Audience Participation in TikTok Metadata

La Participación de la Audiencia en los Metadatos de TikTok

Dra. Amparo Huertas-Bailén. Full professor, Universidad Autónoma de Barcelona (España) (amparo.huertas@uab.cat) (https://orcid.org/0000-0002-8851-5417)
Dra. Natalia Quintas-Froufe. Full professor, Universidade da Coruña (España) (n.quintas.froufe@udc.es) (https://orcid.org/0000-0001-7597-6516)
Dra. Ana González-Neira. Full professor, Universidade da Coruña (España) (ana.gneira@udc.es) (https://orcid.org/0000-0002-6369-0323)

ABSTRACT

With the expansion of digital culture, an in-depth reflection on how to research audiences is necessary. If, formerly, the individual was placed in a social category that defined cultural tastes, now technology identifies patterns of behavior from the direct record of their actions. This text explores the type of knowledge that can be obtained on audience participation on TikTok. We propose a methodology that consists of the analysis of usage metadata. The fieldwork focuses on "Ac2ality", an information account with 4.4 million followers in Spain. We analysed all videos shared over six weeks of the first quarter of 2023 (n=173). The purpose was to find (a) the degree of the linear correlation between the metadata for the same video and (b) the existence of correlations between metadata and type of video/content. For each metadatum available with open access (comments, likes, saves, shares and views), four activity levels have been established (low, intermediate, high and very high). The majority trend indicates that the levels obtained by the metadata of the same content are not coincident, that is, a video will have more or less scope according to the observed metadata. The homogeneity of the videos means that only clear correlations between topic and metadata are detected. Topics with less presence can reach high levels of activity.

RESUMEN

Con la expansión de la cultura digital, es necesaria una reflexión profunda sobre cómo investigar las audiencias. Si antes el individuo era ubicado en una categoría social determinante de los gustos culturales, ahora la tecnología identifica pautas de comportamiento a partir del registro directo de sus acciones. Este texto explora el tipo de conocimiento que puede obtenerse sobre la participación de la audiencia en TikTok. Se propone una metodología que consiste en el análisis de los metadatos de uso visibles. El trabajo de campo se ha centrado en "Ac2ality", cuenta de información con 4,4 millones de seguidores en España. Se han analizado todos los vídeos compartidos durante seis semanas del primer trimestre del año 2023 (n=173). La finalidad era conocer (a) el grado de correspondencia lineal entre los metadatos de una misma pieza y (b) la existencia de correlaciones entre metadatos y tipo de vídeo/contenido. Para cada metadato disponible en abierto (comentarios, "me gusta", guardados, compartidos y reproducción), se han establecido cuatro niveles de actividad (bajo, medio, alto y muy alto). La tendencia mayoritaria señala que los niveles obtenidos por los metadatos de un mismo contenido no son coincidentes, es decir, un vídeo tendrá más o menos alcance según el metadato observado. La homogeneidad de los vídeos hace que solo se detecten correlaciones claras entre tema y metadatos. Los temas con menor presencia pueden alcanzar niveles de actividad altos.

KEYWORDS | PALABRAS CLAVE

Audience, Metadata, Participation, TikTok, Social Media, Video.
Audiencia, Metadatos, Participación, TikTok, Redes Sociales, Vídeo.

1. Introduction and State of the Issue

The continuing changes in the means of production and distribution of media contents forces the constant redefinition of the best way to research audiences. If we focus on the area of metering, we find that the value given to traditional studies on representational population samples is becoming hazy. On the one hand, the ever-growing global hyperconnectivity has weakened the usefulness of working with populational universes the statistical representativity of which is determined by territories with physical limits, and on the other, the extreme personalization of digital consumption has favored an increase in minority consumption with registers beyond the reach of major surveys (García-Orosa et al., 2023; Huertas, 2002, 2015; Livingstone, 2019; Wolton, 1999). With the development of digital technology, age-old controversies have been revived: the use of sociodemographic classifications for the comprehension of media consumption is again open to question as digital traces stand out as the best source of information on audience behaviou r (Andersen, 2020; Fisher & Mehozay, 2019; Fulgoni & Lipsman, 2017; Hallinan & Striphas, 2016; Huertas, 2015; Jensen, 2019; Lawlor, 2020; Livingstone, 2019; Reder, 2012; Webster, 2016). Sociodemographic classifications have never been free of criticism, despite the fact that the classification criteria have been constantly adjusted in accordance with socio-economic changes. Let us not forget the argument between those who defend the concept of "social class", defined as the "undisputed theoretical superhero" (Hartley, 2000), and those who believe it is more appropriate to speak of "lifestyles" (Huertas, 2002). But nowadays these sociodemographic variables are totally debatable. Some believe that the uniformity of traditional social groups was merely an illusion (Jones, 2023). With this background, new modes of researching audiences must be conceived and tested.

In the area of digital culture, the idea that a video may attract different people is growing and, therefore, it is enough to find the "patterns of affinity". The algorithmic epistemology rejects social ascription and classifies the individual into predictive consumer categories (Fisher & Mehozay, 2019). The algorithms function on the basis of the register of user interactions, who they follow and whose decisions they calculate (Fisher, 2015; García-Marín & Serrano-Contreras, 2023; Riemer & Peter, 2021). The users feed the algorithms and this is precisely what makes them work (Andersen, 2020; Jensen, 2019). This vision consolidates the importance of an empathic relationship between the public and the creators of the contents they follow, as they must feel that they have shared tastes, interests and desires (Fisher & Mehozay, 2019) and, in addition, the possibility that the audience may become aware of their continued monitoring means that we can refer to digital activity as a political act, in which it is possible to create "algorithmically imagined audiences", or anti-publics, based on the control of the algorithms (Bucher, 2017; Fisher & Mehozay, 2019; Warner, 2012).

Digital technology fosters self-segmentation processes and this generates another different type of audience knowledge, although this may not be considered more exact than that offered by representative samples (Fisher & Mehozay, 2019). Rifkin (2000) predicted that, in the end, we would refer to relationship technology instead of information technology, and the theory of cultural omnivorousness, based on the rejection of social class as the deciding factor in cultural tastes, may also be re-assumed so as to understand this new type of media consumption (Fernández Rodríguez & Heikkilä, 2011).

A key theoretical issue is whether we may understand the functioning of algorithms. Andersen (2020) states that, as occurs with many other cultural artefacts, it is impossible to have direct access to the meaning of algorithms. He explains that, even if we understand their design, we cannot comprehend their workings, as algorithms are mutant, they vary as an effect of the constant information generated by their continued interaction with the users. With full awareness, then, of the difficulty involved in accepting the challenge of understanding the dynamics of algorithms, the objective of this text is to make a primary exploration based on the analysis of the visible use metadata. It begins with the typology prepared by Fisher (2015), who differentiated between public data (contributions by the direct and visible open users), private (confidentially shared information) and metadata (information generated automatically by the algorithms). In addition, it takes into account the metadata concept of use. Developed in the area of archivistics, it refers to the information on the access to and consulting of documents (Méndez Rodríguez, 2003).

Nevertheless, we must underline that the study of algorithms already exists. Broadly speaking, the technical opacity, the lack of transparency of the process and the presence of ethically questionable biases are the aspects that receive most attention (Lawlor, 2020; Noble, 2018; Pasquale, 2015; Thomas et al.,

2018). In contrast, issues such as the partiality and decontextualization of the information they offer, as stated by Livingstone (2019), barely received any attention. There are two different lines of research: one investigates how algorithms are built, trained and adjusted (Dourish, 2016; Flyverbom & Murray, 2018) and the other focuses on the users. Within the framework of the second line, although studies based on the data generated by the algorithms themselves stand out, (Andersen, 2020; Bucher, 2017), empirical works on population samples also appear. An example of the latter is the proposal by Lomborg and Kapsch (2020) who analyse the algorithmic experience (awareness, feelings it generates and assessment). These authors finally reject the idea that the algorithms know more about the users than they do themselves.

The specific question guiding us on accepting the challenge is regarding what type of information is offered to us by the analysis of the metadata, and, more precisely, the metadata of use. Although it is an issue the answer to which demands continuing empirical reflexive work, we propose to initiate this route and, to do so, we have chosen the "Ac2ality" account, available on TikTok, as our case study. Various scientific studies have focused on investigating the social networks, including TikTok (Estrada-García & Gamir-Ríos, 2023; García-Marín & Salvat-Martinrey, 2022; Pérez-Escolar et al., 2023; Vázquez-Herrero et al., 2021), but most focus on production and offer, and do not exhaustively address new perspectives on audience studies.

TikTok, launched in 2017, is the international version of the social network of Chinese origin based exclusively on airing dances and music, which had been inaugurated the previous year (Abidin, 2021). On TikTok, the users share short videos, to which they may add animated backgrounds, sound effects and various visual effects. At the beginning, the time limit was under one minute and, since 2022, they may last 10 minutes. They play in loop mode, but may also be saved. The available interactions are comments, sharing and likes. In 2020, TikTok announced that its recommendation system was highly individualized based on knowledge of the user interactions, of the videos followed and of the data on the connection (Jones, 2023). It is designed to hook the users (Lovink, 2019). TikTok is one of the favourite social networks for young people, with marked penetration in Latin America (Martínez-Estrella et al., 2023). In Spain, one of the countries with the greatest penetration rate for this network (Newman, 2022), the highest affinity level is among 18 to 24-year-olds, among whom the daily average for visits is over 1500; thus, it is close to the levels for YouTube and Instagram, much higher than for Twitter or Twitch, where the daily average for visits is approximately 500 in this age group (Comscore, 2023). The IAB data (2023) also reflects the success of TikTok in Spain: it is the network with the greatest growth over the last three years (views have increased by 109% in one year); it is among the five favourites and, after Instagram, generates the greatest interaction.

The reason for choosing TikTok was this greater precision in the individualized recommendations, apart from its success in Spain. But it was necessary to narrow down the elements for analysis and, to do so, we opted for "Ac2ality", a Spanish company founded in 2020. This choice has taken into account due to its success and the fact that it is involved in the area of news, as this article comes within the framework of a research project on the following of political affairs by young people in Spain. The main channel content of "Ac2ality" on TikTok, addressing a young audience, is made up of videos which summarize a selection of news items published on other news media. Comscore (2023) states that its number of followers in Spain is close to 4.4 million, and Atresmedia has recently signed an agreement with the company to be one of its shareholders (Martínez, 2023). The Reuters Institute has classified it as a "news medium" and places it as the third TikTok news account with the highest viewing average in the world, after "NBC News" and "G1-Globo" (Newman, 2022).

Our aim is to find what type of knowledge can be obtained on TikTok audience participation based on a methodological proposal of the analysis of the metadata use available with free access. We began with two hypotheses: 1) there is linear correlation between all the metadata, that is to say, all the metadata for one video are closely aligned, either at the range of high or low values; 2) the levels of participation reflected by the metadata depend on the content of the material viewed.

2. Material and Methods

The corpus produced for the analysis of the "Ac2ality" metadata is made up of all the videos uploaded to its TikTok account over 42 days in the first quarter of 2023 (n=173). Between April 3

and 15 2023, we extracted the videos posted in January (third and fourth weeks), February (first and second weeks) and March (second and third weeks). Given that "Ac2ality" focuses on current affairs, these dates were chosen to avoid, in as far as possible, key moments such as holiday periods or the organization of media events.

In order to design the analysis file (definition of variables and categories for analysis), a first exploratory observation of a sample of 25 videos was carried out. The material was viewed by the three authors of this document, and helped us to design a consensual analysis file and to define the classification criteria in order to guarantee objectivity in gathering the information for the definitive corpus. The final analysis file was structured as follows:

- Identification data: ascription of an identification number and registration of the date and day of posting, together with its total length expressed in seconds.
- Format. Four categories were defined: summary (video with five news headlines), news (video news item), advertising (advertising video) and internet clip (re-edited video from internet), and others (videos not fitting into the previous categories).
- Presence of the team responsible. The manner in which the individuals identified as part of the "Ac2ality" team was observed, which allowed us to define for categories: at the start, at the start and the end, constant (in over 80% of the video or completely) and absence.
- Use of hashtags. The number used in each video was registered and, in order to have a list of the hashtags and the periodicity of their appearance, each was copied literally.
- The use of mentions of other profiles. The number used in each video was registered and, in addition, each of the mentions was copied literally.
- Register of metadata use: (visible number of) likes, commentaries, views, shares and saves.

Table 1: Characteristics of the Analysis Corpus (n=173).			
Variable	Category	N	% (of total corpus)
Format	News	110	63.6
	Summary	30	1.4
	Clips	21	12.1
	Advertising	5	2.9
	Others	7	4
Presence of team responsible	Start	79	45.7
	Start and end	20	11.5
	Constant	42	24.3
	Absent	32	18.5
Hashtags	Yes	136	78.6
	No	37	21.4
Mentions of profiles on networks	Yes	10	5.7
	No	163	94.3
Day of posting	Weekdays	159	92
	Weekend	14	8
Length	< 20"	23	13.3
	Between 21" and 1'	142	82
	> 1'	8	4.7

In addition, to further address the specific aspects of the videos in news formats, for this portion of the sample two more elements were added (table 2):

- Subject. Based on our viewing of each video and the reading of the accompanying texts, particularly the hashtags used to identify each subject, nine subject categories were defined: international; private life of public figures; science and health; technology; politics (Spain); gender-based violence; economy, sport and culture. The category "others" was unnecessary, as all the videos in the sample could be classified.
- Information on media cited as origin of news. The headers and channels mentioned were registered, which, apart from resulting in a complete list of the referenced media, established for categories: press, television, digital media and radio, and others (for example, "newsletters").

Table 2: Characteristics of News Items in Analysis Corpus (n=110).			
Variable	Category	N (citations)	% (of total news items)
Subject	International	32	29
	Private life of public figures	21	19
	Science and health	12	11
	Technology	12	11
	Politics (Spain)	11	10
	Gender-based violence	7	6.3
	Economy	6	5.5
	Sport	5	4.5
	Culture	4	3,7
Variable	Category	N (citations)	% (of total citations)
Media type given as origin of news*	Press	52	49.5
	Television	31	29.5
	Digital media	19	18
	Radio	1	1
	Others	2	2

Notes: *In the quantification of the citations, duplications of a medium have not been eliminated, as it is of

interest to know the total weight by category; 27 news items (24.5%) do not indicate any media as the origin of the information.

Before setting on the definitive version, the analysis file was tested by two of the researchers who signed this article and then the data gathering was done manually. We must remember that TikTok did not offer access to an academic API at the moment when the fieldwork was done (TikTok, 2022). But above all, it must be emphasized that careful viewing of all the materials was needed to identify all the variables and establish the corresponding categories. The information gathered was systemized on an Excel spreadsheet for its later analysis.

In the original plan, it was intended to carry out the differentiated analysis of the older videos and the most recent ones, given that the time available for content is considered decisive for its consumption. Initially, it was planned to classify the corpus in accordance with the months of posting (January: 73 videos, February: 41 videos and March: 59 videos). However, once this classification had been made, it was seen that the result did not offer subsamples with homogeneous data in accordance with the timeline. Five levels for each metadatum had been established (low, average, intermediate, high and very high) and the most outstanding case, as regards the slight difference in the metadata in accordance with the exposition time appeared in the number of comments. The differences were minimal. For example, 39% of the January videos, 41% of the February ones and 42% in March were at the low level.

This led us to discard the idea of subdividing the corpus in accordance with the time passed since its posting, so work was begun on the corpus as a whole. Nevertheless, this process of classification and analysis by month of posting was very useful in the end to refine the definitive ranges of the levels of each metadatum.

Table 3: Ranges of Audience Levels.			
Activity	D	C	B	A
Comments	9-299	300-999	1000-4999	+ 5000
Likes	5720-49999	50000-149999	150000-399999	+ 400000
Saved	97-999	1000-4999	5000-19999	+ 20000
Shared	11-499	500-1499	1500-5900	+ 6000
Views	45000-299000	300000-699999	700000-1999999	+ 2 million

In the methodological design, we proposed to work with for levels: low (D), average (C), high (B) and very high (A). That is, we chose to avoid the intermediate range in order to more clearly detect positive or negative tendencies. In addition, each range had to be plainly defined so as to guarantee a significant number of videos for each. Finally, to adapt the methodology to the analysed case, we decided to define the initial and final values for each range based on the analysis of the audience participation in "Ac2ality". For a level D, the initial value was determined by the lowest metadatum registered

3. Analysis and Results

The results have been grouped into two blocks based on the proposed hypotheses: 3.1) presence of linear correlation between the metadata obtained by each video and 3.2) presence of correlation between metadata and video/content type.

3.1. Presence of Linear Correlation Between the Metadata Obtained by Each Media

Table 4, which reflects the information for the total corpus, confirms the diversity of percentages appearing at each level. If the levels for the metadata of each video individually tended to behave homogeneously, either in a framework for a low, average, high or very high level, the percentages for each column should be similar. By detecting that both the number of videos and the percentages that reach the high level (B) and very high level (A) in "shared" and "views" are similar, it has been found to what extent these metadata correspond to the same videos, and this information leads us to believe that this fulfils our first hypothesis. However, the results indicates that this linear correlation for these two metadata is quite insignificant: the number of videos that coincide at level B in both metadata is 25 (and not 46, which table 4 may lead us to think), and those that reach level A in those two metadata make up a total of 8 (and not the 13 classified as potentially coincident in table 4).

On observing the percentage data at each level, we see that at the lowest level (D) 41% of the sample regarding the metadata quantifying the commentaries appears; 45.7% for "likes"; 30%, for "saved"; 37.8%, for "shared" and.34,3 %, for "views". That is to say, the difference in percentage points within the D level for all the meta data between the highest and the lowest values is 15.7, with the highest percentage for "likes" and the lowest for "saved". The mean level (C) shows the highest percentage difference, 22.3 points: the highest value appears in "saved" (48.5%) and the lowest in "shared" (26.2%). At the high level (B), the difference is 15.9 percentage points: the highest value is in "shared" (28%) and the lowest in "likes" (12.1%). For the very high level, the difference is 5.1 percentage points. This is a smaller difference, but we must take into account that the figures are much lower in this range. Here, the highest value appears in "views" (8.2%) and the lowest in "comments" (2.9%). On this absence of linear correlation of the metadata regarding participation levels, also noteworthy is that the highest indexes at each level appear in different metadata and the same occurs with the lowest levels. That is, we do not detect the tendency for one metadatum to stand out in either way.

Table 4: Distribution of Corpus in Line with Audience Level (n=173).								
Activity	D		C		B		A	
	N	%	N	%	N	%	N	%
Comments	71	41	62	35.8	35	20.2	5	2.9
Likes	79	45.7	66	38.1	21	12.1	7	4
Saved	52	30	84	48.5	29	16.7	8	4.6
Shared*	62	37.8	43	26.2	46	28	13	8
Views**	58	34.3	51	30.1	46	27.2	14	8.2

Notes: *This information is only available for 164 videos of the corpus (the % of this total). **This information is only available in 169 videos of the corpus (the % is of this total).

When we check how many videos place all the metadata at the same level, we find there are 43 (25% of the sample). The majority, 31 (18%) remain at the low level, the most numerous. This group brings together videos in all formats: summaries (15), news items (10), others (3), advertising (2) and clips (1). It must be emphasized that this is where 50% of the summaries registered in our fieldwork are to be found, and that this is the precise format that "Ac2ality" most uses to promote themselves. The remainder is distributed as follows: 8 (4.6%) place all their metadata at level C; 2 (1.1%) at level B and another two 2 (1.1%) at level A. The latter two are news items (video 7: "Clara Chía sobre la canción de Shakira" and video 39: "El núcleo de la tierra se ha detenido").

Finally, the level of dispersion has been checked and the majority tendency is that it is placed at two contiguous levels. This occurs with 91 videos (52.6%): 44 (25,4%) present metadata in D and C; 36 (20.8 %) in C and B, and 11 (6.3%) in B and A. Only 24 videos appear in three juxtaposed levels: 17 (9.8%) have metadata at levels D, C and B, and 7 (4%), at C, B and A. Only video 21 (internet clip: "Brote de tiña en España") presents the highest possible spread (very high: "shared"; high: "views"; intermediate: "saved" and

"likes"; low: "comments"). The same occurs with video 59 (news item: "El reality show de Anna Sorokin"), which presents metadata at non-contiguous levels, in A ("comments") and in C (the remaining metadata).

3.2. Presence of Correlation Between Metadata and Content

We will now analyse the behaviour of the metadata in accordance with the format and the presence on the videos of individuals identified as members of the "Ac2ality" team. Having analysed the total variables gathered in the analysis file, these two are the only ones that allow analysis in some depth. In general, the remaining variables have been dismissed as they had a homogenizing effect on the corpus (because one of the categories stood out very much over the remainder): "Ac2ality" concentrated its activities on weekdays (92%); regarding the length of the corpus, although the shortest video lasts 4" and the longest 1'50" (both are internet clips), 82% are between 21" and 60"; regarding hashtags, 79.6% of the corpus includes at least one and, although it has been found that in 48% of the cases hashtags are used to identify the subject and a list of the most frequently used hashtags has been produced (#explicamelofacil: in 52.6% of the videos; #aprendiendoconTikTok: 23.7%; #noticias1minuto2023: 22%; #NoticiasTikTok: 18% and #NoticiasFáciles: 8.6%), no criterion has been found for their choice. Therefore, only the category presence/absence has been established with scientific rigor. The exception to this homogenizing logic is found in the mentions of other profiles on "Ac2ality". This resource is only used in 5.7% of the corpus.

On observing the levels reached by the metadata of the three most habitual formats in the corpus, the first point that attracts attention is that news items are present at all levels, tending to appear with higher percentages in the intermediate ranges (C and B); except for the case of "likes", where they appear at low level (D) and intermediate (C). This expanded presence at all levels is logical, given that it represents 63.6% of the corpus. Hereafter are the internet clips, which, except in "comments" (where they only reach levels D, C and B), in the remaining metadata place videos in all ranges. But, additionally, this format places a greater percentage of videos at the very high level (A) when it reaches this (for example, 19% in the case of "views"). Finally, the metadata of the summaries tend to place this content at the lowest levels: low (D) and intermediate (C). Only in two of the five metadata ("views" and "shared") do some of these videos reach the high level (B).

Table 5: Level of Interaction with the Videos (% Horizontally).							
Category (total)	Comments				Like			
	D	C	B	A	D	C	B	A
Consistent with Format								
News item (110)	24.6	45.4	25.4	4.6	36.4	41.8	17.2	4.6
Summary (30)	90	10	0	0	63.4	36.6	0	0
Clips (21)	38	28.5	33.5	0	52.5	28.5	9.5	9.5
Advertising (5)	100	0	0	0	100	0	0	0
Others (7)	57.2	42.8	0	0	57.2	42.8	0	0
Consistent with Presence of Team Responsible								
Start (79)	52	33	14	1	50.7	36.7	10.1	2,5
Start and end (20)	30	25	35	10	30	50	20	0
Constant (42)	26.2	47.7	21.4	4.7	38	40.6	14 .3	7.1
Absent (32)	40.7	34.3	25	0	53.2	31.3	9.3	6.2

Observation of the scope of the metadata based on the manner in which members of "Ac2ality" appear does not permit us to indicate clearly that some stimulate more participation than others. We can merely state that appearing at the start and the end tends to generate a better result of a metadatum related with the use of the video, but, even so, they remain chiefly at the intermediate (C) and low (B) levels. However, this profile can be seen to have somewhat more potential to generate comments: 10% reach the very high level, where only 4.7% of the videos with a constant presence appear and 1% of those in which the presence is concentrated at the start. That said, these latter two options do have some videos at the highest level in all the metadata. From the perspective of audience activity, the most ill-advised option would be absence. Nevertheless, the reaction is also uneven depending on which metadatum is observed. In the interaction with the audience the values dominate at the lowest level ("comments": 40.7%; "like": 53.2%). But, conversely, regarding the use of the video, the metadatum of the views of the material in which the staff of "Ac2ality" does reach the highest level (12.5%).

Category (total)	Saved				Shared				Viewed			
	D	C	B	A	D	C	B	A	D	C	B	A
Table 6: Level of Video use (% Horizontally).												
Consistent with format												
News item (110)	21	51	22.6	5.4	28.1	29	31.9	9	27.2	30	31.9	9
Summary (30)	63.4	36.6	0	0	73.4	13.4	10	0	53	33.5	13.5	0
Clips (21)	23.8	52.5	14.2	9.5	9.5	23.8	23.8	14.2	23.8	23.8	28.5	19
Advertising (5)	40	40	20	0	40	40	20	0	80	20	0	0
Others (7)	42.8	57.2	0	0	71.5	0	28.5	0	57.2	28.5	14.3	0
Consistent with Presence of Team Responsible												
Start (79)	36.7	48.1	11.4	3.8	46.8	27.8	16.5	6.3	39.2	29.1	25.3	5
Start and end (20)	5	65	30	0	15	25	55	5	15	35	45	5
Constant (42)	28.4	40.6	23.9	7.1	31	26.1	31	9.5	35.6	31	21.4	12
Absent (32)	31.3	50	12.5	6.2	28.1	15.6	28.1	9.3	28.1	25	25	12.5

Interesting correlations have been detected between metadata and subject regarding news items. In contrast, the application of the variable of the cited media as the original source did not allow for an in-depth analysis; the diversity of the media mentioned meant that occasional appearances dominated, which prevented the establishment of clear correlations.

The news items that deal with international matters, which are those that have the greatest weight in the body of news (32,29%), tend to place their metadata at the intermediate level (C): 18 appear in "comments" and "likes" respectively, and 21 in "saved". As differential elements, it must be highlighted that 10 do reach high level (B) in "views" and that one reaches very high (A) in "comments".

The metadata of news on the private life of public figures, the second subject matter by volume (21,19%), are present at all levels and, also, in quite a balanced way. In "comments" and "likes", a similar number appears (between 5 and 8) in D, C and B. In "saved", the highest proportion (10) appears at the intermediate level (C); in "shared", 12 appear at the low level (D) y, in "views", 9 at a high level (B). But, in all the metadata, some videos always reach the highest level: 1, (in "saved"); 2, (in "comments", "likes" and "shared"), and 3, in "views".

The metadata for news on science and health, despite making up only 11% of the total, with 12 videos, also tend to be present at all levels, but, in comparison with the news items on the private life of public figures, they are usually placed at higher levels, between C and A. As regards interaction, the highest number of videos are to be found at the intermediate level (5: "comments" and 6: "likes"). In the metadata, the greatest number of videos appears in the high range (5: "saved"; 4: "shared" y 6: "views"). But perhaps the most striking point is that five videos reach the highest level in "shared"; two reach this in three metadata ("likes", "saved" and "views") and one does so in "comments".

The news items on technology, with the same weight on the corpus as those on science and health, also tend to appear at all levels of the metadata. However, although they also reached the highest level (one in "likes"; three in "saved" and two in "shared" and "views"), the tendency is to remain at lower levels (D and C).

The metadata for the 10 news items on politics (Spain), which make up 10% of the news items, tend to be concentrated at the intermediate level (C). Here there are between five and seven videos in the metadata for "comments", "likes", "saved" and "views". Only one video reaches the highest level in a metadatum ("shared").

The remainder of the subjects have little weight in the corpus. In their analysis, we can emphasize that only one video on sports (in "comments") and two on culture (in "views") reach the very high level (A), which are precisely the subjects with the least weight 4.5 and 3.7%, respectively). For the remainder the metadata on news regarding gender-based violence tend to be placed at the intermediate levels (C and B), and in the case of the economy, at the lowest level (D).

4. Discussion and Conclusions

Although we cannot access the meaning of the algorithms directly (Andersen, 2020), we consider our methodology venture to be very suitable. This has allowed us to obtain information of great interest to consider the way in which the audiences interact with contents produced by people who are presented

as peers (Fisher & Mehozay, 2019; Rifkin, 2000). In a first phase of the fieldwork, on observing that the length of time a content remains open does not influence the metadata, we realised that the consumption of "Ac2ality" on TikTok is ephemeral. Later, during the development of the fieldwork, the analysis based on the "ad hoc" definition of ranges to determine the levels of participation registered by each metadatum of use has been very rewarding.

After the refutation of the first hypothesis, it is clear that the metadatum chosen as the parameter to measure success on TikTok is not trivial, given that the metadata of use for a single content do not tend to coincide at the same activity level (low, intermediate, high and very high). Linear correlation has only been detected in 25% of the sample and occurs mainly at low levels, which are those where most videos come together. Therefore, on analysing the reach of a content, the metadatum observed must be taken into account and, in its interpretation, it cannot be taken as given that there is a linear correlation, particularly when the metadatum used indicates high activity. Secondly, based on the work for verification of the second hypothesis, we have detected phenomena presenting very useful knowledge for the design of network contents. The analysis of the "Ac2ality" case has allowed us to observe, on the one hand, that the contents with greater presence do not always have the highest levels of participation, and the same is valid for the contents that the channel uses to promote itself by identifying them as distinctive or unique –here we are referring to summaries of news items– and, on the other hand, that the absence of people on the videos that represent the channel is not an obstacle to obtaining good results. Thus, as an example of this reflection, issues that "Ac2ality" barely handles can reach high levels of activity, as has been verified in news items on science and health, or that, although the absence of the figure of the presenter does seem to reduce interaction, this is no obstacle to achieving high levels in the viewing metadatum.

We believe that this work may be of great assistance to address the study of audiences with the new focus demanded by the current digital socio-technological context. The growth of audience activity online is generating an increase in the available range of metadata of use and, therefore, it would be convenient to consider how to reach a consensus on which to use to determine the positioning reached by a content. We also believe it would be interesting to enquire into the pertinence of differentiating between audiences and followers. This idea arose after finding that it is easier to achieve higher levels of participation in the metadata that measure the use of the video (save, share and view) than in those that register the direct interaction with same (comments and likes). The audience would be those who only use the video without interacting with it, while the followers would be those who, in addition, become involved by having dialogue with the content and marking their acceptance. In this way, it is logical that the higher activity would occur in the metadata of video use, where the two profiles meet.

Nevertheless, carrying out this study has also brought many questions to mind. Due to space constraints, we ask only two of them here. First of all, what influence do the algorithms have when it comes to reinforcing (a) one type of participation over another; (b) one content over another or (c) one format over another? And, secondly, is it possible that the algorithms stimulate participation on contents and strategies whose success is already marked by the traditional media? The emergence of news regarding people's private lives or the presenting of sensationalist texts to attract attention or contents and strategies that already work outside the web and, at least in the case of "Ac2ality", also seem to bring about successful outcomes.

That said, to counter the problems that Livingstone (2019) quite rightly noted, the partiality and decontextualization of much of the available studies, we consider that it is indispensable to complement this type of work with representative population samples. For example, the results and conclusions obtained here are a good starting point for the development of specific surveys on samples of "Ac2ality" users on TikTok.

CONOCIMIENTO", within the framework of the Programa Estatal para impulsar la Investigación Científico-Técnica y su Transferencia, Plan Estatal de Investigación Científica, Técnica y de Innovación 2021-2023.

References

Abidin, C. (2021). Mapping internet celebrity on TikTok: Exploring attention economies and visibility labours. *Cultural Science Journal*, *12*(1), 77-103. https://doi.org/10.5334/csci.140

Andersen, J. (2020). Understanding and Interpreting Algorithms: Toward a Hermeneutics of Algorithms. *Media, Culture & Society*, *42*(7-8), 1479-1494. https://doi.org/10.1177/0163443720919373

Bucher, T. (2017). The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms. *Information, Communication & Society*, *20*(1), 30-44. https://doi.org/10.1080/1369118X.2016.1154086

Comscore. (2023, March 15). *Comscore Insights: nuevos tiempos, nuevos consumos, perspectiva 2023*. https://bit.ly/43wDsuZ

Dourish, P. (2016). Algorithms and Their Others: Algorithmic Culture in Context. *Big Data & Society*, *3*(2), 1–11. https://doi.org/10.1177/2053951716665128

Estrada-García, S., & Gamir-Ríos, J. (2023). Soft News in Original Videos. Adaptation to Tiktok of the Main Spanish Online Media. *Profesional De La Información*, *32*(2), 320222. https://doi.org/10.3145/epi.2023.mar.22

Fernández Rodríguez, C. J., & Heikkilä, R. (2011). El Debate Sobre El Omnivorismo Cultural. Una Aproximación a Nuevas Tendencias en Sociología Del Consumo. *Revista Internacional De Sociología*, *69*(3), 585-606. https://doi.org/10.3989/ris.2010.04.15

Fisher, E. (2015). 'You Media': Audiencing as Marketing in Social Media. *Media, Culture & Society*, *37*(1), 50-67. https://doi.org/10.1177/0163443714549088

Fisher, E., & Mehozay, Y. (2019). How Algorithms See Their Audience: Media Epistemes and the Changing Conception of the Individual. *Media, Culture & Society*, *41*(8), 1176-1191. https://doi.org/10.1177/0163443719831598

Flyverbom, M., & Murray, J. (2018). Datastructuring—organizing and Curating Digital Traces Into Action. *Big Data & Society*, *5*(2), 1–12. https://doi.org/10.1177/2053951718799114

Fulgoni, G. M., & Lipsman, A. (2017). Measuring Television in the Programmatic Age: Why Television Measurement Methods Are Shifting Toward Digital. *Journal of Advertising Research*, *57*(1), 10-14. https://doi.org/10.2501/JAR-2017-009

García-Marín, D., & Salvat-Martinrey, G. (2022). Viralizar la verdad. Factores predictivos del engagement en el contenido verificado en TikTok. *Profesional De La Información*, *31*(2), 310210. https://doi.org/10.3145/epi.2022.mar.10

García-Marín, J., & Serrano-Contreras, I.-J. (2023). (Un) Founded Fear Towards the Algorithm: Youtube Recommendations and Polarisation. *Comunicar*, *31*(74), 61-70. https://doi.org/10.3916/C74-2023-05

García-Orosa, B., Canavilhas, J., & Vázquez-Herrero, J. (2023). Algorithms and Communication: a Systematized Literature Review. *Comunicar*, *31*(74), 9-21. https://doi.org/10.3916/C74-2023-01

Hallinan, B., & Striphas, T. (2016). Recommended for You: the Netflix Prize and the Production of Algorithmic Culture. *New Media & Society*, *18*(1), 117-137. https://doi.org/10.1177/1461444814538646

Hartley, J. (2000). *Los Usos De La Televisión*. Paidós. https://go.revistacomunicar.com/vccDVk

Huertas, A. (2002). *La Audiencia Investigada*. Gedisa. https://go.revistacomunicar.com/2IeYgE

Huertas, A. (2015). *Yo Soy Audiencia. Ciudadanía, Público Y Mercado*. UOC Press. https://go.revistacomunicar.com/dJwgaX

IAB. (2023, May 10). *Estudio De Redes Sociales*. IAB Spain. https://go.revistacomunicar.com/lLOSmV

Jensen, K. B. (2019). The Double Hermeneutics of Audience Research. *Television & New Media*, *20*(2), 142-154. https://doi.org/10.1177/1527476418811103

Jones, C. (2023). How to train your algorithm: the struggle for public control over private audience commodities on Tiktok. *Media, Culture & Society*, *45*(6). https://doi.org/10.1177/01634437231159555

Lawlor, B. (2020). Questioning the legitimacy of data. *Information Services & Use*, *40*(3), 259-272. https://doi.org/10.3233/ISU-200098

Livingstone, S. (2019). Audiences in an Age of Datafication: Critical Questions for Media Research. *Television & New Media*, *20*(2), 170-183. https://doi.org/10.1177/1527476418811118

Lomborg, S., & Kapsch, P. H. (2020). Decoding Algorithms. *Media, Culture & Society*, *42*(5), 745-761. https://doi.org/10.1177/0163443719855301

Lovink, G. (2019). *Tristes Por Diseño: Las Redes Sociales Como Ideología*. Consonni. https://doi.org/https://doi.org/10.7203/rd.v1i7.208

Martínez-Estrella, E.-C., Samacá-Salamanca, E., García-Rivero, A., & Cifuentes-Ambra, C. (2023). Generation Z in Chile, Colombia, México, and Panama: Interests and new digital consumption habits. Their use of Instagram and TikTok. *Profesional De La Información*, *32*(2), 1-15. https://doi.org/10.3145/epi.2023.mar.18

Martínez, J. (2023, May 5). *Atresmedia Entra en El Canal De Tiktok Ac2uality a Cambio De Un Pacto Comercial*. La Información. https://go.revistacomunicar.com/M6CABa

Méndez Rodríguez, E. M. (2003). La descripción de documentos electrónicos a través de metadatos: una visión para la Archivística desde la nueva e-Administración. *Revista d'Arxius*, *2003*, 47-82. https://go.revistacomunicar.com/dOUT60

Newman, N. (2022). *How Publishers Are Learning to Create and Distribute News on TikTok* (1914566041). Reuters Institute for the Study of Journalism. https://bit.ly/3OJ2RO0

Noble, S. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press. https://doi.org/10.2307/j.ctt1pwt9w5

Pasquale, F. (2015). *The Black Box Society: the Secret Algorithms That Control Money and Information*. Harvard University Press. https://doi.org/10.4159/harvard.9780674736061.c8

Pérez-Escolar, M., Alcaide-Pulido, P., & Del Toro, A. (2023). Nuevos referentes informativos de la generación Z: Estudio del rol de los y las influencers en TikTok como divulgadores/as de contenidos. *Prisma Social: Revista De Investigación Social*, (40), 262-288. https://go.revistacomunicar.com/RWmfH4

Reder, M. (2012). *Globalización Y Filosofía*. Herder. https://doi.org/10.2307/j.ctvt9k0dr

Riemer, K., & Peter, S. (2021). Algorithmic Audiencing: Why We Need to Rethink Free Speech on Social Media. *Journal of Information Technology*, *36*(4), 409-426. https://doi.org/10.1177/02683962211013358

Rifkin, J. (2000). *La Era Del Acceso. La Revolución De La Nueva Economía*. Paidós. https://go.revistacomunicar.com/7Ahqlz

Thomas, S. L., Nafus, D., & Sherman, J. (2018). Algorithms as Fetish: Faith and Possibility in Algorithmic Work. *Big Data & Society*, *5*(1), 1-11. https://doi.org/10.1177/2053951717751552

TikTok. (2022, Nov 17). *An update on our platform API for researchers*. https://bit.ly/3IGWBIO

Vázquez-Herrero, J., Negreira-Rey, M.-C., & Rodríguez-Vázquez, A.-I. (2021). Intersections Between Tiktok and Tv: Channels and Programmes Thinking Outside the Box. *Journalism and Media*, *2*(1), 1-13. https://doi.org/10.3390/journalmedia2010001

Warner, M. (2012). *Público, Públicos, Contrapúblicos*. Fondo De Cultura Económica (FCE). https://go.revistacomunicar.com/YdSkEh

Webster, J. G. (2016). Why Study Measures of Exposure? From Exposure to Attention. *Communication Methods and Measures*, *10*(2-3), 179-180. https://doi.org/10.1080/19312458.2016.1150974

Wolton, D. (1999). *Sobre La Comunicación*. Acento Editorial. https://go.revistacomunicar.com/b94hh5