UNDERSTANDING THE IMPACTS AND MOTIVATIONS OF DUPLICATE REVIEWS ON TRIPADVISOR

Nuno António*
Universidade Nova de Lisboa, CiTUR-Algarve
https://orcid.org/0000-0002-4801-2487

Marisol B. Correia**
Universidade do Algarve, CiTUR-Algarve
https://orcid.org/0000-0002-1788-6114

Filipa Perdigão Ribeiro***
Universidade do Algarve, CiTUR-Algarve
http://orcid.org/0000-0003-3267-7166

ABSTRACT

TripAdvisor is a popular review platform, where users post reviews for the same place, including duplicate reviews. This duplication can skew research results and visitors' perceptions. To address this issue, we analyze TripAdvisor reviews in 3 languages from 20 attractions in 2 UNESCO heritage-listed cities. We identify 3 types of motivations for multiple reviews: hedonic, utilitarian, and publishing issues. Our study recommends that online review platforms implement strategies to mitigate this and advises researchers to on how to overcome duplicate reviews in their research.

Keywords: bias; duplicate observations; eWOM; Online reviews; Tourism; UGC.

Comprensión de los impactos y las motivaciones de las reseñas duplicadas en Tripadvisor

RESUMEN

TripAdvisor es una plataforma de reseñas, donde los usuarios publican reseñas para el mismo lugar, incluidas las reseñas duplicadas. Esta duplicación puede sesgar los resultados de la investigación y las percepciones de los visitantes. Para abordar este problema, analizamos las reseñas de TripAdvisor en 3 idiomas de 20 atracciones en 2 ciudades declaradas Patrimonio de la Humanidad por la UNESCO. Identificamos 3 tipos de motivaciones para las revisiones múltiples: cuestiones hedónicas, utilitarias y editoriales. Nuestro estudio recomienda que las plataformas de revisión en línea implementen estrategias para mitigar esto y asesora a los investigadores sobre cómo superar las revisiones duplicadas en su investigación.

Palabras clave: sesgo; observaciones duplicadas; eWOM; Reseñas en línea; Turismo; CGU.

1. INTRODUCTION

User-Generated Content (UGC) and Electronic Word-of-Mouth (eWOM) are essential tools for both consumers and companies today (Velicia-Martin *et al*., 2022). Consumers recur to eWOM to decide which products or services to buy (Pyle *et al*., 2021, Walther *et al*., 2023). Companies' marketers take advantage of UGC and eWOM not only to understand consumers' likes and dislikes but also to assess the progress of their actions (Antonio *et al*., 2018a; Chatterjee *et al*., 2021; Oliveira *et al*., 2022).

In tourism, eWOM is one of the main influencers of travelers' trip planning and purchase intentions (Cantallops & Salvi, 2014; Chen & Law, 2016; Filieri *et al*., 2015). For this reason, research on these topics has been continuously increasing and gaining importance (Chen & Law, 2016; Kwok *et al*., 2017). Online reviews are the primary form of UGC that most affect online reputation and impact consumers' decisions (Cantallops and Salvi, 2014; Kwok *et al*., 2017). According to Kwok *et al*. (2017), in tourism, research has been focusing on four major topics, listed below according to the number of studies dedicated to each, the first being the most popular amongst researchers: (1) evaluating quantitative features (valence or review rating, variance, volume, and verbal features), (2) verbal evaluation features ("deciphering" or interpreting the reviews' content), (3) reputation features (understanding how the online review context, such as consumers demographic information or online community status impacts on reputation features), and (4) social features (interaction between users and managers).

However, there has been an increasing concern about the quality of online reviews, particularly with reliability, manipulation, spamming, and fake reviews published to influence the social reputation of goods and services (Walther *et al*., 2023; Antonio *et al*., 2018b; Banerjee and Chua, 2021; L. Chen *et al*., 2019; Díaz and Rodríguez, 2018;

Filieri *et al*., 2015; Heydari *et al*., 2015; J. Li *et al*., 2014; Lin *et al*., 2014; Pyle *et al*., 2021; Thakur *et al*., 2018). Nevertheless, despite this concern with reviews' reliability, research seems to neglect a crucial aspect that may negatively impact the quality of the research that employs online reviews, which are the motivations for publishing multiple reviews for the same place, and how that may influence eWOM research results and the perception of reviews readers.

To study the volume and the possible reasons for users to post multiple reviews for the same place, we decided to take an empiric, data-driven approach. With this approach, we aim to answer our research questions and identify clear examples that illustrate the type of duplicate/near-duplicate reviews.

Our dataset consists of reviews collected from two UNESCO world heritage cities, Coimbra in Portugal and Salamanca in Spain. We assumed that these two Iberian cities would be comparable as tourist attractions since both have medieval universities and city centers (Antonio *et al*., 2020). Although other platforms, such as Expedia or Booking.com, are considered less susceptible to the influence of fake reviews, these platforms limit users' opinions of the product and services they purchase (Mayzlin *et al*., 2014). For this reason, we gathered reviews from TripAdvisor, one of the largest platforms for travel-related online reviews and by far the most popular source of data used in UGC and eWOM tourism and travel research (Chen and Law, 2016; Kwok *et al*., 2017). Indeed, the Scopus database (search: TITLE-ABS-KEY(tripadvisor)) confirmed TripAdvisor's popularity in April 2023, with 1,568 documents including the term "TripAdvisor" in their title, abstract, or keywords.

2. LITERATURE REVIEW

2.1. Manipulation and reliability of online reviews

Some authors dub reviews that are posted with the intention to manipulate as "fake" reviews (Banerjee and Chua, 2021; Ben Khalifa *et al*., 2020; L. Chen *et al*., 2019; Choi *et al*., 2017; L. Li *et al*., 2020; Walther *et al*., 2023), while others as "spam" (Heydari *et al*., 2015; Lau *et al*., 2011; Lin *et al*., 2014). Both terms are used interchangeably to describe reviews not based on any prior product or service experimentation (Banerjee and Chua, 2021; Choi *et al*., 2017) or "authentic" or "real" reviews that result from a product or service compensation, another type of offer, or service provider encouragement (Petrescu *et al*., 2018; Salehi-Esfahani and Ozturk, 2018). Individuals or groups publish these reviews to gain profit or publicity for the service providers and organizations who pay them or for themselves (Choi *et al*., 2017; Heydari *et al*., 2015; Petrescu *et al*., 2018). Fake or spam reviews harm not only consumers but also products/service providers. On the one hand, due to the discrepancy between previous reviews' ratings and their own experience with the product or service, consumers may feel deceived. On the other hand, product/service providers lose essential information on which areas to improve or emphasize (H. Li *et al*., 2023; Salehi-Esfahani and Ozturk, 2018).

Due to the importance of eWOM to businesses in general and to tourism and travel in particular, the reliability and manipulation of UGC, specifically of online reviews in tour-

ism, has been investigated from multiple dimensions. Whereas several studies have investigated the motivations of service providers to promote the publication of fake reviews (Gössling *et al*., 2018; Mayzlin *et al*., 2014), others have focused on the motivations of consumers/individuals (Choi *et al*., 2017) or the methods to distinguish fake reviews from authentic ones (Antonio *et al*., 2018b; L. Li *et al*., 2020).

Although prior studies in tourism have investigated the motivations behind fake reviews, to our knowledge, no prior studies have empirically demonstrated examples of such reviews to understand the motivation of these fake reviews better and simultaneously assess their volume and impact.

2.2. Duplicate/near-duplicate reviews

In probability theory, statistical independence implies that one event does not affect the probability of another event (Kac, 1959; Kruskal, 1988), meaning that one event should not change the belief in another event. Duplicate observations (additional instances of one same observation) or "near duplicates" (observations that share a high number of features with other observations) are a problem in social science studies because they violate the requirement of statistical independence between observations (Kuriakose and Robbins, 2016). The more non-independent observations, the more significant bias of the coefficient and the increase in significance. Non-independent observations increase statistical power and decrease variance, which results in smaller confidence intervals (Berk and Freeman, 2009; Kuriakose and Robbins, 2016; Sarracino and Mikucka, 2017).

Although statistically independent observations should be an obligatory requirement, this issue, at least to the best of our knowledge, is not addressed in online reviews' research literature. The reason may be due to the fact that researchers assume that statistic independence is guaranteed by online reviews platforms, such as TripAdvisor. TripAdvisor's content guidelines state that "[The user] may write one review on any given accommodation, restaurant, or attraction per visit" (TripAdvisor, n.d.-a). To ensure this guideline, TripAdvisor defines a 90-day interval between the publication of new reviews for the same hotel/attraction and a one-month interval for the same restaurant.

3. METHODS

The reviews employed for this study were extracted using a custom-built web robot or simply bot - an application developed to execute tasks in an automated way (Dunham and Melnick, 2009). A custom-built bot was employed to automate the web scraping process, i.e., to simulate the process of a human reading the reviews' web page and save the data on a file. This bot was custom built for TripAdvisor website. The bot extracted all the reviews published before the COVID-19 pandemic, during the years 2017 and 2018, in English, Portuguese, and Spanish, on each city's ten most popular attractions under analysis, Coimbra and Salamanca. Table 1 summarizes the reviews' frequencies and distribution per language and city.

Table 1
REVIEWS' FREQUENCY AND DISTRIBUTION SUMMARY

		English	Portuguese	Spanish	TOTAL
Coimbra	Frequency	884	1,643	465	2,992
	Distribution (in relation to total)	10.2%	19.0%	5.3%	34.6%
	Distribution (in relation to city)	29.6%	54.9%	15.5%	100.0%
Salamanca	Frequency	1,162	587	3,897	5,646
	Distribution (in relation to total)	13.5%	6.8%	45.1%	65.4%
	Distribution (in relation to city)	20.6%	10.4%	69.0%	100.0%
TOTAL	Frequency	2,046	2,230	4,362	8,638
	Distribution (in relation to total)	23.7%	25.8%	50.5%	100.0%

A dataset was created with all the reviews extracted. This dataset comprises the following fields:

- *City*: name of the city (Coimbra or Salamanca).
- *FullText*: qualitative assessment assigned by the user.
- *GlobalRating*: quantitative assessment that each attraction presents at the moment of extraction. On TripAdvisor, this quantitative rating is on a 1-5 scale, in which 3 is average and 5 is excellent (TripAdvisor, 2017).
- *IsMobile*: indication if the review was posted from a mobile device (1) or not (0).
- *Language*: the language of the review.
- *Location*: user registered location.
- *Name*: name of the user. To post content and comments, users must create a TripAdvisor profile. This profile must be associated with a name/pseudonym.
- *PublishDate*: the date on which the review was published.
- *ReviewRating*: quantitative assessment assigned by the user.
- *SiteDesignation*: attraction designation.

To evaluate if users had published more than one review for the same attraction during the two years-period, the dataset was loaded into R (R Core Team, 2016). Using the "dplyr" R package (Wickham *et al*., 2018), reviews where aggregated by *Name* and *SiteDesignation*. At the same time, several measures were calculated to help identify "duplicated" reviews. The R source code employed in creating these measures can be seen in Figure 1. The measures created were:

amplitude: number of days between the oldest and the most recent review.
avgDaysBetweenReviews: average days between each review publication.
count: the number of reviews published by the same user for the same attraction.
isLessThan90Days: indication if reviews are published within less than 90 days (1: yes, 0: no).

isLocal: indication if the user's registered location on TripAdvisor is in the city of the attraction (1: yes, 0: no).

minDaysSincePreviousReview: minimum number of days between reviews' publication.

Figure 1
AGGREGATION SOURCE CODE

```
# Create column with number of days from previous review
fullDataset<-fullDataset %>%
  arrange(PublishedDate) %>%
  group_by(Name,SiteDesignation) %>%
  mutate(daysSincePreviousReview=PublishedDate-lag(PublishedDate))
fullDataset[c("daysSincePreviousReview")]
  [is.na(fullDataset[c("daysSincePreviousReview")])] <- 0
fullDataset$daysSincePreviousReview<-as.numeric(fullDataset$daysSincePreviousReview)


# Aggregate and create measures
cityToCheck<-"Coimbra"
df<-fullDataset%>%
  filter(City==cityToCheck) %>%
  arrange(PublishedDate) %>%
  group_by(Name,SiteDesignation) %>%
  summarise(count=n(),
            amplitude=max(PublishedDate)-min(PublishedDate),
            avgDaysBetweenReviews=sum(daysSincePreviousReview)/(n()-1),
            minDaysSincePreviousReview=
              min(daysSincePreviousReview[daysSincePreviousReview>0]),
            isLocal=min(ifelse(grepl(toLower(cityToCheck),Location,fixed=TRUE,1,0)))
      %>%
  arrange(desc(count))
df<-df[df$count>1,]
df$minDaysSincePreviousReview[is.infinite(df$minDaysSincePreviousReview)]<-0
df$isLessThan90Days<-ifelse(df$minDaysSincePreviousReview<90,1,0)
```

Using this approach, we searched for duplicates and near-duplicates on the two contours this type of reviews may assume: text duplication and conceptual similarity (Heydari *et al*., 2015).

4. RESULTS AND DISCUSSION

The aggregation measures illustrated in Table 2 show that 60 users had published more than one review for the same attraction for Coimbra in the two-year period. One of these

users had published three reviews for the same attraction. As for Salamanca, 64 users had published more than one review for the same attraction, five of which had published three reviews for the same attraction.

Table 2
SUMMARY OF USERS WITH MULTIPLE REVIEWS

Users with…	Coimbra	Salamanca
Two reviews for the same attraction	59	59
Three reviews for the same attraction	1	5
Two reviews for the same attraction in less than 90 days	46	19
Location registered in the city under analysis	1	5
Multiple reviews for two attractions	0	7

For both cities, we found users who had published reviews for the same attraction with less than the 90-day interval stipulated by TripAdvisor (a total of 46 users for the Coimbra attractions and 19 users for the Salamanca attractions) (Table 2). These numbers show that TripAdvisor's policy of only accepting reviews from the same user for the same attraction, written 90 or more days apart, is not fully enacted.

A close reading of the reviews unveils some reasons why users deliberately and otherwise post multiple reviews. Following the consumers' motivations model loosely, these were organized into three preliminary categories, according to (1) hedonic and (2) utilitarian outcomes (Chen and Schwartz, 2010; H. Li *et al*., 2023), and a third category which we have tentatively designated as (3) publishing problems or forgetfulness, for lack of a more precise designation.

4.1. Hedonic/"In love" with their local attractions

Some users reside nearby the attractions, and because they seem to enjoy their place of residence immensely, they feel compelled to advertise it to others. These users tend to submit multiple reviews regularly, with some months apart. For example, as we can see in table 3, user no.1 published three reviews on Plaza Mayor, Salamanca. All three reviews were written in Spanish, using a mobile device, within a span of 483 days. In the 1st review, the user begins by recognizing his/her great fortune for being a Salamanca native-born. In the 2nd review, the same user begins by stating, "What to say, of the Plaza Mayor of Salamanca, which I have not already said, in my previous opinions", recognizing s/he had already posted reviews of the same attraction. In the third review, s/he repeats the exact introductory phrase: "What to say about the Plaza Mayor of Salamanca, which has not been said…".

Table 3
USER NO.1 - COMMENTS ON PLAZA MAYOR, SALAMANCA

Review Number	English Review (authors' translation)	Spanish Review (original language)
1st review	We, people from Salamanca, have the privilege and great fortune, to have a Plaza Mayor, considered one of the most beautiful in the world, whenever we can we come to pay a visit, have a drink or food at any of its terraces, this Plaza is the hallmark of our beloved Salamanca and the meeting place for local residents and visitors, it is a world famous square and that increasingly, welcomes more tourists, for everyone who does not know it, your visit is a must and you will be captivated by [the square], without a doubt, we, the people of Salamanca, do not tire of seeing it and admiring it. (Review rating: 5; 2017-09-01)	*Los salmantinos, tenemos el privilegio y la enorme suerte, de tener una Plaza Mayor, considerada, de las más bellas del mundo; siempre que podemos, nos acercamos a hacerle una visita y de paso, tomarnos algo en cualquiera de sus terrazas; esta Plaza, es la seña de identidad, de nuestra querida Salamanca y el lugar de encuentro de paisanos y visitantes; es una Plaza que tiene fama mundial y que cada vez, acoge a más turistas; para todo aquel que no la conozca, su visita es obligada y quedará prendado de ella, sin duda; nosotros, los salmantinos, no nos cansamos de verla y admirarla.*
2nd review	What to say, of the Plaza Mayor in Salamanca, which I have not already said, in my previous opinions, that it is The Most Beautiful I've ever seen, that it is a Jewel, that it is unique, that is inimitable, just by walking through it or by having a drink or food in any of its terraces, you inflate your lungs, you feel increasingly proud to belong to Salamanca and to have such a Precious and Appreciated Monument, all the people that visit such a beautiful city and make a stop at the Plaza Mayor, are captivated and amazed by so much Beauty and want to immortalize such an amazing moment, with some photos or selfies, it is how every visitor behaves, to take a little bit of the Plaza Mayor of Salamanca, to their home, there is no other like it (Review rating: 5; 2018-04-16)	*Qué decir, de la Plaza Mayor de Salamanca, que no haya dicho ya, en mis anteriores opiniones; que es, La Más Bonita que he visto, que es una Joya, que es Única, que es Inigualable; solo con pasear por Ella o tomarte algo en cualquiera de sus terrazas, se te ensanchan los pulmones, te sientes cada vez más orgulloso de ser salmantino y de tener tan Preciado y Apreciado Monumento; toda la gente que visita tan Bella ciudad y hace una parada en la Plaza Mayor, se queda prendada y admirada de tanta Belleza y quiere inmortalizar tan grato momento, con unas fotos o selfies; es la manera que tiene todo visitante, de llevarse un trocito de la Plaza Mayor de Salamanca, a su casa; no hay otra igual.*

Review Number	English Review (authors' translation)	Spanish Review (original language)
3rd review	What to say about the Plaza Mayor of Salamanca, which has not been said, that it is UNIQUE, besides being the Agora and the Heartbeat of the city, it is the most visited monument by local residents and tourists, it is a Plaza where uncountable local, national and international events take place, in it concerts, car rallies, fairs, national sports celebrations, such as the Tour of Spain, are held it has been the backstage for several films, parades ... (Review rating: 5; 2018-12-28)	*Qué decir de La Plaza Mayor de Salamanca, que no se haya dicho ya, que es ÚNICA; además de ser el Ágora y el Centro Neurálgico de la ciudad, es el Monumento más visitado por locales y turistas; es una Plaza donde se realizan numerosos eventos de índole local, nacional e internacional; en Ella se celebran conciertos, concentraciones automovilísticas, ferias, celebraciones deportivas de ámbito nacional, como la Vuelta Ciclista a España, ha sido escenario de varias películas, de desfiles...*

We also find local users who did not explicitly register their location on their TripAdvisor profile but hold their local attractions in the highest possible esteem. An example of this type of user is illustrated with user no.2's three comments about Plaza Mayor, Salamanca (table 4). Although the user states s/he is originally from Madrid ("I am from Madrid") in the 2nd review, this user clearly states in the 1st review that s/he has the pleasure of crossing the square every day.

Table 4
USER NO.2 - COMMENTS ON PLAZA MAYOR, SALAMANCA

Review Number	English Review (authors' translation)	Spanish Review (original language)
1st review	I have the pleasure of crossing [Plaza Mayor] every day, it is always full of life, sitting on its terraces is an immense pleasure (Review rating: 5; 2017-04-11)	*Tengo el placer de atravesarla todos los días, está siempre llena de vida, sentarse en sus terrazas es un placer inmenso.*
2nd review	Plaza Mayor de Salamanca very authentic, I am from Madrid and I like it more than Madrid, more intimate and a very cool atmosphere (Review rating: 5; 2017-09-16)	*Plaza mayor de Salamanca muy autentica, yo soy de Madrid y me gusta más que la de Madrid, más íntima y un ambiente muy chulo.*
3rd review	Plaza with a lot of charm and beautifully illuminated in the night, but during the November holidays, they had put up a Book Fair and metal poles all around the square, which totally tarnished the square during a period when there are a huge number of tourists, they should have put up the fair somewhere else (Review rating: 5; 2018-11-04)	*Plaza con mucho encanto y de noche preciosa iluminada, el pero es que durante el puente de noviembre, habían puesto en la. plaza una feria del libro y los puestos metálicos deslucian toda la plaza en unas fechas donde hay muchísimos turistas, deberían ponerla en otra ubicación.*

The publication of multiple reviews about a particular place, whether it be a place of living or a much-loved tourist destination, can be attributed to various psychological theories. This study will focus on four theories that seem to help explain a case in point, namely (1) social comparison theory, (2) self-determination theory, (3) social identity theory, and (4) cognitive dissonance theory. Understanding these theories can provide insights into the motivations behind duplicate reviews, which in turn can help businesses and organizations better understand their customers and improve their products and services. The first theory is the social comparison theory (Festinger, 1954), which posits that people evaluate themselves and their experiences based on those of others and that these comparisons can influence their attitudes and behaviors. In the context of online reviews, users may be motivated to share positive reviews of their place of living to create a positive self-image (e.g., portraying themselves as happy or successful when compared to others). The same could also apply to tourist destinations, as users may want to showcase their travel experiences to friends and family, influencing their attitudes and behaviors.

The self-determination theory (Ryan *et al*., 2000) suggests that people have innate psychological needs for autonomy, competence, and relatedness and that these needs can influence their motivation and behavior. In the context of online reviews, users may be motivated to share positive reviews of their place of living as a way to demonstrate their autonomy (e.g., by choosing where to live), competence (e.g., by finding a great place to live), and relatedness or connect with others (e.g., by sharing recommendations with friends and family).

Social identity theory (Tajfel *et al*., 1979) suggests that people derive part of their sense of self from their group memberships and the social categories they belong to. In the context of online reviews, users may be motivated to share positive reviews of their chosen place of living as a way to affirm their identity as a member of a particular geographic community or social group.

Cognitive dissonance theory (Festinger *et al*., 1959) posits that people experience discomfort when they hold conflicting beliefs or attitudes and will often take action to reduce this discomfort. In the context of TripAdvisor, users who have strong positive or negative feelings about a particular attraction or place may feel compelled to write multiple reviews to reconcile their conflicting opinions. For example, a user who loves to live in Salamanca but dislikes over-tourism may be compelled to write duplicate positive reviews about Plaza Mayor to mitigate the negative effect of over-tourism.

4.2. Utilitarian and altruistic motivations

As we have seen above with user no.2 (3rd review), users, probably motivated by the utilitarian and altruistic objective of helping others, or at least to make them avoid a less satisfactory experience, offer advice on what the local authorities should change in the attractions (user no. 2 suggests that the book fair should be placed somewhere else). This behaviour could be explained by the self-presentation theory proposed by Goffman (1959), which suggests that individuals are motivated to present themselves in a favourable light to others. This motivation may lead them to engage in strategic behaviours, such as selectively sharing information or emphasizing positive qualities, to enhance their

social standing or reputation. In the context of duplicate reviews, users may write duplicate reviews to highlight their experiences with a particular attraction to establish themselves as knowledgeable or experienced travellers, thereby bolstering their reputation.

One example of this type of utilitarian motivation is users who write reviews in a language that appears not to be their native one to showcase their expertise and gain recognition from others. As an example of this, user no. 3 from Coimbra, Portugal, writes about the Biblioteca Joanina, Coimbra, in Spanish (table 5).

Table 5
USER NO.3 - COMMENTS ABOUT COIMBRA

Review Number	English Review (authors' translation)	Spanish Review (original language)
1st review	In Coimbra there are few interesting things to visit but this is essential. It can be done as a single visit, but it is better to buy the pack and visit it together with the university and the science museum. (Review rating: 4; 2017-11-19)	*En Coimbra hay pocas cosas interesantes de visitar, pero esta es imprescindible. Se puede hacer como visita única pero mejor comprar el pack y visitar junto con la universidad y el museo de la ciencia.*
2nd review	A splendid library, which deserves to be visited and will not leave you indifferent. Take the opportunity to visit the University. (Review rating: 4; 2018-08-09)	*Una biblioteca espléndida, que merece ser visitada y no te dejará indiferente. Aprovecha para darte una vuelta por la Universidad.*

Another instance of reviews being written in a language other than the reviewer's native one is the comments posted by user no. 4, from Madrid, Spain, about Plaza Mayor, Salamanca. These two reviews were published within a five-day gap. In the first review, posted from a computer and written in Portuguese, the user recommends visiting Plaza Mayor and rates it a 4. In the second review, posted from a mobile device and written in Spanish, the user interestingly raises the rating to a 5 (table 6).

Table 6
USER NO. 4 - COMMENTS ABOUT PLAZA MAYOR, SALAMANCA

Review Number	English Review (authors' translation)	Spanish/Portuguese Review (original language)
1st review	The square is very beautiful!!! When we went to visit still had Christmas ornaments ... there was a huge gift in the centre of the square!! We had lunch in a café and we enjoyed the movement of the square and its beauty!!!! The visit is super recommend. (Review rating: 4; 2018-07-04)	*A Praça é muito bonita!!! Quando fomos visitar estava ainda com detalhes de Natal... no centro da praça tinha um presente enorme!! Almoçamos num café e ficamos apreciando o movimento da Praça e sua beleza!!!! Super recomendo a visita.*

Review Number	English Review (authors' translation)	Spanish/Portuguese Review (original language)
2nd review	I have traveled through many cities in Spain, without a doubt the Plaza Mayor of Salamanca is the most beautiful. Spectacular! (Review rating: 5; 2018-07-09)	*He recorrido muchas ciudades de España, sin duda la Plaza Mayor de Salamanca es la más bonita. Espectacular!*

4.3. Publishing problems or forgetfulness

As illustrated by user no. 2, many users publish reviews for the same attraction within a time frame of less than the required 90 days. On the one hand, this clearly highlights TripAdvisor's difficulties in enforcing its guidelines. On the other hand, it also suggests that users may have forgotten that they had already published a review about their recent visit to a particular attraction. Therefore, they feel compelled to submit additional, albeit further, very similar, reviews. Another pertinent explanation for these repetitions could be the non-immediate publication of the reviews, a regular topic of discussion on the TripAdvisor Support forum (e.g., https://www.tripadvisor.com/ShowTopic-g1-i12105-k11236967-Reviews_will_not_publish_Long_term_problem-TripAdvisor_Support.html or https://www.tripadvisor.co.nz/ShowTopic-g1-i12105-k7602197-Reviews_not_published-TripAdvisor_Support.html). The delay between submission and online publication, as recognized by TripAdvisor in its help centre (TripAdvisor, n.d.-b), could be related to the fact that TripAdvisor may have flagged a particular review as needing additional assessment to confirm it does not violate TripAdvisor's guidelines. In fact, TripAdvisor states that the review assessment could take from 24 hours to several weeks. Any user unfamiliar with these guidelines may decide to submit a new review if s/he cannot locate his or her previously submitted review. For instance, user no. 5 (table 7) published two reviews in English one day apart, both with a quantitative rating of 5 but with different content.

Table 7
USER NO. 5 - COMMENTS ABOUT PLAZA MAYOR, SALAMANCA

Review Number	English Review (original language)
1st review	A must see when in Salamanca. Medieval looking plaza with a variety of shops and cafes under the columns. (Review rating: 5; 2018-01-22)
2nd review	Well attended plaza by tourists and locals alike and when you are in it you can see why, it is very grand and lovely. lined with restaurants and cafes, people sit outside even in winter (Review rating: 5; 2018-01-23)

User no. 6 presents a different situation (table 8). The user published two reviews for Catedral Vieja, in Salamanca, within a four-day period. In the first review, published from a standard computer, the user rated the cathedral with 5 and commented that it was "beautiful, full of art." Still, unfortunately, s/he was only able to see it from the outside. In the second review, published from a mobile device, the user gave it a rating of 1 and explained in the comments why s/he could not visit the inside. As such, the same user wrote two opposite opinions for the same attraction. In this case, the latter publication could be explained by the cognitive dissonance theory (Festinger *et al*., 1959). The user clearly had conflicting beliefs about the visit. However, it is noteworthy that the first review is an assessment of the Cathedral itself, while the second review focuses on a complaint about a specific staff member who did not allow the user to visit the interior of the Cathedral. This kind of personal attack is something that TripAdvisor advises against in its guidelines, and it is likely one of the primary reasons why submitted reviews are not published (TripAdvisor, n.d.-b).

Table 8
USER NO. 6 - COMMENTS ABOUT CATEDRAL VIEJA, SALAMANCA

Review Number	English Review (authors' translation)	Spanish Review (original language)
1st review	Beautiful, full History and Art, we saw it from the outside, but we liked the surroundings and the doorway (Review rating: 5; 2018-03-23)	*Preciosa, llena Historia y Arte, la vimos por fuera, pero nos gustó mucho el entorno y las portadas.*
2nd review	[…]we haven't been let in for arriving 5 minutes late, even after the boss had told the lady at the teller (a brunette with short hair and bad-tempered) to let us in. How disgusting ... They do not realize that they are running out of tourists (Review rating: 1; 2018-03-27)	*No nos han dejado entrar por llegar 5 minutos tarde, aún diciéndole el jefe a la señorita de la caja (una morena de pelo corto y malhumorada) que nos dejará entrar. Que asco… No se dan cuenta que se están quedando sin turismo.*

A similar situation, but with less dichotomous ratings, is illustrated in the reviews of user no. 7 (table 9). This user published the first review on April 18, 2017, assigning a quantitative rating of 4. However, 20 days later, he published a new review mentioning his prior visit in April and upgraded the rating to 5.

Table 9
USER NO. 7 - COMMENTS ABOUT CONVENTO SAN ESTEBÁN, SALAMANCA

Review Number	English Review (original language)
1st review	There is some much to see in the detail that you need to allow twice as much time and not rush as you will miss the best bits (Review rating: 4; 2017-04-18)

Review Number	English Review (original language)
2nd review	The Dominican friars may have taken a vow of poverty, but their order definitely did not. This is a massive place - huge imposing church, with a facade full of statues unharmed by passing wars and infidels (or well-restored), and a very large two-storey cloister. The cloister's decorations on the ground floor are ornate and lovely. On the second floor, in April 2017 at least, they had a display of the work of the Dominicans in Latin America, particularly Peru. It sounded a bit like making up for the role of the clergy in smoothing the way for the brutality of the Spanish conquest and occupation of the region for several centuries. There is also a museum which we did not visit. (Review rating: 5; 2017-05-08)

One of the most common situations we have identified is when users write two reviews on the same day or just one day apart, giving the same quantitative rating but with different content. However, the second review generally presents similarities to the first one. Due to these multiple publications within a brief period, it could be argued that the users may have forgotten that they had already submitted a review. Another possible explanation could be that users post a similar review because they have not seen the previous one published. User no. 8 (table 10) and user no.9's reviews (table 11) further illustrate this point. For instance, user no. 9 posted the first review from a mobile device and the second from a standard computer, which proves the problem is not linked to a technical limitation or mistake since the user submitted each review from a different device.

Table 10
USER NO. 8 - COMMENTS ABOUT BIBLIOTECA JOANINA, COIMBRA

Review Number	English Review (original language)
1st review	How many library have real gold and treasures as decorations? how many library are like living in the set of Harry Potter? Well, this one does and it's super cool. (Review rating: 5; 2017-06-11)
2nd review	University was set up before America was a country. Site was also a palace. It was cool to see. The library is an absolute must to visit. We went on a Sunday when the town was virtually dead quiet. Got in when it first open. (Review rating: 5; 2017-06-22)

Table 11
USER NO. 9 - COMMENTS ABOUT BIBLIOTECA JOANINA, COIMBRA

Review Number	English Review (authors' translation)	Spanish/Portuguese Review (original language)
1st review	Amazing. It reminds us of the Harry Potter films. You cannot take pictures, it has three large halls with shelves up close to the ceiling. Entrance through the side gate on the left after descending the stairs. (Review rating: 5; 2017-12-03)	*Belíssima. Lembra os filmes do Harry Potter. Não se pode fotografar, possui três grandes salões com prateleiras até próximo ao teto. Entrada pelo portão lateral à esquerda depois de descer as escadas.*

Review Number	English Review (authors' translation)	Spanish/Portuguese Review (original language)
2nd review	It consists of three large halls with shelves from floor to ceiling. Impeccable decor. It was created by Dom João V, hence the name. The ceiling is painted with frescoes. (Review rating: 5; 2017-12-04)	*Consta de três grandes salões com prateleiras do chão ao teto. Decoração impecável. Foi criada por Dom João V, daí o nome. Possui o teto pintado de frescos.*

The reasons why people reassess their reviews differently can be framed in the context of the expectancy disconfirmation theory (Oliver, 1980) or the attribution theory (Ross, 1977). The expectancy disconfirmation theory posits that people form expectations about a product, service, or experience before encountering it and that their subsequent satisfaction or dissatisfaction with that product, service, or experience is influenced by how well it meets or exceeds their expectations. This theory could explain why users revise their online reviews in this context. They may do so because their initial expectations were inaccurate or their experience changed over time, leading them to reassess and revise their evaluation. The attribution theory suggests that people make inferences about the causes of their own behaviour and that of others, and these attributions can influence their subsequent attitudes and behaviours. In the context of online reviews, people may revise their reviews if they attribute their initial evaluation to irrelevant factors or discover new information that changes their attribution about the cause of their experience.

5. CONCLUSION

This study focuses on users' motivations for duplicating reviews and possible ways to overcome this duplication. It reveals that the intentional or unintentional publication of multiple reviews by the same user for the same attraction is not uncommon. This problem was detected in reviews from different cities in two countries and three languages. Results show that some users even take the trouble of writing reviews for the same visit in multiple languages. We demonstrated that the motivations for publishing the so-called "fake" reviews go beyond the need to gain profit or publicity. It seems that the motivations for writing more than one review could be linked to the need to publish "authentic" reviews for hedonic or utilitarian reasons.

We also demonstrated that some of the "fake" reviews might happen due to publishing problems, users' forgetfulness, or the intention to reassess previous evaluations. Overall, the reasons for publishing hedonic, utilitarian, or revising online reviews are likely complex and may depend on the individual, the context, and the specific circumstances of their experience. However, these reasons can be interpreted under several scientific theories.

Even though TripAdvisor clearly and publicly advises users against posting more than one review for the same attraction within less than a 90-day interval, this study demonstrates that TripAdvisor's guidelines are not being fully enforced. As such, this situation can misrepresent the online reputation of the attractions and impact the statistical results of online reviews research. Therefore, we recommend that online reviews

platforms, such as TripAdvisor, take measures to enforce their publication guidelines. Otherwise, they may be regarded as not trustworthy. We also recommend that researchers tease out duplicate/near-duplicate reviews when exploring online reviews. According to the time span between the publication dates of the reviews, researchers should remove duplicates/near-duplicates from the collected sample. Removal of duplicate reviews is particularly relevant for duplicates published by the same user within a very short period. Otherwise, research results will be unreliable, especially when the reviews sample is small.

To ensure the reviews sample does not contain duplicates/near-duplicates, researchers can rely on a simple aggregation of reviews, as described in the methods section. Additionally, researchers should only employ reviews or textual data from sources in which users are identified. In addition to collecting the reviews' data, we recommend that researchers collect the reviews' metadata (name of the user, location, publication date, and other information available), as this information enables the identification of duplicates or near-duplicate reviews.

We consider our work raises awareness regarding online reviews research that relies on samples with duplicate/near-duplicate reviews. Finally, to confirm our initial assumptions on motivations, more research needs to be conducted on users' motivations and reasoning for posting repeated comments on the same attraction, service, or product. Therefore, future research could explore another context (e.g., hotels or retail goods) and other languages, as results may differ by context or language.

This study analyzed the patterns and characteristics of users who submitted multiple reviews for the same attraction on TripAdvisor. Our analysis revealed that a significant number of users submit more than one review for the same attraction and, in most cases, these reviews have the same quantitative rating but different content. This suggests that users may have forgotten that they had already submitted a review or that their previous review had not been published due to a violation of TripAdvisor's guidelines.

The study also found that users may have revised their online reviews due to the expectancy disconfirmation theory or the attribution theory. In particular, users could have revised their reviews if their initial expectations were inaccurate or if they discovered new information that changed their attribution about the cause of their experience.

Overall, the findings suggest that TripAdvisor may face challenges in enforcing its guidelines and preventing users from submitting multiple reviews for the same attraction. To address this issue, TripAdvisor could consider implementing additional measures to educate users about its guidelines and to remind them of their previous submissions. Additionally, TripAdvisor could explore new technologies and approaches to detect and prevent the submission of multiple reviews for the same attraction.

Finally, this study contributes to the literature on online reviews by shedding light on a seldomly explored area of research. While most studies on online reviews focus on their impact on consumer behaviour and decision-making, our study provides insights into the patterns and characteristics of users who submit multiple reviews, in different languages, for the same attraction. By highlighting the challenges and opportunities associated with multiple reviews, this study could inform future research and help practitioners to develop more effective strategies to manage online reviews.

Authorship statement: The authors declare no conflict of interest. All authors contributed equally to all aspects of this research including the conceptualization of the study, methodology development, data collection and analysis, and preparation of the manuscript. Authorship statement: The authors declare no conflict of interest. Conceptualization: Nuno António, Marisol B. Correia, Filipa Perdigão Ribeiro. Consulting and treatment of original sources: Nuno António, Marisol B. Correia, Filipa Perdigão Ribeiro. Research and writing: Nuno António, Marisol B. Correia, Filipa Perdigão Ribeiro.

REFERENCES

ANTONIO, N., ALMEIDA, A. DE, NUNES, L., BATISTA, F. and RIBEIRO, R. (2018a): «Hotel online reviews: Different languages, different opinions», *Information Technology and Tourism*, vol. 18 (1-4), pp. 157-185. https://doi.org/10.1007/s40558-018-0107-x

ANTONIO, N., ALMEIDA, A. DE, NUNES, L., BATISTA, F. and RIBEIRO, R. (2018b): «Hotel online reviews: Creating a multi-source aggregated index», *International Journal of Contemporary Hospitality Management*, vol. 30 (12), pp. 3574-3591. https://doi.org/10.1108/IJCHM-05-2017-0302

ANTONIO, N., CORREIA, M. B. and RIBEIRO, F.P. (2020): «Exploring User-Generated Content for Improving Destination Knowledge: The Case of Two World Heritage Cities», *Sustainability*, vol. 12 (22), 9654. https://doi.org/10.3390/su12229654

BANERJEE, S. and CHUA, A.Y.K. (2021): «Calling out fake online reviews through robust epistemic belief», *Information and Management*, vol. 58 (3), 103445. https://doi.org/10.1016/j.im.2021.103445

BEN KHALIFA, M., ELOUEDI, Z. and LEFÈVRE, E. (2020): «Multiple Criteria Fake Reviews Detection Using Belief Function Theory», In A. ABRAHAM, A. K. CHERUKURI, P. MELIN and N. GANDHI (Eds.), *Intelligent Systems Design and Applications* (pp. 315-324), Springer International Publishing. https://doi.org/10.1007/978-3-030-16657-1_29

BERK, R. A. and FREEMAN, D.A. (2009): «Statistical assumptions as empirical commitments», In D. COLLIER, J.S. SEKHON and P.B. STARK (Eds.), *Statistical Models and Causal Inference: A Dialogue with the Social Sciences*, Cambridge University Press. https://doi.org/10.1017/CBO9780511815874

CANTALLOPS, A.S. and SALVI, F. (2014): «New consumer behavior: A review of research on eWOM and hotels», *International Journal of Hospitality Management*, vol. 36, pp. 41-51. https://doi.org/10.1016/j.ijhm.2013.08.007

CHATTERJEE, S., GOYAL, D., PRAKASH, A. and SHARMA, J. (2021): «Exploring healthcare/health-product ecommerce satisfaction: A text mining and machine learning application», *Journal of Business Research*, no. 131, pp. 815-825. https://doi.org/10.1016/j.jbusres.2020.10.043

CHEN, C.-C. and SCHWARTZ, Z. (2010): «The Impact of hedonic and utilitarian motivations on the hotel customers' risk perception», Emerging Issues and Trends in *Hospitality and Tourism Research*, Paper 5. http://digitalscholarship.unlv.edu/hhrc/2010/june2010/5/

CHEN, L., LI, W., CHEN, H. and GENG, S. (2019): «Detection of Fake Reviews: Analysis of Sellers' Manipulation Behavior», *Sustainability*, vol. 11 (17), 4802. https://doi.org/10.3390/su11174802

CHEN, Y.-F. and LAW, R. (2016): «A review of research on electronic word-of-mouth in hospitality and tourism management», *International Journal of Hospitality and Tourism Administration*, vol. 17 (4), pp. 347-372. https://doi.org/10.1080/15256480.2016.1226150

CHOI, S., MATTILA, A. S., VAN HOOF, H. B. and QUADRI-FELITTI, D. (2017): «The Role of Power and Incentives in Inducing Fake Reviews in the Tourism Industry», *Journal of Travel Research*, vol. 56 (8), pp. 975-987. https://doi.org/10.1177/0047287516677168

DÍAZ, M.R. and RODRÍGUEZ, T.F.E. (2018): «Determining the reliability and validity of online reputation databases for lodging: Booking.com, TripAdvisor, and Holiday Check», *Journal of Vacation Marketing*, vol. 24 (3), pp. 261-274. https://doi.org/10.1177/1356766717706103

DUNHAM, K, and MELNICK, J. (2009): «Malicious bots: An Inside look into the cybercriminal underground of the internet», CRC Press.

FESTINGER, L. (1954): «A theory of social comparison processes», *Human Relations*, vol. 7 (2), pp. 117-140.

FILIERI, R., ALGUEZAUI, S. and MCLEAY, F. (2015): «Why do travelers trust TripAdvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth», *Tourism Management*, vol. 51, pp. 174-185. https://doi.org/10.1016/j.tourman.2015.05.007

GOFFMAN, E. (1959): *The presentation of self in everyday life*, Doubleday.

GÖSSLING, S., HALL, C.M. and ANDERSSON, A.-C. (2018): «The manager's dilemma: A conceptualization of online review manipulation strategies», *Current Issues in Tourism*, vol. 21 (5), pp. 484-503. https://doi.org/10.1080/13683500.2015.1127337

HEYDARI, A., TAVAKOLI, M. ALI, SALIM, N. and HEYDARI, Z. (2015): «Detection of review spam: A survey», *Expert Systems with Applications*, vol. 42 (7), pp. 3634-3642. https://doi.org/10.1016/j.eswa.2014.12.029

KAC, M. (1959): «Statistical independence in probability, analysis and number theory», The Mathematical Association of America.

KRUSKAL, W. (1988): «Miracles and statistics: The casual assumption of independence», *Journal of the American Statistical Association*, vol. 83 (404), pp. 929-940. https://doi.org/10.1080/01621459.1988.10478682

KURIAKOSE, N. and ROBBINS, M. (2016): «Don't get duped: Fraud through duplication in public opinion surveys», *Statistical Journal of the IAOS*, vol. 32 (3), pp. 283-291. https://doi.org/10.3233/SJI-160978

KWOK, L., XIE, K.L. and RICHARDS, T. (2017): «Thematic framework of online review research: A systematic analysis of contemporary literature on seven major hospitality and tourism journals», *International Journal of Contemporary Hospitality Management*, vol. 29 (1), pp. 307-354. https://doi.org/10.1108/IJCHM-11-2015-0664

LAU, R.Y.K., LIAO, S.Y., KWOK, R.C.-W., XU, K., XIA, Y. and LI, Y. (2011): «Text mining and probabilistic language modelling for online review spam detection», *ACM Transactions on Management Information Systems*, vol. 2 (4), pp. 1-30. https://doi.org/10.1145/2070710.2070716

LI, H., MENG, F. and HUDSON, S. (2023): «Are Hotel Guests Altruistic? How Positive Review Disconfirmation Affects Consumers' Online Review Behavior», *Journal of Hospitality and Tourism Research*, 47(3), 528-548. https://doi.org/10.1177/10963480211030313

LI, J., OTT, M., CARDIE, C. and HOVY, E. (2014): «Towards a general rule for identifying deceptive opinion spam», In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics* (Vol. 1: Long Papers), pp.1566-1576, Baltimore, Maryland. Association for Computational Linguistics.

LI, L., LEE, K. Y., LEE, M. and YANG, S.-B. (2020): «Unveiling the cloak of deviance: Linguistic cues for psychological processes in fake online reviews», *International Journal of Hospitality Management*, vol. 87, 102468. https://doi.org/10.1016/j.ijhm.2020.102468

LIN, Y., ZHU, T., WANG, X., ZHANG, J. and ZHOU, A. (2014): «Towards online review spam detection», In *Proceedings of the 23rd International Conference on World Wide Web - WWW '14 Companion*, pp. 341-342. https://doi.org/10.1145/2567948.2577293

MAYZLIN, D., DOVER, Y. and CHEVALIER, J. (2014): «Promotional Reviews: An Empirical Investigation of Online Review Manipulation», *American Economic Review*, vol. 104 (8), pp. 2421-2455. https://doi.org/10.1257/aer.104.8.2421

OLIVEIRA, A.S., RENDA, A.I., CORREIA, M.B. and ANTONIO, N. (2022): «Hotel customer segmentation and sentiment analysis through online reviews: an analysis of selected European markets», *Tourism and Management Studies*, vol. 18 (1), pp. 29-40. https://doi.org/10.18089/tms.2022.180103

OLIVER, R.L. (1980): «A cognitive model of the antecedents and consequences of satisfaction decisions», *Journal of Marketing Research*, vol. 17 (4), pp. 460-469.

PETRESCU, M., O'LEARY, K., GOLDRING, D. and BEN MRAD, S. (2018): «Incentivized reviews: Promising the moon for a few stars», *Journal of Retailing and Consumer Services*, vol. 41, pp. 288-295. https://doi.org/10.1016/j.jretconser.2017.04.005

PYLE, M.A., SMITH, A.N. and CHEVTCHOUK, Y. (2021): «In eWOM we trust: Using naïve theories to understand consumer trust in a complex eWOM marketspace», *Journal of Business Research*, vol. 122, pp. 145-158. https://doi.org/10.1016/j.jbusres.2020.08.063

R CORE TEAM. (2016): «R: A language and environment for statistical computing», R Foundation   for Statistical Computing. https://www.R-project.org/

ROSS, L. (1977): The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process. In L. BERKOWITZ (Ed.), *Advances in Experimental Social Psychology* (pp. 173-220). New York: Academic Press.

RYAN, R. M. and DECI, E. L. (2000): «Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being», *American Psychologist*, vol. 55 (1), pp. 68-78.

SALEHI-ESFAHANI, S. and OZTURK, A. B. (2018): «Negative reviews: Formation, spread, and halt of opportunistic behaviour», *International Journal of Hospitality Management*, vol. 74, pp. 138-146. https://doi.org/10.1016/j.ijhm.2018.06.022

SARRACINO, F. and MIKUCKA, M. (2017): «Bias and efficiency loss in regression estimates due to duplicated observations: A Monte Carlo simulation», *Survey Research Methods*, vol. 11 (1). https://doi.org/10.18148/srm/2017.v11i1.7149

TAJFEL, H. and TURNER, J. C. (1979): «An integrative theory of intergroup conflict», In W. G. Austin and S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Brooks/Cole.

THAKUR, R., HALE, D. and SUMMEY, J. H. (2018): «What Motivates Consumers to Partake in Cyber Shilling?», *Journal of Marketing Theory and Practice*, vol. 26 (1-2), pp. 181-195. https://doi.org/10.1080/10696679.2017.1389236

TRIPADVISOR. (n.d.-a): «How often can I write a review?», TripAdvisor Help Center. Retrieved August 16, 2022, from http://www.tripadvisorsupport.com/hc/en-us/articles/200614897-How-often-can-I-write-a-review-

TRIPADVISOR. (n.d.-b): «Why isn't my review posted yet?», TripAdvisor Help Center. Retrieved August 16, 2022, from http://www.tripadvisorsupport.com/hc/en-us/articles/200614817-Why-isn-t-my-review-posted-yet-

TRIPADVISOR. (2017, September 11): «All about your TripAdvisor bubble rating», TripAdvisor. https://www.tripadvisor.com/TripAdvisorInsights/w810

VELICIA-MARTIN, F., FOLGADO-FERNANDEZ, J.A., PALOS-SANCHEZ, P. and LOPEZ-CATALAN, B. (2022): «MWOM business strategies: Factors affecting recommendations», *Journal of Computer Information Systems*, 2041504. https://doi.org/10.1080/08874417.2022.2041504

WALTHER M., JAKOBI T., WATSON S.J. and STEVENS G. (2023): «A systematic literature review about the consumers' side of fake review detection - Which cues do consumers use to determine the veracity of online user reviews?», *Computers in Human Behavior Reports*, vol. 10. https://doi.org/10.1016/j.chbr.2023.100278

WICKHAM, H., FRANÇOIS, R., HENRY, L. and MÜLLER, K. (2018): «dplyr: A grammar of data manipulation (R package version 0.7.8)». https://CRAN.R-project.org/package=dplyr