

**Dr. José-Luis ROJAS-TORRIJOS**

Universidad de Sevilla. España. jlrojas@us.es

**Dr. Jesús GARCÍA-CEPERO**

Universidad de Sevilla. España. jesusgcepero@gmail.com

## **Percepción del periodismo deportivo de datos entre usuarios habituales. Estudio de caso del modelo predictivo de El País para el Mundial de Fútbol de 2018**

### ***Perception of sports data journalism among heavy users. Case study: predictive model during the 2018 Football World Cup in El País***

**Fechas** | Recepción: 13/01/2020 - Revisión: 20/03/2020 - En edición: 16/03/2020 - Publicación final: 01/07/2020

#### **Resumen**

El periodismo de datos ha hallado en los últimos años un terreno abonado para su expansión en las coberturas deportivas, en particular las futbolísticas. El sustrato estadístico de la competición y su carácter cíclico han favorecido el desarrollo de esta nueva modalidad de análisis con datos que se apoya en matemática aplicada para expandir el periodismo deportivo actual. Este artículo profundiza en el estudio de la aplicación, por primera vez en el diario *El País*, de un modelo matemático de predicción de resultados en una gran cobertura como el Mundial de fútbol de Rusia en 2018. Asimismo, evalúa el grado de aceptación y comprensión de esta metodología de estadística avanzada entre los consumidores más habituales de este tipo de información a partir de cuestionarios semiestructurados a estudiantes universitarios de Periodismo Deportivo y periodistas deportivos de medios españoles. Los resultados del estudio ponen de manifiesto que el uso de estadística avanzada para informar de probabilidades en un torneo aún encuentra una difícil aceptación periodística y un desigual grado de entendimiento entre la audiencia. Pese a ello, el periodismo de datos es percibido mayoritariamente como una gran posibilidad para mejorar la diversidad y la calidad de las coberturas deportivas.

#### **Palabras clave**

Estadística; fútbol; infografía; periodismo de datos; periodismo deportivo; visualización de datos

#### **Abstract**

*In recent years, sports and, most notably, football coverage has become a breeding ground for data journalism. The vast number of statistics collected from scheduled matches have helped data analytics techniques based on applied mathematics to be expanded in current sports journalism. This article examines the statistical model in predicting results developed by El País for the first time to enhance its coverage of a mega sporting event: the 2018 FIFA World Cup held in Russia. This research also analyses the level of acceptance and understanding of this advanced statistical method among heavy users of sports content. To this end, a semi-structured questionnaire involving both open-ended and closed questions was conducted to know the views of both Sports Journalism students and sports reporters from twelve Spanish media outlets. The results reveal that the news values of applying advanced statistics to report on probabilities in a sport tournament encounter a reluctant attitude and an uneven level of understanding among professionals and students. Despite this, data journalism is mainly perceived as a huge opportunity to diversify the agenda and improve the quality of sports coverage.*

#### **Keywords**

*Data journalism; data visualisation; football; infographics; sports journalism; statistics*

## 1. Introducción

El periodismo de datos se ha constituido en los últimos años como una de las nuevas modalidades de trabajo en los medios digitales a la hora de representar visualmente las historias y de hallar nuevas formas de contar a partir del análisis de conjuntos ordenados de datos. El *data journalism*, término que fue acuñada por primera vez por Simon Rogers en 2008 en su *Datablog* en *The Guardian* (Knight, 2015), se ha ido implantando desde entonces de manera gradual en las redacciones gracias a la accesibilidad a los datos que propician las nuevas plataformas digitales.

Para muchos autores, el periodismo de datos no es más que la evolución natural del periodismo de precisión (Dader, 1997), mientras que otros prefieren definirlo como una combinación entre la aplicación de métodos científicos de investigación al desempeño periodístico con el uso del ordenador como herramienta imprescindible de trabajo. Así, Crucianelli lo describe como "una suma de métodos conocidos a los que se añaden tres frutos de la innovación tecnológica": los grandes volúmenes de datos a los que se tiene acceso hoy día, la visualización interactiva y la incorporación del programador al equipo periodístico (2013: 107).

Sin embargo, el periodismo de datos, pese a ser un fenómeno reciente, aún se encuentra en una fase expansiva en la que cada vez más redacciones están construyendo equipos para trabajar bases de datos (Rogers, 2014). Además, su grado de implantación por países es aún desigual. Mientras algunas grandes cabeceras de Europa y América como *The Guardian*, *Financial Times*, *The New York Times* o el argentino *La Nación* cuentan con departamentos de Datos y Periodismo Visual, en los medios de España la apuesta editorial en este sentido es todavía menor, debido a "la falta de tradición y de formación dentro y fuera de las redacciones, así como el escaso interés que suscita entre los responsables" de los medios (Ferreras, 2013: 130).

En mayor o menor medida, los medios de comunicación han ido incorporando a sus rutinas de producción digital de noticias las técnicas provenientes del periodismo de datos para enriquecer y diversificar sus coberturas. Mientras Bradshaw (2011) definió esta modalidad de trabajo periodístico como una suma de tareas de recolección, depuración, contextualización, combinación y comunicación de datos, Veglis y Bratsas (2017a) se refirieron a este tipo de periodismo como un proceso que consiste en extraer información útil de bases de datos, en escribir artículos a partir de esa información y en incorporar a esos artículos visualizaciones (a menudo interactivas) que ayuden a los lectores a comprender mejor esas historias.

Asimismo, la evolución del periodismo de datos ha traído consigo la publicación de diferentes tipologías de trabajos informativos, los cuales han sido objeto de clasificación por varios autores. Esas clasificaciones se han realizado en función del nivel de análisis e interpretación que requieren los trabajos de datos para ser producidos, pero sobre todo en función de la estructura y el volumen de datos que incluyen y de la metodología empleada para presentar la información (Veglis y Bratsas, 2017b).

Así, por ejemplo, Kang (2015) categorizó proyectos de datos según sus posibilidades de interacción y visualización, pero sobre todo según sus fines periodísticos: piezas que ofrecen una visión panorámica o muestran la evolución de una realidad a lo largo del tiempo, gráficos/mapas globales que permiten al lector extraer datos de interés local, historias visuales que comparan situaciones y subrayan diferencias, historias con fines más explicativos que suelen construirse a partir de la intersección de diferentes fuentes de datos o proyectos de investigación.

En un proceso de progresiva democratización de las estadísticas, con un mayor número de bases de datos disponibles en acceso abierto, el periodismo se ve obligado a trabajar con grandes volúmenes de información. Esto hace que el periodismo de datos "en ocasiones se convierta en una curación" (Rogers, 2013: 16), ya que el periodista necesita depurar esas bases de datos, analizarlas y proveer de datos clave a cada historia buscando siempre aquellos más pertinentes para cada información y que sean comprensibles para el lector.

En esta fase expansiva, el periodismo de datos se ha ido ocupando de todo tipo de temáticas, y entre ellas la deportiva despunta como una de las de mayor potencial a la hora de contar y visualizar historias con estadísticas (Segel y Heer, 2010; Silver, 2015). La información deportiva es una de las especializaciones donde tiene más peso la estadística y donde ya resulta muy complicado que una crónica de partido esté desprovista de datos que aclaren lo sucedido (Marrero-Rivera, 2010: 131).

De hecho, el periodismo de datos va ganando presencia en los contenidos deportivos de una forma natural, ya que las competiciones generan una importante cantidad de datos, derivados de los resultados (rendimiento, dinámicas, puntos...), los cuales tienen, además, un carácter acumulativo en series históricas que facilitan mediciones avanzadas. Era, por tanto, cuestión de tiempo que el periodismo deportivo aprovechara todo este caudal estadístico para enriquecer sus coberturas. Tal como expone Arias-Robles:

Cuesta imaginar una especialidad que se ajuste mejor al periodismo de datos que la información deportiva, primero, porque cualquier tipo de competición genera una gran cantidad de información cuantificable; segundo, porque un tratamiento metódico de estos datos permite demostrar casos o tendencias (2017: 217).

En su aplicación a la información deportiva, la finalidad del periodismo de datos va a ser la de otorgar un valor añadido y un nuevo significado a los registros estadísticos de equipos y deportistas. De esta forma, como apuntan Rojas y Rivera:

Más que reunir y mostrar datos, el periodismo de datos constituye una estadística avanzada aplicada a la información donde la clave reside en analizar, a partir de una metodología científica, relaciones entre diferentes variables para llegar a unos datos que no suele mostrar la estadística convencional (2016: 3).

Por tanto, se puede delimitar una clara diferencia entre la estadística básica, la mera exposición o enumeración de los datos y registros estadísticos, como se hace, por ejemplo, durante las retransmisiones deportivas en directo; y la estadística avanzada, es decir, la conversión que se otorga a los datos estadísticos por parte del periodismo de datos para ir más allá de esa estadística básica y llegar a nuevos conocimientos que se puedan comunicar a través de gráficos y visualizaciones.

Dentro de ese uso periodístico de la estadística avanzada, se encuentra el aprovechamiento de datos masivos para elaborar predicciones de resultados. En este caso, el desarrollo de modelos probabilísticos dentro de una cobertura deportiva tiene como objetivo extraer lecturas e interpretaciones nuevas a partir de un conjunto de información datificada ya existente. Este conjunto de datos ordenados ya lo sirven a medios de comunicación, ligas y entidades deportivas empresas especializadas en el almacenamiento y tratamiento de estadísticas deportivas como son la estadounidense *Stats LLC* o la británica *Opta Sports*.

Los modelos de predicción de resultados deportivos ya han sido empleados por medios de comunicación como *Financial Times* o la web *Five Thirty Eight*, propiedad de ABC News, que ha hecho de los datos su verdadera seña de identidad (Arias-Robles, 2017: 217). Estas publicaciones han desarrollado este tipo de trabajos para diversificar sus coberturas deportivas, ofreciendo nuevos ángulos y perspectivas antes que buscando contar certezas. A este respecto, Mayer-Schönberger y Cukier se muestran partidarios de que se pueda sacrificar un poco la exactitud si con ello se puede descubrir una tendencia de comportamiento estadístico general. No obstante, agregan que "los datos masivos convierten los cálculos aritméticos en algo más probabilístico que preciso" (2013: 52).

Por tanto, los modelos basados en estadística avanzada, como el confeccionado por *El País* para el Mundial de fútbol 2018, miden probabilidades y hacen proyecciones futuras a partir de datos ya existentes. Se trata, más que hacer meros pronósticos, de determinar qué escenarios son los más factibles dentro de una competición según la tendencia general y cómo informar de esas probabilidades al lector, siguiendo una metodología que caracteriza al periodismo de datos.

## 2. Metodología

Teniendo en cuenta este contexto, este artículo aborda el estudio de caso del modelo probabilístico de resultados desarrollado por *El País* para su cobertura informativa del Mundial de fútbol de Rusia celebrado entre el 14 de junio y 15 de julio 2018. Se trata de la primera vez que este diario generalista español empleó técnicas y herramientas de estadística avanzada para formular predicciones con valor periodístico dentro de su sección de Deportes.

El modelo matemático de predicción de *El País* para el Mundial de Rusia, creado por el analista Kiko Llaneras, se publicó inicialmente el 4 de junio, más de una semana antes de que comenzase el torneo. Tras ese primer pronóstico, el rotativo madrileño iría dando sucesivas predicciones para cada equipo a medida que fue avanzando la competición.

La situación de punto de partida que plantea la revisión teórica efectuada para esta investigación es, en realidad, doble. Por un lado, el hecho de que el periodismo de datos sea aún incipiente dentro de la información deportiva que desarrollan los medios de comunicación en España hace que esta nueva modalidad todavía requiera más tiempo para que los consumidores habituales de contenidos deportivos asimilen, comprendan y acaben aceptando el valor informativo que tiene la proyección de datos en modelos probabilísticos de resultados en el marco de la cobertura de un gran evento deportivo. En este sentido, el modelo probabilístico desarrollado por *El País* se topó con una difícil aceptación y un desigual grado de entendimiento entre la audiencia.

En segundo lugar, cabe señalar que el periodismo deportivo de datos goza de un gran potencial que puede propiciar, a partir del empleo de técnicas de extracción, análisis y visualización de estadísticas, el desarrollo de nuevos ángulos en las coberturas de las competiciones y un mejor aprovechamiento de los recursos gráficos y multimedia de representación visual de la información. Esta potencialidad no solo constituye un incentivo para que la información deportiva avance hacia un periodismo de precisión y, por tanto, de más calidad, sino también para que haga sus contenidos más diversos y atractivos para los consumidores más habituales. Esta es precisamente la apuesta que ha realizado *El País* al aplicar por primera vez esta metodología de trabajo en la cobertura de un gran evento deportivo.

Por todo ello, los objetivos del presente estudio son los siguientes:

1. Examinar el grado de desarrollo actual del periodismo de datos en la especialización periodística deportiva en España.
2. Analizar el valor periodístico de los modelos de predicción de resultados a partir de estadísticas avanzadas en la cobertura de un gran evento deportivo.
3. Estudiar en concreto el modelo probabilístico desarrollado por *El País* en su cobertura del Mundial de fútbol de Rusia 2018, y compararlo con otros modelos similares empleados por otros medios internacionales en coberturas deportivas.
4. Conocer el impacto y el grado de comprensión de las técnicas que se emplean en el periodismo de datos en los principales grupos de consumidores habituales de este tipo de información, como son los estudiantes de Periodismo Deportivo y los profesionales que trabajan en medios y secciones de Deportes en España.
5. Ampliar la reflexión sobre el carácter innovador, las aportaciones, las posibilidades que ofrece la óptica del análisis de los datos para el futuro del periodismo deportivo.

Para tratar de dar respuesta a los objetivos de investigación formulados, se ha llevado a cabo una investigación cualitativa basada, por un lado, en el estudio de caso del modelo probabilístico desarrollado por *El País* en su cobertura informativa del Mundial de fútbol de Rusia. A este respecto, no solo se ha analizado dicha cobertura a través de las distintas predicciones de cada partido que fueron publicándose en la edición online del citado diario a lo largo del torneo, sino también la propia visión del medio. Para ello, se hizo una entrevista en profundidad al analista Kiko Llaneras, autor del modelo matemático objeto de estudio, para conocer la metodología que desarrolló y su propia valoración de la experiencia periodística referida al término del campeonato. Dicha entrevista se realizó el 9 de septiembre de 2018.

En una segunda fase de la investigación, se elaboraron cuestionarios semiestructurados, dirigidos por un lado a estudiantes de Periodismo Deportivo y, por otro, a periodistas que trabajan en redacciones de Deportes de medios de comunicación impresos y digitales españoles, tanto generalistas como especializados. De esta forma, siguiendo la propuesta de Rodríguez-Gómez, Gil-Flores y García-Jiménez (1999: 73) sobre cómo diseñar una investigación cualitativa, el muestreo seguido en la selección de informantes ha seguido un criterio intencionado y no aleatorio, con el fin de confrontar actitudes y percepciones de dos tipos de público sobre la realidad objeto de estudio.

Además, considerando el cuestionario como un instrumento básico para la obtención de datos y para el estudio de las actitudes respecto a un problema (Igartua y Humanes, 2009: 94-95), con los resultados que se pudieran obtener de esas baterías de preguntas se persiguió confrontar testimonios, opiniones y valoraciones entre dos grupos cualificados de público que son consumidores habituales de información deportiva.

El primero de los cuestionarios tenía como objetivos comprobar, por un lado, la consideración general por parte de futuros periodistas hacia las estadísticas y su relación con el periodismo deportivo y, por otro, su valoración del modelo probabilístico desarrollado por *El País* para Rusia 2018. En total, se recogieron 52 respuestas válidas, todas ellas de estudiantes del Grado de Periodismo de la Facultad de Comunicación de la Universidad de Sevilla y matriculados en la asignatura optativa de Periodismo Deportivo, correspondiente al tercer curso de la carrera. El cuestionario se pasó a los estudiantes de forma online a través de la herramienta Formularios de Google el 4 de abril de 2019.

Posteriormente, un cuestionario similar se envió por correo electrónico a 16 periodistas deportivos de 12 medios diferentes. Las respuestas se recolectaron entre el 11 y el 21 de octubre de 2019 gracias también a un trabajo de refuerzo a través del teléfono (WhatsApp) y las redes sociales (mensajes directos en Twitter). Los periodistas participantes han sido: Carles Vila (*Mundo Deportivo*), Javier Sánchez (*El Mundo*), Víctor García (*El Confidencial*), Ignacio Labarga y Alberto Benítez (*Marca*), Antonio Medina (*Estadio Deportivo*), Ignacio Delgado y Álvaro Ramírez (*El Desmarque*), Juan Luis Rodríguez Cudeiro (*El País*), Pablo

Salvago (*Diario de Sevilla*), Mateo González (*ABC*), Jorge Fernández Maldonado (*As*), Eduardo Casado (*20 Minutos*) y Enrique Julián Gómez, Cristina Caparrós y Borja Pardo (*Sphera Sports*).

Ambos cuestionarios, aunque son muy similares y tienen en común la mayor parte de las preguntas, se han diseñado ad hoc teniendo en cuenta las características de ambos tipos de público y se han basado en los interrogantes derivados de la revisión teórica. Dado el carácter exploratorio y cualitativo de la investigación, en ambos casos se han combinado preguntas cerradas, más fácilmente medibles a partir de un reducido número de opciones de respuesta, con una mayoría de preguntas abiertas con el fin de que las personas participantes pudieran explicar su experiencia personal, así como sus percepciones y valoraciones sobre el tema objeto de estudio.

De esta forma, el cuestionario realizado a estudiantes constó de seis preguntas, referidas en su totalidad a la valoración de trabajos publicados por medios en clave de periodismo de datos, salvo una primera cuestión, más introductoria, que tenía como finalidad medir el interés y la aceptación del alumnado por las estadísticas en general. Entretanto, la batería de preguntas enviada a los periodistas comprendió siete cuestiones. De ellas, cuatro fueron similares a las planteadas a los alumnos, mientras que otras dos se reformularon para conocer la visión de profesionales sobre el valor informativo de trabajos basados en modelos de predicción y sobre la implantación del periodismo de datos en las redacciones deportivas en España. Por último, una séptima pregunta trasladó a los periodistas algunos resultados del cuestionario previamente realizado a estudiantes para que interpretasen su significado y se pudieran confrontar así puntos de vista procedentes de ambos tipos de público.

### 3. Resultados

Esta investigación cualitativa se detiene, en primer lugar, en la explicación y el análisis del modelo matemático de predicción de *El País* como ejemplo de un periodismo deportivo de datos innovador. El estudio de este modelo se realiza a partir de la comparación de las actualizaciones que el citado medio fue publicando en su web en el transcurso de la competición con los resultados que finalmente se produjeron, y de la propia explicación que hizo de la herramienta su autor, Kiko Llaneras, en una entrevista.

#### 3.1. Explicación del modelo

El modelo matemático de predicción desarrollado por *El País* para el Mundial de Rusia se dio a conocer a los lectores el 4 de junio de 2018, es decir, diez días antes del inicio de la competición. En ese primer texto, titulado “¿Quién ganará el mundial? Así hacemos la predicción de *El País*” (1), se ofreció un primer pronóstico y se explicó con detalle la metodología empleada para confeccionar dicho modelo.

En esa primera explicación, *El País* indicaba que el modelo tenía tres partes fundamentales: 1) un ranking que medía la fuerza de cada selección; 2) un modelo estadístico para estimar los posibles resultados de cada partido; y 3) un simulador de la competición. Según detalló Llaneras (2018) en un artículo posterior, el ranking de punto se basa en el sistema de puntuación Elo, inspirado en el que se emplea en ajedrez desde la década de los 50 del siglo XX y que desde hace años se emplea también para predecir resultados en fútbol (Hvattum y Arntzen, 2010). El sistema Elo consiste en calcular la habilidad relativa de un jugador y sus probabilidades de victoria a partir de sus resultados contra otros adversarios (Glickman y Jones, 1999). De esta manera, se tuvieron en cuenta los resultados de cada equipo nacional, las ocasiones de gol producidas y concedidas (“Elo esperado”) y los datos de las plantillas de cada selección.

Según explicó Llaneras (2018) en un artículo publicado al término del campeonato, el ranking se basó en los datos de goles esperados de más de 200 partidos de selecciones disputados desde 2017, que les proporcionó la empresa especializada en estadísticas deportivas *Opta Sports*, mientras que para calibrar el valor de cada plantilla se usaron datos de 352 clubs y 800 jugadores.

El analista de datos de *El País* en la entrevista realizada para esta investigación afirma que el modelo se inspiró en trabajos académicos y en publicaciones de otros medios de comunicación basadas en metodologías similares como *Five Thirty Eight* o *Financial Times*. Entre esos trabajos académicos, se refiere a un estudio desarrollado en Alemania para estimar los resultados de los partidos en la Eurocopa de fútbol en 2016 a partir de una fórmula de distribución de goles denominada Poisson (Arroyo, Bravo, Llinás y Muñoz, 2014; Groll, Kneib, Mayr y Schauburger, 2016), en la que se tiene en cuenta el factor campo, es decir, el hecho de jugar como local, hacerlo visitante o medirse en una cancha neutral.

Igualmente, el autor del modelo probabilístico de *El País* aclara que optaron por predecir antes goles que victorias directamente “porque tiene dos ventajas: resuelve la fase de grupos y sirve para predecir prórrogas” (2018). A pesar de estudios que sugieren que los modelos que predicen goles infraestiman los

empates que realmente se producen (Dixon y Coles, 1997: 266-267), que, Llaneras considera que este sesgo es mucho menor en torneos de selecciones que en los de clubes.

El analista de *El País* agrega que para calibrar este modelo se usó una base de datos de 17.000 partidos de selecciones y, para calcular probabilidades a lo largo de las distintas fases del torneo, se realizaron unas 10.000 simulaciones de partidos.

Tras ese primer pronóstico, *El País* fue publicando sucesivas predicciones para cada equipo a medida que la competición fue avanzando (2). Sin embargo, los resultados de la competición no siempre coincidieron con las previsiones y generaron críticas al modelo. Así, por ejemplo, Alemania, una de las selecciones favoritas, cayó eliminada en la primera fase del Mundial pese a partir con casi un 90% de opciones de clasificarse para octavos de final; o España, que partía con más de un 84% de posibilidades de pasar a cuartos, quedó apeada en octavos ante Rusia.

Estas divergencias entre las probabilidades publicadas y los resultados acaecidos durante el Mundial de Rusia plantean un debate sobre el valor periodístico y la idoneidad de publicar informaciones basadas en predicciones y estimaciones a partir de datos estadísticos acumulados en coberturas de grandes eventos deportivos. Por otra parte, conducen a preguntarse también sobre el grado de comprensibilidad para el lector medio, e incluso especializado, de determinados análisis y visualizaciones fundamentados en estadística avanzada cuando no son habituales en los medios y, por tanto, cuando los usuarios no están acostumbrados a interpretarlos y valorarlos adecuadamente.

### **3.2. Cuestionarios**

Con el fin de comprobar hasta qué punto el modelo de *El País* en concreto y el periodismo de datos en la información deportiva en general es aceptado y entendido por los lectores, se realizaron cuestionarios semiestructurados tanto a estudiantes universitarios de Periodismo Deportivo como a periodistas que trabajan en redacciones de Deportes de medios de comunicación impresos y digitales españoles, tanto generalistas como especializados.

#### **3.2.1. Valor y significado de los datos estadísticos**

En los cuestionarios se planteó una primera pregunta común a estudiantes y periodistas con el objeto de medir el grado de aceptación de los usuarios participantes en este estudio por la estadística deportiva en general y de comprobar cuál es la relevancia que le otorgan como contenido, ya sea periodístico o no periodístico. La primera lectura que pudo extraerse es que la estadística deportiva suscita mucho interés: un 78,8% de los universitarios declararon que les gusta o les produce curiosidad comprobar los datos de un evento deportivo mientras lo siguen en directo o tras haberlo visto. Esta cifra en el caso de los profesionales de los medios alcanzó el 81,25%. Igualmente, los usuarios participantes (61,5% y 68,75%, respectivamente) reconocieron que los datos estadísticos resultan ciertamente atractivos para el espectador, usuario o lector medio, lo cual podría explicar que los medios hagan un uso cada vez mayor de ellos.

Ahora bien, esta consideración no se traduce en unos resultados tan positivos cuando se pregunta a ambos grupos de usuarios sobre el valor y el significado de los datos estadísticos. Según sus respuestas, no son una herramienta que permita explicar por sí misma un acontecimiento deportivo o que logre definir al máximo las características de un deportista. De este modo, solo para el 25% de los alumnos y el 12,5% de los periodistas basta con mirar las estadísticas más relevantes de un partido para hacerse una idea muy aproximada de lo ocurrido sin necesidad de haberlo visto. En este mismo sentido, un reducido porcentaje (30,8% y 12,5%, respectivamente) sostiene que es posible conocer cómo es un deportista simplemente con mirar su registro estadístico de pases, remates, victorias, etc. Es decir, desde este prisma, las estadísticas no lo abarcan ni lo explican todo, pero pueden aportar curiosidades.

#### **3.2.2. El papel de los datos en el periodismo deportivo**

El último aspecto de este apartado introductorio del cuestionario tiene que ver con la relación entre estadística y periodismo deportivo: ¿Qué papel tienen los datos en las informaciones periodísticas deportivas? Si bien hay mucha variedad de opiniones sobre el empleo de los datos en este tipo de contenidos, parece existir una premisa básica: el 88,5% de los alumnos de Periodismo y el 87,5% de los profesionales de los medios concuerdan en que la presencia de datos en contenidos periodísticos deportivos resulta de gran utilidad para complementar y enriquecer la información base. Igualmente, una

gran parte de los usuarios participantes en el estudio coinciden en que una crónica o una previa siempre estarán incompletas si no incluyen datos estadísticos relevantes. Así lo indicaron el 67,3% de los futuros periodistas y el 56,25% de quienes actualmente desempeñan la profesión.

### 3.2.3. Aceptación y entendimiento del periodismo de datos

Si bien parece haber consenso en cuanto a la importancia de la sustancia estadística dentro de la información deportiva, existe una mayor disparidad a la hora de entender y asimilar el periodismo de datos como una nueva forma de trabajo en las coberturas. Así se pone de manifiesto en el análisis concreto del modelo probabilístico elaborado *El País* para el Mundial de Fútbol de 2018 (Imagen 1), un trabajo que, por su originalidad y sus peculiaridades, generó tanta admiración como críticas. Los resultados que arroja el cuestionario realizado en la presente investigación así lo corrobora. Por un lado, hasta un 19,2% de los estudiantes consideraron el modelo como algo fallido a tenor de las diferencias entre los resultados pronosticados y los que se dieron finalmente en el torneo. Mientras, para el 80,8% restante el estudio realizado por el periódico les resultó interesante. No obstante, dentro de este último porcentaje un 34,6% lo entendieron como un estudio periodístico serio frente a un 46,2% que sostuvieron que esa manera de usar los datos rebasó los límites estrictamente periodísticos.

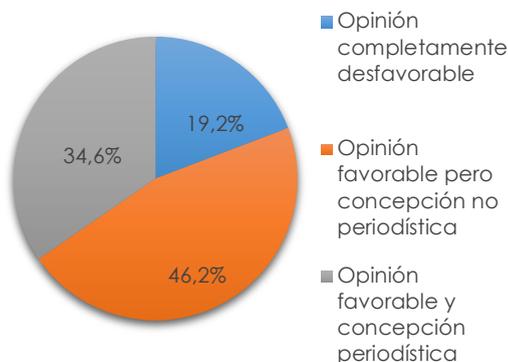
Imagen 1: Predicciones en la previa del Mundial

SELECCION	OCTAVOS	CUARTOS	SEMIFINAL	FINAL	GANAR
 Brasil	90,3%	63,6%	45,5%	29,3%	17,9%
 Alemania	89,1%	60,2%	41,9%	25,5%	15,6%
 España	84,1%	65,9%	42,6%	25,6%	15,3%
 Argentina	85,6%	58,8%	34,8%	19,9%	11,3%
 Portugal	75,0%	53,4%	30,9%	16,0%	7,9%
 Francia	79,5%	46,1%	25,5%	12,9%	6,2%
 Inglaterra	80,8%	50,8%	23,6%	12,1%	5,3%
 Bélgica	79,3%	49,1%	23,1%	11,2%	4,7%
 Colombia	74,6%	40,8%	17,4%	7,6%	3,3%

Fuente: Gráfico publicado por *El País* el 4 de junio de 2018.

El modelo predictivo de *El País* causó sorpresa entre los alumnos, hasta el punto de que solo el 36,5% conocía que los datos pudieran emplearse para intentar predecir eventos futuros. Este manejo innovador de la estadística en el periodismo es mucho más popular entre los profesionales, con un 62,5% que afirman conocerlo. No obstante, el debate vuelve a radicar en que si esa utilización responde a fines periodísticos o no, y este modelo probabilístico es un gran ejemplo porque ni los estudiantes de Periodismo ni los periodistas terminan de ponerse de acuerdo a este respecto.

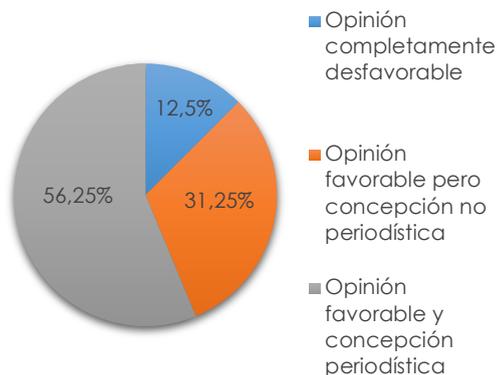
**Gráfico 1: Valoración de alumnos del modelo predictivo de *El País***



Fuente: elaboración propia

Como ilustran los Gráficos 1 y 2, el 56,25% de los profesionales, por solo un 34,6% de los estudiantes, lo entendieron como un estudio periodístico serio, mientras que al 31,25% de profesionales y al 46,2% de los alumnos les resultó atractivo pese a salirse de los márgenes estrictamente periodísticos. Para el 12,5% restante de periodistas y el 19,2% de estudiantes, el modelo es algo inútil y carente de cualquier valor.

**Gráfico 2: Valoración de periodistas del modelo predictivo de *El País***



Fuente: elaboración propia.

### 3.2.4. El valor periodístico de la predicción de resultados

Pero, ¿por qué algunos periodistas defienden que constituye un trabajo periodístico en toda regla y otros no se muestran convencidos de que se atenga por completo a los códigos y procedimientos propios del oficio periodístico, independientemente de que los resultados pronosticados se hayan cumplido en mayor o menor medida? Algunas de las opiniones de profesionales recogidas en la presente investigación insisten en las limitaciones que plantea el uso de la estadística avanzada para informar y explicar acontecimientos desde un punto de vista periodístico.

Es el caso, por ejemplo, de Mateo González, redactor jefe de Deportes en el ABC de Sevilla, quien sostiene que "la estadística avanzada está bien para apuntar tendencias y momentos de forma, pero aventurarse a una predicción cierta del futuro es demasiado arriesgado y quizás inconveniente para un trabajo periodístico", lo cual incluso haría que el modelo "viva limítrofe con el mundo de las apuestas deportivas". Por su parte, Carles Vila, redactor de *Mundo Deportivo*, también mencionó las casas de apuestas al considerar que la metodología que utilizó el modelo de *El País* es parecida a la que utilizan estas para establecer sus cuotas, por lo que dicho modelo "puede servir como paso previo a informar de que en otras ocasiones ha ocurrido un resultado, siendo informativo en ese caso, pero no aporta mucho en el conjunto de la explicación total de la competición".

Asimismo, Víctor García (*El Confidencial*) y Borja Pardo (*Sphera Sports*) coinciden en que este trabajo con datos puede ser un valor añadido excelente de la información, pero que nunca deberían cimentarla. Rescatamos, por último, el juicio de Javier Sánchez, periodista de *El Mundo*, que no le ve más valor que al ranking FIFA u otro similar, y que "los datos pueden ser interesantes para el análisis y levemente para la crónica, pero no para la predicción", que tampoco sería exactamente la labor de un periodista.

Frente a este tipo de críticas, que cabe considerar como entendibles dado lo novedoso de este tipo de trabajos en el periodismo español, Kiko Llaneras (2018), el máximo responsable del proyecto, esgrimió una vez finalizado el Mundial que el modelo realmente sí fue confiable, porque, por ejemplo, los eventos que tenían una probabilidad de entre 0 y 10% de ocurrir solo se materializaron el 3% de las veces, mientras que los sucesos que tenían una probabilidad de darse en una probabilidad superior al 75% ocurrieron en el 86% de las ocasiones.

De esta forma, más allá de esperar una confirmación de los resultados para ver si coinciden con las predicciones del modelo, lo que realmente ocupa el debate es discernir el valor periodístico que contiene este tipo de planteamientos editoriales, porque las cifras sí que comunican (en este caso, igualdad, incertidumbre...), y, además, cuenta con una importante propósito divulgativo, dado que se trata de explicar un evento deportivo de un modo diferente, por medio de datos y de cruces de variables.

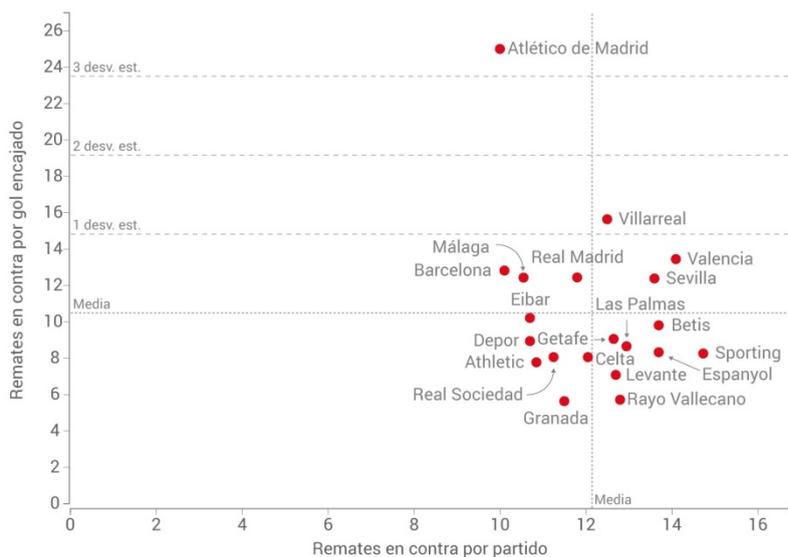
Según Llaneras, un periódico no solo debe obcecarse en buscar hechos noticiosos, sino que también puede explicar un suceso desde otro punto de vista a la par que entretiene a sus lectores. Y lo cierto es que muchos de los profesionales que han participado en la presente investigación parecen abrazar esa idea. Este es el caso de Pablo Salvago, periodista de *Diario de Sevilla*, quien considera que "la estadística avanzada se ha convertido en una herramienta más con la que elaborar informaciones periodísticas válidas e incluso certeras", lo cual hace que el estudio de *El País* sea, a priori, "riguroso y objetivo". También Cristina Caparrós, de *Sphera Sports*, premia su atractivo y su laboriosidad, apuntando que "no se trata de una opinión o de un procedimiento basado en un dato simple, sino de un proceso que recoge diversos factores en su metodología".

### 3.2.5. Un segundo ejemplo de periodismo de datos: gráfico de control

Para calibrar mejor el grado de dificultad planteado por el modelo de *El País*, en el cuestionario se adjuntó un segundo ejemplo de periodismo deportivo de datos, el cual también ha suscitado muy variadas opiniones. Se trata de un gráfico que apareció en una pieza elaborada por *El Español* en enero de 2016 (Imagen 2) y que analizaba la solidez defensiva del Atlético de Madrid frente a otras retaguardias de la Liga española de fútbol (3), cruzando las estadísticas de remates en contra por partido y de remates en contra por gol encajado de los diferentes equipos de la competición.

Era interesante comprobar cómo ha pasado el tiempo para una de las visualizaciones de datos más innovadoras introducidas hasta la fecha por el citado medio digital, que a su vez fue uno de los primeros en España en apostar por el periodismo deportivo de datos (Rojas-Torrijos y Rivera, 2016: 180). Por tanto, se trataba de evaluar si lo que fue disruptivo hace unos años (el gráfico se convierte en la pieza central de la explicación ofreciendo múltiples lecturas a partir de la combinación de variables y el texto queda reducido a la mínima expresión como un mero apoyo de lo visual), ahora se entiende mejor.

Imagen 2: Gráfico de control



Fuente: Opta Sports y elaboración propia

Luis Sevillano / EL ESPAÑOL

Fuente: Gráfico publicado por *El Español* el 20 de enero de 2016.

Sin embargo, los resultados del cuestionario dejan claro que este tipo de trabajos de análisis y visualización de datos deportivos siguen sin comprenderse del todo, incluso entre los grupos de usuarios más habituales. Así, los estudiantes de Periodismo valoraron el gráfico en una dificultad media de 3,9 en una escala de 1 a 5, con hasta el 50% asignándole una dificultad de 4 e incluso el 23,1% otorgándole la máxima dificultad (5). Más llamativa aún resulta la opinión de los profesionales, pues la media de la valoración de la dificultad que hicieron de la citada visualización fue de 3,75 (muy similar a la de los estudiantes), y hasta el 62,5% otorgaron un valor de 4 o superior.

Con el fin de comprobar si realmente el gráfico era tan complejo como parecía, se decidió preguntar a los alumnos cuál era, a su juicio, el mensaje o la idea principal de dicha visualización. En efecto, cruzando las dos estadísticas, se podía llegar a la conclusión de que el Atlético de Madrid es el equipo de toda la Liga al que más cuesta hacerle un gol, lo cual habla muy bien de su sistema defensivo porque se dan simultáneamente dos circunstancias: por un lado, recibe muy pocos remates en contra por partido, y por otro lado, tiene que recibir de media muchos remates antes de que logren hacerle un gol.

Dicho esto, solo el 7,7% del alumnado extrajo esa idea central, mientras que un 17,3% se aproximó mucho a la respuesta ideal, ya que vieron con claridad la solidez defensiva del equipo rojiblanco, pero no llegaron a lo más hondo de la cuestión. Luego, un 34,6% de los estudiantes no pudo sacar nada en claro del gráfico, reconociendo abiertamente que no sabían interpretarlo, mientras que un 7,7% intentó resolverlo pero lo hizo de forma errónea, con algunas respuestas incluso completamente opuestas a la correcta. Finalmente, entre el 32,8% restante se hallan respuestas bastante ambiguas y que no permiten saber si captaron la esencia del gráfico, pues algunos alumnos se limitaron a enunciar el tema genérico (los remates en contra recibidos por los equipos de la Liga), mientras que otros extrajeron lecturas alternativas.

### 3.2.6. Más formación en periodismo de datos para una mejor comprensión

Viendo estos resultados, resultaba interesante preguntar a los profesionales por la respuesta dada por los estudiantes y a qué puede deberse. En concreto se les preguntó a los periodistas si el hecho de que la mayoría de los alumnos de Periodismo no lograra interpretar el gráfico tenía que ver más con deficiencias en su formación o era porque el gráfico era excesivamente complejo. Sorprendentemente, hasta un 56,25% de los periodistas apuntaron más a esta última opción, como subraya, por ejemplo, Jorge Fernández Maldonado, redactor de Deportes de AS. Este sentencia que "es obligación del periodista y del medio que su trabajo final sea entendido por la mayoría de sus receptores", y que "si estos en su mayoría no lo entienden, habrá que buscar otra alternativa". Por su parte, Javier Sánchez (*El Mundo*) está

convencido de que "aunque casi cualquier lector pueda llegar a entenderlo, para conseguirlo debe pararse y examinarlo con detenimiento, y ese esfuerzo va a requerir de "un tiempo que muy posiblemente no va a dedicar a ello". La solución sería, por tanto, intentar simplificar el gráfico o tratar de buscar otro método para exponer los datos, adaptándose así al nivel de sus potenciales lectores, tal como añade Víctor García (*El Confidencial*).

Entretanto, hay profesionales que consideran que el problema hay que situarlo más en el lado de la formación. Enrique Julián Gómez, de *Sphera Sports*, considera que en las facultades "la formación en estadística para los estudiantes, y en general en matemáticas, es escasa", algo que, a su juicio, se antoja preocupante en el caso de los periodistas porque "la estadística es una disciplina cada vez más fundamental para exponer e interpretar la realidad, y por tanto imprescindible para un periodista. Por tanto, "aunque se trate de un gráfico complejo para un lector medio, un periodista debe entender a la perfección y de manera clara este tipo de información estadística".

Al problema de la formación, Carles Vila añade "la falta de experiencia en el manejo de datos por parte del alumnado", lo cual implica que aún no es habitual que los periodistas trabajen desde pronto esa especialidad. Finalmente, hubo usuarios que dieron la misma importancia al problema de la complejidad del gráfico que a posibles carencias en la formación de los futuros periodistas, sin subrayar especialmente ninguno de los dos factores como el causante principal de la falta de entendimiento estudiada.

**Gráfico 3: Valoración de la funcionalidad del gráfico de *El Español* (I)**



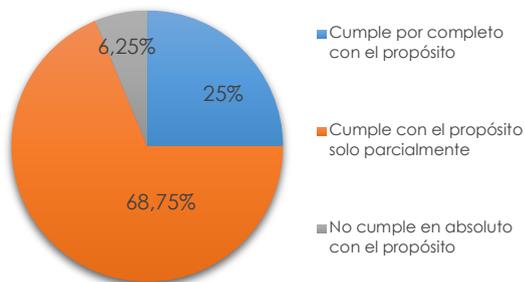
Fuente: elaboración propia.

Junto con el grado de dificultad percibido, otro aspecto que cabía analizar es la funcionalidad del gráfico, es decir, si los dos grupos de usuarios participantes consideran que esa visualización de datos cumple con el propósito de representar con claridad la idea que pretende (la solidez defensiva de los equipos). También en este caso, las opiniones recogidas son bastante divergentes.

Como indica el Gráfico 3, solo para el 17,3% de los estudiantes la visualización cumplía por completo ese propósito, al entender que una defensa que concede menos ocasiones es una defensa más sólida; mientras, para el 57,7% resulta útil porque sirve para explicar una variable del juego relacionada con la labor defensiva, pero no cumple completamente con el propósito inicial del medio digital; y según el 25% el gráfico no sirve para ese propósito, al entender que el hecho de que a un equipo le rematen menos no significa exactamente que defienda mejor.

**Gráfico 4: Valoración de la funcionalidad del gráfico de *El Español* (II)**

### Valoración de los profesionales



Fuente: elaboración propia.

En el caso de los profesionales, mientras tanto, los porcentajes fueron, 25%, 68,75% y 6,25%, respectivamente (ver Gráfico 4), con lo que tenemos una amplia mayoría que considera que dicho trabajo de datos tiene cierto interés pero no termina de convencer al no representar la idea del modo más adecuado.

#### 3.2.7. El futuro del periodismo deportivo de datos

El último apartado de la presente investigación consistió en una pregunta abierta realizada expresamente a los profesionales para saber si, en el momento actual, ven viable que el periodismo de datos se asiente como técnica de trabajo en las redacciones de los medios de comunicación españoles.

Y mientras hay testimonios como el de Antonio Medina (*Estadio Deportivo*), quien cree que "el periodismo de datos ya está muy introducido en las redacciones, al menos en las deportivas", y que incluso "de ahí salen más informaciones que del periodismo de declaraciones", la inmensa mayoría de las opiniones apuntan en la dirección contraria: consideran que el periodismo de datos en España aún no ha dado un paso al frente y que su presencia en los medios es demasiado baja como para que sea considerada como significativa.

Sobre el futuro del periodismo de datos español, hay visiones optimistas. Así, Víctor García señala que los datos son un elemento cada vez más importante en nuestra sociedad y, en consecuencia, están llegando con más fuerza al periodismo. Mientras, para Jorge Fernández Maldonado el asentamiento del periodismo de datos en la información deportiva es perfectamente viable y de hecho muy próximo si tenemos en cuenta que "el manejo de datos a nivel masivo está cada vez más extendido en el mundo del deporte".

Hay otros periodistas que consideran seriamente la posibilidad pero con importantes matices. Por ejemplo, Juan Luis Rodríguez (*El País*) apunta que mientras la precariedad siga siendo la tónica predominante en el entorno periodístico, solo "aquellos medios que sí se puedan permitir destinar recursos humanos a periodismo de datos" podrán implantarlo en sus redacciones. Por su parte, Carles Vila se mostró convencido de que solo será una realidad si los periodistas gozan de la inquietud suficiente para desarrollarlo. Finalmente, Javier Sánchez cree que tendrá éxito pero no en el caso concreto del periodismo deportivo, ya que a su juicio las estadísticas no aportan mucho sobre todo a las piezas de los medios generalistas, que no requieren entrar en aspectos demasiado técnicos.

Por último, varios de los profesionales participantes en el estudio entienden que el periodismo de datos tiene poca cabida en el contexto periodístico actual dentro de nuestras fronteras. La cuestión económica es la causa más recurrente para justificar este argumento, pero también podemos encontrar muchas alusiones a la cuestión del gusto del público receptor. Así, según Eduardo Casado, periodista de *20 Minutos*, antes de poder hablar de crecimiento en el periodismo de datos "es necesario despertar el interés del público, que por ahora no da más valor a estos datos o estadísticas que el meramente anecdótico". Habla, asimismo, de "reeducar" a dicho público, el cual "sigue estando más interesado por los aspectos más 'forofos' de la información deportiva".

#### 4. Discusión

Como hemos visto en los resultados de la investigación, no existe una percepción unívoca sobre el periodismo de datos entre los distintos grupos de público investigados en tanto que consumidores habituales de información deportiva. Las diferencias evidenciadas en cuanto al grado de aceptación y de comprensión del uso de técnicas avanzadas de visualización y de predicción de datos, en el caso concreto de *El País* durante su cobertura del Mundial de fútbol Rusia 2018, plantean una discusión sobre tres cuestiones clave: por un lado, la utilidad y el valor informativo de la probabilidad como concepto asumido por los periodistas; por otro, la apuesta editorial que los medios estén dispuestos a hacer para potenciar el uso del periodismo de datos en las coberturas informativas; y en tercer término, los desafíos que plantea el desarrollo de esta nueva modalidad de periodismo en las redacciones para la formación de futuros profesionales desde las universidades.

En primer término, tal como señalan los resultados, el desarrollo futuro del periodismo de datos en el ámbito de la información deportiva en España afronta como uno de sus más importantes desafíos el de afianzarse como una técnica de trabajo habitual dentro de las redacciones profesionales especializadas. De esta forma, tal como han señalado diferentes estudios, seguiría la estela de medios de otras latitudes, sobre todo anglosajonas (Rojas-Torrijos y Rivera, 2016: 173), pero también en otros países como Alemania (Horky y Pelka, 2017) o países nórdicos (Fink y Anderson, 2015), donde el periodismo deportivo de datos ya opera dentro del ciclo diario de producción de noticias como un campo de innovación emergente que trae consigo nuevas formas de contar historias, de reunir información y de distribuir noticias (Segel y Heer, 2010; Borges-Rey, 2016).

Para su implantación dentro de las redacciones, no obstante, el periodismo de datos ha de incorporarse también al esquema mental de los profesionales, quienes han de explorar las nuevas posibilidades que ofrecen las diferentes tecnologías para hacer ese periodismo posible y también han de asumir nuevas formas de pensar y trabajar como parte de un proceso integral de innovación (Gynnild, 2013). Así, el periodismo de datos podrá trasladarse y "hacerse mucho más visible al consumidor de noticias" (Bradshaw, 2015: 202).

Evidentemente, se trata de un proceso que lleva tiempo y que también pasa por la formación de los futuros profesionales desde las universidades, los mismos que han respondido a las preguntas de esta investigación y que apuntan también a la necesidad de mejorar la capacitación en esta materia dentro de los planes de estudio existentes en el Grado de Periodismo. Como apuntan diferentes autores, los estudios de Periodismo han de estar ligados a los cambios tecnológicos, y precisan de una permanente y crítica actualización en sintonía con el nuevo panorama digital y la realidad de las empresas mediáticas (López, 2012; Saavedra, Grijalba y Pedrero, 2018).

Mientras esa implantación gradual se produce, los profesionales perciben con claridad el importante potencial del periodismo de datos para aportar nuevos enfoques y posibilidades a las coberturas de eventos deportivos, particularmente los futbolísticos, dado el gran volumen de datos en torno a la competición y también el carácter cíclico y repetitivo de los partidos. La naturaleza estadística de la competición futbolística ha facilitado en estos últimos años el desarrollo de tecnologías de automatización a partir de la generación de lenguaje natural dentro de muchas redacciones de Deportes, que producen crónicas de datos robotizadas (Graefe, 2016; Túñez, Toural y Cacheiro, 2018; Rojas-Torrijos y Toural, 2019).

Esta tecnología, a medida que se incorpora a los procesos de creación y distribución de contenidos en los medios deportivos españoles, también implica la asunción por parte de los profesionales de nuevos métodos y rutinas de trabajo basadas en el manejo y el análisis de datos con fines informativos, de forma que luego puedan comunicarlas con eficacia a sus respectivos públicos. Precisamente, el modelo estadístico presentado por *El País* dio un valor periodístico a la probabilidad y trató de innovar a la hora de buscar de qué forma es posible comunicar la incertidumbre, pero se encontró con un grado de aceptación y entendimiento desigual entre los usuarios.

La investigación aquí desarrollada, pese a abordar el fenómeno del periodismo deportivo de datos desde una perspectiva novedosa, como es el grado de percepción de una iniciativa periodística concreta por parte de públicos objetivos, tiene limitaciones. Por un lado, la muestra de público que ha participado en la presente investigación es ampliable a estudiantes de otros centros universitarios de España, de forma que se puedan comparar y evaluar diferentes percepciones de una misma modalidad informativa en la cobertura de competiciones.

Por otra parte, el estudio se fundamenta en un solo caso que, aun siendo innovador y correspondiente a un medio de referencia como *El País* dentro de una cobertura importante como un Mundial de fútbol, es susceptible de ser comparado con modelos similares desarrollados por otros medios de comunicación en sus coberturas deportivas, no solo futbolísticas. A este respecto, cabe plantear futuras líneas de la

investigación sobre el uso de periodismo de datos en otros eventos deportivos relevantes como los Juegos Olímpicos o Mundiales de otras disciplinas.

## 5. Conclusiones

Los resultados permiten confirmar, a través de las respuestas recabadas, las dos cuestiones planteadas al inicio de la presente investigación sobre cuál es la situación del periodismo de datos en las redacciones de Deportes en medios españoles.

Por un lado, la aparición del periodismo de datos como técnica periodística en el ámbito de la información deportiva en España es demasiado reciente y los consumidores habituales de este tipo de contenidos, tal como han manifestado en sus respuestas a los cuestionarios realizados, todavía no parecen haberse acostumbrado, por lo general, a piezas informativas que tienen en el análisis de los datos y las visualizaciones de variables estadísticas sus principales protagonistas. Como consecuencia de esa realidad, dichos contenidos son percibidos a menudo como poco accesibles y excesivamente complejos, tanto que resultan difíciles de comprender.

Quizá lo más significativo es que no solo los usuarios aficionados que se encuentran en fase de formación universitaria están aún en proceso de asimilación de esta nueva modalidad de trabajo periodístico, sino que además los propios profesionales de la comunicación no parecen ponerse de acuerdo sobre lo que significa el periodismo de datos, sus posibilidades de presente y de futuro, así como su valor y sus fines informativos.

En este sentido, el carácter novedoso del modelo de predicción de resultados desarrollado por *El País* durante el Mundial de fútbol en Rusia se encontró con un público no lo suficiente habituado ni preparado para poderlo asimilar e interpretar plenamente ni para darle el valor periodístico que los promotores del proyecto, en particular Kiko Llaneras, pretendió al poner en marcha esta innovación.

No obstante, la evaluación del modelo que efectuó *El País* al compararlo con otros similares que desarrollaron otros medios y empresas internacionales con motivo del Mundial fue muy positiva. Desde el propio periódico argumentan que su modelo estuvo muy bien calibrado, ya que de la predicción inicial los pronósticos de baja probabilidad casi nunca ocurrieron y no prometió más precisión de la que tuvo. El propio Llaneras deja claro que una cosa es calcular probabilidades que tienen un valor informativo y otra muy distinta es hacer meras conjeturas y que quizá este fuera un motivo por el que el modelo no se entendió bien.

En segundo lugar, los resultados del estudio refrendan la idea de que el periodismo de datos es una modalidad de trabajo incipiente que goza de numerosas posibilidades de desarrollo dentro del periodismo deportivo español en los próximos años. Los grupos de público consultados coinciden al destacar las valiosas aportaciones que puede ofrecer el análisis de datos a las coberturas deportivas, no solo propiciando la creación de nuevas historias y de nuevos modos de construir y relatar la información, sino también enriqueciéndolas visualmente con los recursos gráficos que genera.

En suma, una mayor implantación en el futuro del periodismo de datos en las redacciones de Deportes en España implicaría abordar las coberturas informativas de los principales eventos desde una perspectiva original e innovadora. Esta potencialidad no solo puede suponer que la información deportiva avance hacia un periodismo de precisión y de más calidad, sino también propiciar que sus contenidos sean más atractivos y actúen de reclamo para despertar un mayor interés entre los consumidores más habituales.

## 6. Referencias bibliográficas

- [1] Arias-Robles, F. (2017). Nuevas narrativas digitales en el periodismo deportivo. En J. L. Rojas-Torrijos (Coord.), *Periodismo deportivo de manual* (pp. 203-232). Valencia: Tirant lo Blanch
- [2] Arroyo, I.; Bravo, L. C.; Llinás, H. y Muñoz, F. L. (2014). Distribuciones Poisson y Gamma: Una Discreta y Continua Relación. *Prospectiva*, 12(1), 99-107. <http://doi.org/dp72>
- [3] Borges-Rey, E. (2016). Unravelling Data Journalism. *Journalism Practice*, 10(7), 833-843. <http://doi.org/dp73>
- [4] Bradshaw, P. (07/07/2011). The inverted pyramidofdatajournalism [Blog]. *Online Journalism Blog*. <https://bit.ly/2wum7sy>

- [5] Bradshaw, P. (2015). Data Journalism. In L. Zion & D. Criag (Eds.), *Ethics for Digital Journalists: Emerging Best Practices* (pp. 202-218). Nueva York/Abingdon: Routledge
- [6] Crucianelli, S. (2013). ¿Qué es el periodismo de datos? *Cuadernos de Periodistas*, 26, 106-124. <http://bit.ly/32TQCUw>
- [7] Dader, J. L. (1997). *Periodismo de precisión. Vía socioinformática de descubrir noticias*. Madrid: Síntesis
- [8] Dixon, M. J. & Coles, S. G. (1997). Modelling Association Football Scores and Inefficiencies in the Football Betting Market. *Journal of Applied Statistics*, 46(2), 265-280. <http://doi.org/bc265q>
- [9] Ferreras, E. M. (2013). Aproximación teórica al perfil profesional del periodista de datos. *Icono* 14, 11(2), 115-140. <http://doi.org/dp74>
- [10] Fink, K. & Anderson, C. W. (2015). Data Journalism in the United States. *Journalism Studies*, 16(4), 467-481. <http://doi.org/dp75>
- [11] Glickman, M. E. & Jones, A. C. (1999). Rating the chess rating system. *Chance*, 12(2), 21-28. <https://bit.ly/2J1BdZt>
- [12] Graefe, A. (2016). *Guide to automated journalism*. New York: Tow Center.
- [13] Groll, A.; Kneib, T.; Mayr, A. & Schaubberger, G. (2016). *Who's the Favourite? A Bivariate Poisson Model for the UEFA European Football Championship 2016*. Technical Report Number, 195. Munich: Institut für Statistik, Universität München. <http://doi.org/dp76>
- [14] Gynild, A. (2013). Journalism innovation leads to innovation journalism: The impact of computational exploration on changing mindsets. *Journalism*, 15(6), 713-730. <http://doi.org/dp77>
- [15] Horky, T. & Pelka, P. (2017). Data Visualisation in Sports Journalism. *Digital Journalism*, 5(5), 587-606. <http://doi.org/gf3mwn>
- [16] Hvattum, L. M. & Arntzen, H. (2010). Using ELO ratings for match result prediction in association football. *International Journal of Forecasting*, 26(3), 460-470. <http://doi.org/fgs4rg>
- [17] Igartua, J. J. y Humanes, M. L. (2009). *Teoría e investigación en comunicación social*. Madrid: Síntesis
- [18] Kang, M. (15/06/2015). Exploring the 7 Different Types of Data Stories. *Mediashift*. <https://bit.ly/2vErEMR>
- [19] Knight, M. (2015). Data Journalism in the UK: a preliminary analysis of form and content. *Journal of Media Practice*, 16(1), 55-72. <http://doi.org/dp78>
- [20] Llaneras, K. (10/09/2018). Evaluando nuestras predicciones durante el Mundial de Rusia. *Medium*. <http://bit.ly/2QETRuy>
- [21] López, X. (2012). La formación de los periodistas para los entornos digitales actuales. *Revista de Comunicación*, 11, 178-195. <https://bit.ly/33BfctO>
- [22] Marrero-Rivera, O. (2011). *Fundamentos del periodismo deportivo*. San Juan: Terranova
- [23] Mayer-Schönberger, V. y Cukier, K. (2013). *Big data. La revolución de los datos masivos*. Madrid: Turner Publicaciones
- [24] Rodríguez-Gómez, G.; Gil-Flores, J. y García-Jiménez, E. (1999). *Metodología de la investigación cualitativa*. Málaga: Aljibe
- [25] Rogers, S. (2013). *Facts are sacred. The power of data*. Londres: Guardian Books
- [26] Rogers, S. (2014). Data journalism is the new punk. *British Journalism Review*, 25(2), 31-34. <http://doi.org/dp79>
- [27] Rojas-Torrijos, J. L. y Rivera, A. (2016). El Español y El Confidencial, exponentes del periodismo deportivo de datos en los medios nativos digitales españoles. *Revista Doxa Comunicación*, 23, 171-193. <http://doi.org/dp8b>
- [28] Rojas-Torrijos, J. L. y Toural, C. (2019). Periodismo deportivo automatizado. Estudio de caso de AnaFut, el bot desarrollado por El Confidencial para la escritura de crónicas de fútbol. *Revista Doxa Comunicación*, 29, 235-254. <http://doi.org/dp8c>

- [29] Saavedra, M.; Grijalba, N. y Pedrero, L. M. (2018). Hacia una redefinición de las competencias y perfiles profesionales del comunicador audiovisual en el ecosistema digital. *Doxa Comunicación*, 27, 369-385. <http://doi.org/dp8d>
- [30] Segel, E. & Heer, J. (2010). Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6), 1139-1148. <http://doi.org/dnqx5m>
- [31] Silver, N. (17/08/2015). Nate Silver Comes Full Circle With Sports Data Journalism Mecca. *Sporttechie.com*. <https://bit.ly/2WwFtbb>
- [32] Túniz, J. M.; Toural, C. y Cacheiro, S. (2018). Uso de bots y algoritmos para automatizar la redacción de noticias: percepción y actitudes de los periodistas en España. *El Profesional de la Información*, 27(4), 750-758. <http://doi.org/dp8f>
- [33] Veglis, A. & Bratsas, C. (2017a). Reporters in the age of data journalism. *Journal of Applied Journalism & Media Studies*, 6(2), 225-244. <http://doi.org/dp8g>
- [34] Veglis, A. & Bratsas, C. (2017b). Towards a taxonomy of data journalism. *Journal of Media Critiques*, 3(11), 109-121. <http://doi.org/dp8h>

## Notas

1. La primera explicación del modelo de predicción de *El País* se puede leer en la siguiente dirección: <http://bit.ly/2QLQEJU>
2. *El País* publicó a lo largo del Mundial hasta cuatro actualizaciones más del modelo. Fueron las siguientes:
  - a. "Las opciones de cada selección para estar en octavos del Mundial" (2018, 24 de junio): <http://bit.ly/37TfCq>
  - b. "De los equipos en octavos, ¿cuál es el favorito para ganar el Mundial?" (2018, 29 de junio): <http://bit.ly/2tQm495>
  - c. "El "big data" es el pulpo Paul del Mundial de Rusia de 2018 y decía que España será subcampeona" (2018, 1 de julio): <http://bit.ly/36Puo7A>
  - d. "Los favoritos para ganar su cruce de cuartos y el Mundial" (2018, 4 de julio): <http://bit.ly/37Z8LBR>
3. El gráfico de control de *El Español* que ha formado parte del cuestionario se puede consultar en la siguiente dirección: <http://bit.ly/2RcWZwT>

