProCivicStat Project. Available from: <http://IASE-web.org/islp/pcs>*
- Pérez Tornero, J. M. (2008). Media Literacy. New Conceptualisation, New Approach. In *Empowerment through Media Education and Intercultural Dialogue* (1st ed., pp. 103–116). <https://doi.org/https://ddd.uab.cat/record/220411>
- Poljičak Sušec, M., Jerak Muravec, N. y Stančić, H. (2014). Statistical literacy as an aspect of media literacy. *Medijska Istraživanja: Znanstveno-Stručni Časopis Za Novinarstvo i Medije*, 20(2), 131–155.
- Prasad, G. V. R. J. S. y Ojha, A. (2012). Text, table and graph - Which is faster and more accurate to understand? *2012 IEEE 4th International Conference on Technology for Education, T4E 2012*, 126–131. <https://doi.org/10.1109/T4E.2012.18>
- Rodríguez-Muñiz, L. J., Muñiz-Rodríguez, L., Vásquez, C. y Alsina, Á. (2020). ¿Cómo promover la alfabetización estadística y de datos en contexto? Estrategias y recursos a partir de la COVID-19 para Educación Secundaria. *Números - Revista de Didáctica de las matemáticas*, 104, 217-238.
- Rodríguez Pérez, C. (2019). No diga fake news, di desinformación: una revisión sobre el fenómeno de las noticias falsas y sus implicaciones. *Comunicación*, 40, 65–74. <https://doi.org/10.18566/comunica.n40.a05>

- Rumsey, D. J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education*, 10(3).
<https://doi.org/10.1080/10691898.2002.11910678>
- Scheufele, D. A. y Krause, N. M. (2019). Science audiences, misinformation, and fake news. *Proceedings of the National Academy of Sciences of the United States of America*, 116(16), 7662–7669.
<https://doi.org/10.1073/pnas.1805871115>
- Sinayev, A., Peters, E., Tusler, M. y Fraenkel, L. (2015). Presenting Numeric Information with Percentages and Descriptive Risk Labels: A Randomized Trial. *Medical Decision Making*, 35(8), 937–947.
<https://doi.org/10.1177/0272989X15584922>
- Watson, J. y Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, 2(2), 3–46.
- Watson, J. (1997). Assessing statistical literacy using the media. In I. Gal y J. B. Garfield (Eds.), *The Assessment Challenge in Statistics Education* (pp. 107–121). IOS Press and The International Statistical Institute.
<https://doi.org/10.2307/2685944>

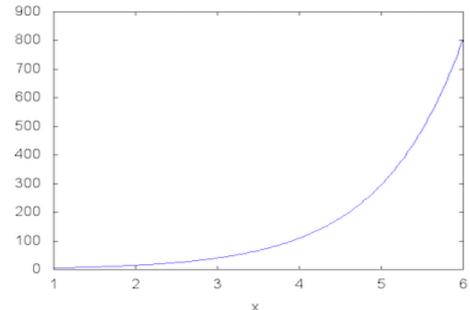
ANEXO I. Representaciones de las noticias usadas en los distintos cuestionarios.

Representaciones de la noticia 1, referida a la mutilación genital femenina. Fuente: La Sexta.

Frecuencia	Porcentaje El 0,04% de la población mundial puede sufrir mutilación genital al día.
	Probabilidad La probabilidad de que alguien sufra una mutilación genital cada día es de 0,0004.

<p>laSexta 6 feb. a las 11:01 • 🌐</p> <p>Casi tres millones de niñas son sometidas a la ablación cada día en el mundo, y 30 son susceptibles de serlo en la próxima década http://atres.red/ftlbnk4</p>  <p>Más de 18.000 niñas pueden sufrir la mutilación genital en España lasexta.com</p> <p>👍👎👏 56 40 comentarios • 24 veces compartido</p>	<p>Gráfico</p> <p>Mutilación genital de la población mundial al día</p>  <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Mutilación genital al día</td> <td>0,04%</td> </tr> <tr> <td>Resto de población</td> <td>99,96%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Mutilación genital al día	0,04%	Resto de población	99,96%
Categoría	Porcentaje						
Mutilación genital al día	0,04%						
Resto de población	99,96%						

Representaciones de la noticia 2, referida a las muertes por arma de fuego de niños estadounidenses. Fuente: NBC.

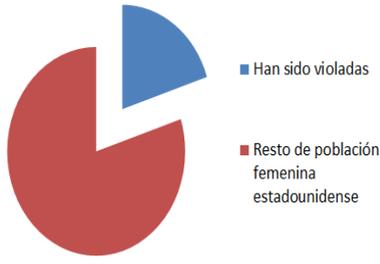
<p>Frecuencia Cada año, desde 1972, el número de niños estadounidenses abatidos por arma de fuego se ha doblado.</p>	<p>Gráfico</p>  <p>Número de niños estadounidenses abatidos por arma de fuego cada año desde 1972 (los números del eje x se refieren a cada año transcurrido).</p>
<p>Porcentaje Desde 1972, cada año se dobla el porcentaje del número de niños estadounidenses abatidos por arma de fuego.</p>	
<p>Probabilidad Desde 1972, cada año aumenta el doble la probabilidad de ser niño estadounidense y ser abatido por arma de fuego.</p>	

Representaciones de la noticia 3, referida a las muertes por violencia de género en España en el periodo 1999-2013. Fuente: eldiario.es.

<p>Frecuencia España alcanza las 898 víctimas mortales por violencia de género entre 1999 y 2013. Entre estas, más de 140</p>	<p>Probabilidad Probabilidad de resultar víctima mortal por violencia de género por comunidades autónomas</p>
--	--

<p>se han identificado en las comunidades autónomas de Andalucía y Cataluña, seguidas de la Comunidad de Madrid y la Comunidad Valenciana, donde se han registrado entre 120 y 140.</p>	<ul style="list-style-type: none"> • Andalucía y Cataluña: más de un 0,156. • Comunidad de Madrid y Comunidad Valenciana: entre un 0,09 y un 0,11.
<p>Porcentaje Un 0,000019% de la población española resultó víctima mortal por violencia de género entre 1999 y 2013. Un 0,000017% de la población andaluza, un 0,000019% de la población catalana y un 0,000013% de la población madrileña resultó víctima mortal por violencia de género entre esos años.</p>	<p>Gráfico 1999/2013 víctimas mortales por comunidades autónomas</p>  <p>898 víctimas mortales</p>

Representaciones de la noticia 4, referida a las violaciones en Estados Unidos.
Fuente: El País digital.

<p>Frecuencias</p>  <p>Una de cada cinco estadounidenses ha sido violada</p> <ul style="list-style-type: none"> ▪ La violencia se ceba en jóvenes ▪ Una de cada cuatro ha sido agredida por su pareja ▪ El caso de las latinoamericanas ▪ España, a la cola en violencia machista 	<p>Gráfico Violaciones en Estados Unidos</p>  <ul style="list-style-type: none"> ■ Han sido violadas ■ Resto de población femenina estadounidense
<p>Porcentaje El 20% de las estadounidenses ha sido violada</p>	
<p>Probabilidad La probabilidad de haber sido violada siendo mujer estadounidense es de 0,20.</p>	