BEHAVIOUR & INFORMATION TECHNOLOGY
2026, VOL. 45, NO. 4, 694–710
https://doi.org/10.1080/0144929X.2025.2528764

A grounded theory of how consumers determine the veracity of online user reviews

Michelle Walther [a,b], Steven Watson[b], Alexander Boden[a] and Marielle Stel[b]

[a]Department for Economic Sciences, Institute for Consumer Informatics, Bonn-Rhein-Sieg University of Applied Sciences, Sankt Augustin, Germany; [b]Faculty for Psychology of Conflict, Risk, and Safety, Department for Behavioural, Management and Social Sciences (BMS), University of Twente, Enschede, The Netherlands

ABSTRACT
Consumers use online reviews to decide which products to purchase. Cybercriminals produce fake reviews to influence unknowing consumers into buying products of lower quality, which can lead to financial, emotional and physical damage. However, there is still limited understanding of how consumers assess the veracity of online reviews, or incorporate online reviews into purchasing decisions, especially outside of laboratory settings. Therefore, this study uses a grounded theory approach to explore how consumers determine the veracity and trustworthiness of online user reviews. Twenty-five interviews with consumers were held to identify veracity cues, thought processes and other markers of online shopping behaviour. The results show that consumers use online reviews differently depending on context (e.g. product value, consumer knowledge). Our findings support the development of a theory suggesting that consumers evaluate reviews through a two-step process. First, consumers scan the review for relevance and then subsequently evaluate trustworthiness, credibility, and veracity. The different deception cues that are used by consumers are also identified and classified. These findings offer new insights of how consumers identify fake reviews online.

ARTICLE HISTORY
Received 5 March 2025
Accepted 26 June 2025

KEYWORDS
Fake review detection; user-centred research; grounded theory; human fake review detection

1. Introduction

In recent years, worldwide digitisation has fuelled the growth of online shopping (Brüggemann and Olbrich 2023; Hou and Zhu 2023). This trend is evident in Germany, where e-commerce revenue increased from around 20 billion in 2010 and around 84 billion in 2022 (Statista Research Department 2023). This increase in digitisation and online shopping is mirrored by a rise in online criminality, including online shopping fraud (Chevalier 2021; Spithoven 2020; Trivedi 2021). Cybercriminals use multiple strategies to obtain personal data or extort money from potential victims (Spithoven 2020). Different strategies relevant to online shopping include fake shopping websites and fake reviews to trick costumers into buying products they normally would not (Beltzung et al. 2020; Sánchez-Paniagua et al. 2021).

Despite a vast amount of research addressing algorithmic and AI facilitated fake review detection (e.g. Hussain et al. 2019; Jindal and Liu 2008; Ott et al. 2011; Plotkina, Munzel, and Pallud 2020), research shows that automatic fake review detection nonetheless fails to identify many fake reviews. Additionally, fake reviewers are constantly developing techniques to make fake reviews less detectable, meaning they are often one step ahead of current automated protections (Paul and Nikolaev 2021). Therefore, fake reviews continue to pose a significant threat to consumers, highlighting the need to support their ability to detect and navigate fake reviews in real-world settings (Paul and Nikolaev 2021; Walther et al. 2023).

It is critical we understand how consumers assess the veracity of reviews within the broader context of online shopping, as behaviours examined in isolation may differ significantly from those observed in a real-world context (Hofmann and Grigoryan 2023; Winkler and Murphy 1973). To address this, we conducted 25 semi-structured interviews with consumers to answer the research question: '*How do consumers determine the veracity and trustworthiness of online user reviews within the context of genuine purchases online?*'.

1.1. Theoretical framework

Fake reviews are reviews created to mislead customers in their purchase decision (Zhang et al. 2016). These

CONTACT Michelle Walther michelle.walther@h-brs.de Faculty for Psychology of Conflict, Risk, and Safety, Department for Behavioural, Management and Social Sciences (BMS), University of Twente, Drienerlolaan 5, 7522, Enschede, The Netherlands

reviews are often written by people with little or no experience of the products or services they are reviewing (Zhang et al. 2016). Fake reviews can generally be categorised into three distinct types: untruthful opinions, brand reviews, and advertisements (Ansari, Gupta, and Dewangan 2018; Jindal and Liu 2008; Ren and Ji 2019). In this study, we focus on untruthful opinions, as these are the most difficult to recognise as fraudulent (Cohen 2020; Jindal and Liu 2008; Kronrod, Gordeliy, and Lee 2023).

Prior research shows that people are significantly less accurate than current algorithmic methods at detecting fake reviews (Kim et al. 2021; Kronrod, Lee, and Gordeliy 2017; Ott et al. 2011). Plotkina, Munzel, and Pallud (2020) found that while their algorithmic detection tool detected fake reviews with 81% accuracy, human judges only reached 57% accuracy. Wang et al. (2016) achieved algorithmic accuracy of 91.2%. Although automated methods outperform humans in detecting fake reviews, they have limitations. Even with error rates of only 9% many fake reviews still evade detection (Mohawesh et al. 2021). This limitation is due to inconsistencies in behaviour and similarities in behaviour between truthful and untruthful reviewers, as well as the increasingly sophisticated tactics employed by spammers (Paul and Nikolaev 2021). Additionally, detection models are complex, expensive, and sometimes limited due to data privacy requirements. These requirements mean not all platforms are able or willing to use the most effective up-to-date fake review detection systems. Furthermore, since automated fake review detection is trained using established markers of fakeness, fake reviewers are always a step ahead. Therefore, it is important we better understand the cues people rely on to identify fake consumer reviews and to use this knowledge to develop appropriate interventions to improve human performance.

To improve how well humans can identify fake reviews, we first need to know how they decide if a review is genuine or fake. Research shows that human judges generally assume reviews are truthful, and they are especially likely to trust negative reviews over positive ones (Azimi, Chan, and Krasnikov 2022). However, when consumers feel deceived by the seller in any way, their intention to rebuy a product lessens significantly (Wang et al. 2022). Identifying fake reviews can therefore significantly harm the business of legitimate sellers even when they make efforts to eliminate fake reviews.

A recent systematic review (Walther et al. 2023), shows that research on human fake review detection is divided into different focus areas – employing a diverse array of theories and methods – and borrowing mostly from existing theories rather than developed specifically for this new problem (Ansari, Gupta, and Dewangan 2018; Ansari and Gupta 2019; DeAndrea et al. 2018; Jensen et al. 2013; Luo et al. 2013; Román, Riquelme, and Iacobucci 2019). These focus areas centre on how individuals assess credibility, trustworthiness, and veracity in online reviews. While these assessments often rely on the same cues, the focus areas themselves reflect different conceptualisations of these terms. For this paper we define trustworthiness as mirroring the perceived competence, benevolence, and integrity of the reviewer, while credibility entails the experience with a product (Mayer, Davis, and Schoorman 1995; Mayer and Davis 1999). Veracity is defined by whether the review is believed to reflect honestly held opinions based on genuine experiences with a product (Zhang et al. 2016).

Walther et al. (2023) identify seven different theories that have been used to explain how humans detect fake reviews. However, most theories only addressed one of two key aspects: (1) what cues are used to indicate if something is real or fake (uncertainty reduction theory, speech act theory, credibility theory, language expectancy theory) or (2) how are those cues processed to make a veracity judgment (warranting theory, the elaboration likelihood model (ELM), cognitive dissonance theory). Subsequent research by Chatterjee et al. (2023) used a deductive quantitative approach to build a theoretical framework merging cognitive dissonance theory with cue identification. Additional research by Petrescu et al. (2022) used a multi-method approach guided by interpersonal deception theory and the persuasion knowledge model to address how fake reviews are processed.

According to the systematic review by Walther et al. (2023), the cues consumers use to determine whether a review is fake can be divided into one of five types: review characteristics, such as argument quality (Ansari, Gupta, and Dewangan 2018; Peng et al. 2016; Román, Riquelme, and Iacobucci 2019; Walther et al. 2023), textual characteristics, such as word count (Ansari, Gupta, and Dewangan 2018; Kronrod, Lee, and Gordeliy 2017), and contextual factors, including whether the review appears on a business's own website or on a third-party platform (Ananthakrishnan, Li, and Smith 2020; Munzel 2015, 2016), reviewer characteristics, such as user name, and seller characteristics, such as seller reputation (DeAndrea et al. 2018; Peng et al. 2016).

Walther et al. (2023) argue that while the existing literature gives valuable insights, it only partially answers the question which cues consumers use to identify the veracity of a consumer review. Firstly, the authors argue, most research is based on laboratory situations where the main task was the identification of (depending on the study) the trustworthiness, credibility or veracity of the consumer review. Secondly, since most of these previous studies are quantitative laboratory studies based on existing theoretical frameworks, this deductive approach may have limited the generation of novel insights into how consumers detect the veracity of consumer reviews in real-world shopping contexts (Walther et al. 2023).

An additional problem is that many of the theoretical models used in previous research have considerable conceptual overlap, underscoring the need to synthesise these different models in a simpler unifying model. For example, language expectancy theory, uncertainty reduction theory and credibility theory all suggest that an online review is perceived as suspicious if the review does not match the raters' schema, or mental model, of what an online review should be like (Walther et al. 2023). Furthermore, the different theories focus on different aspects of the detection process. In contrast to credibility theory, the ELM focuses on how consumer motivation and ability lead to (sub)optimal decisions depending on whether consumers are willing or able to process online reviews via a central or peripheral route (Walther et al. 2023). However, when synthesising different models, it is currently difficult to evaluate how different aspects of the models weigh into the judgement process. Therefore, developing a naturalistic theory based on consumer reports of actual shopping behaviour would allow triangulation between the newly developed theory and existing theories.

The Perceived Deception in Online Consumer Reviews framework (PDOCR) by Román, Riquelme, and Iacobucci (2019), along with the theoretical framework by Filieri (2016), address both what cues are used and how these are processed, while also capturing actual user behaviour outside a laboratory setting. However, the PDOCR was developed on the basis of 18 in-depth interviews using the Cognitive Dissonance Theory and Elaboration Likelihood Model as the initial theoretical framework (Román, Riquelme, and Iacobucci 2019). This deductive approach may have limited the capacity for novel insights in a similar way to the quantitative studies (Gilgun 2013). Deductive research begins with the assumption that an existing theory will apply to the data, followed by the testing of that theory (Armat, Assarroudi, and Rad 2018; Elo and Kyngäs 2008). In contrast, qualitative research produces more, in-depth data than questionnaires (Bedos et al. 2009). Despite the value of this qualitative work, it does not address the risk of deductive work failing to identify context-specific considerations (Armat, Assarroudi, and Rad 2018). Inductive research takes the opposite approach: rather than using data to test a theory, it involves collecting data to develop one (Elo and Kyngäs 2008). A particularly promising inductive method for theory generation is grounded theory.

There have been two attempts to apply grounded theory to gain insights into how consumers perceive reviews. Filieri (2016) purportedly utilised a grounded theory methodology to identify cues for detecting fake reviews and explore the impact of these reviews on consumer decision-making on TripAdvisor. However, the authors indicated that credibility theory served as an organising framework in shaping their analysis approach. They also noted that the study was predominantly focused on triangulating its findings. Therefore, their approach appears to have been primarily deductive rather than inductive. Hou and Zhu (2023) used grounded theory interviews to identify nine features (e.g. number of likes on comments, number of other products to be viewed in the store, and number of comments viewed) human judges used to evaluate the usefulness of online reviews to complement a multi-layer perception neural network. Since usefulness is closely related to, but not the same as veracity, this study gives valuable insights into how to improve the automated classification of reviews. However, it does not provide clear input into human fake review detection.

To sum up, consumers rely heavily on online consumer reviews without knowing their truthfulness. The automated systems in place are insufficient in detecting all fake reviews, which leaves consumers to detect the rest, with evidence suggesting they do so poorly. To help consumers improve at identifying real and fake reviews, we must understand how consumers determine the veracity of online reviews within the wider context of online shopping (Ansari, Gupta, and Dewangan 2018; Burgoon and Miller 2018; Walther et al. 2023). Since the research on this is scarce and mostly deductive (Walther et al. 2023), there is a need for exploratory inductive work. This inductive approach will allow the triangulation of participants' statements with the cues identified in prior research and help uncover cues which may have been missed by previous studies. Additionally, our study considers the wider context of shopping behaviour with two major benefits. First, it will help develop a unified theory that can be triangulated against extant theory and tested against actual user experiences. Second, it will allow us to see how veracity decisions fit into the wider context of actual purchasing behaviour outside of the laboratory.

2. Materials and methods

To investigate how consumers perceive the veracity of online reviews, and which factors they use to identify fake reviews, semi-structured in-depth interviews based on a grounded theory approach (Corbin and Strauss 2015) were held with online shoppers.

This research was approved by the university's ethics committee (approval number 220401) and aligns with the legal requirements of the researchers' country. All participants gave written informed consent in accordance with the Declaration of Helsinki. The materials and data, along with supplementary material for the analysis, results and discussion are available in the web appendix on the Open Science Framework (https://osf.io/ju5x7/?view_only=63bd9c2d2b054f6a9dc6f5788e342a88).

Sampling and Participants. Since almost everyone engages in online shopping (Statista Research Department 2024b, 2024a), the study's target group is heterogeneous in socio-demographic characteristics, such as age, gender and education level (Marshall 1996; Shaheen and Pradhan 2019; Statista Research Department 2024b). Therefore, to build a theoretical framework that is robust across socio-demographic differences of consumers and mirrors the behaviour of a wider range of people (Mason 2010; Robinson 2014) we sought a heterogeneous sample of Dutch and German online shoppers using social media advertisements (Corbin and Strauss 2015; Mason 2010; Robinson 2014). Furthermore, within the analysis we compared the developing theory constantly across all demographic groups. The sample was therefore a purposive, partially theoretical, convenience-based sample (Corbin and Strauss 2015; Marshall 1996; Shaheen and Pradhan 2019).

To participate in the study, participants needed to be older than 18 years, live in Germany or the Netherlands, speak Dutch, German or English, and have shopped online at least twice in the last year.

The final sample consisted of 25 participants (for demographic information, see Table 1). The participants had a mean age of 34.84 years ($SD = 12.12$ years), with the youngest being 19 and the oldest 58 years old. Frequency of online shopping varied between two to three times per week and once every four months, indicating that the sample was heterogeneous regarding shopping frequency.[1] After 15 interviews a theoretical saturation was reached, since there were only few refinements to the theory from this point on (Mason 2010; Urquhart 2013). However, to ensure our findings were robust, we interviewed the final 10 interviewees to achieve our desired heterogeneous sample (Creswell and Poth 2016; Mason 2010).

Interview Procedure. The interviews were conducted by the first researcher. Before the interview, each participant received a consent form including consent to video and audio recordings which they signed either electronically or in person. Since participants lived in different parts of Germany and the Netherlands, 21 interviews were conducted and recorded via video call using Webex[2] software. After the interviews, video-recordings were immediately converted into voice-only mp3 files with the help of VLC Plus Player.[3] In accordance with GDPR guidelines the anonymised transcript files and the consent forms were the only materials stored post interview. The interviews were between 21 and 68 min ($M = 49.22$; $SD = 11.99$).

Table 1. Overview over socio-demographic characteristics of participants.

Socio-demographic characteristics	Type	Number of participants
Gender	Male	15
	Female	10
Nationality	Dutch	8
	German	16
	Indian	1
Residency	The Netherlands	10
	Germany	15
Profession	Students	2
	Higher education /research	6
	White-collar	11
	Blue-collar	6
Highest Education	Highschool	7
	Bachelor	7
	Master/state examen	8
	Professional education	3

Note: to ensure protecting the identity of the participants, there will be no demographic listing per participant.

Interview Guideline. Semi-structured interviews were chosen to ensure that the research questions could be thoroughly explored, while maintaining the flexibility consistent with a grounded theory approach. We could iteratively incorporate the insights of the current and of previous participants (Corbin and Strauss 2015). That is, this approach allowed us to incorporate the emergent insights of participants, following their priorities, exploring their meanings through further elaboration (Karatsareas 2022), and to ask specific questions to challenge our developing theory following the principle of seeking disconfirming cases (Booth et al. 2013; Corbin and Strauss 2015). It furthermore helped to maintain rapport by varying the question order to maintain conversational flow (Miller 2019). However, the guideline itself was not altered because it proved to be exhaustive.

To prevent priming participants with cues to fake review detection or a specific way of doing online shopping, we first invited participants to provide a full narrative of their online shopping process. We then built on their stories by asking what kinds of products they buy and how often they shop online to confirm the heterogeneity of the sample. Afterwards, we asked how participants decided which products or services they buy.

We followed up by specifically probing their use of online consumer reviews – whether they read them, how much they trusted them, and how they determined their veracity (the interview-guideline can be found in the web appendix on OSF).

Analysis. We used a grounded theory approach for analysis, meaning that the theory was developed through line-by-line coding of the data, thereby grounding it in the data (Biaggi and Wa-Mbaleka 2018; Corbin and Strauss 2015; Urquhart 2013). First, the audio files were automatically transcribed via the online service Amberscript[4] into verbatim transcripts. Afterwards, transcription errors were corrected and potentially identifiable information removed to anonymise the transcripts and to facilitate data familiarisation (Backman and Kyngäs 1999; Clarke and Braun 2017). The transcripts were then analysed in four steps using MAXQDA.[5] To verify initial interpretations of the data support the development of concepts, the data was constantly compared to previously coded data, and participants' experiences – based on their socio-demographic characteristics – was considered throughout all four steps (Corbin and Strauss 2015; Urquhart 2013).

Open Coding. First, transcripts were open coded by applying line-by-line descriptive codes (Biaggi and Wa-Mbaleka 2018; Corbin and Strauss 2015; Urquhart 2013). An example of the initial descriptive code is 'bad grammar untrustworthy' to describe that the participant interprets a bad grammar as a cue to detect fake reviews. Afterwards, codes that were similar to each other were merged into concepts (Corbin and Strauss 2015). For example, multiple codes that indicated participants were looking at grammar to determine the veracity and trustworthiness of a review were combined into a single code labelled 'grammar'.

Axial Coding. Second, to explore the relationships between the concepts the open codes were organised into sub-categories (Corbin and Strauss 2015). For example, some initial codes seemed to identify what features of reviews are thought to be important, while others identified what affected motivation to attend to those cues. Our analysis will show that these sub-categories cut across demographic strata.

Selective Coding. Third, the axial codes were integrated into more abstract codes around central core categories and developed precise coding definitions to refine the emerging theory (Corbin and Strauss 2015). To refer to the earlier example, the code 'language' was grouped into the core category 'Cues' which captures all aspects that are directly attended to during the veracity judgement. Developing selective codes were discussed and alternative explanations considered to ensure the emerging theory was robust and captured individual variation (Corbin and Strauss 2015).

The initial open coding process identified 412 codes. After selective coding there were 32 categories. After excluding categories that did not directly address the research question (Corbin and Strauss 2015), there were seven selective codes. Across these seven codes 'information sufficiency' appeared to be the core category, since the processes described by participants were centred around gathering enough information to make a judgement (Corbin and Strauss 2015).

Theory Construction. During theory construction disconfirming examples were sought to test the robustness of the developing theory to apparently contradictory data (Corbin and Strauss 2015). For example, ten participants indicated grammar was an important indicator of reviewer competence. However, one person disagreed indicating that issues such as dyslexia or emotional stress could affect grammar. This suggests that the participant does believe that grammatical errors can be indicative of untrustworthiness, but that grammar alone would not be sufficient to determine if a review is real or fake. Thus, the apparently disconfirming data rather supported and enriched the developing theory.

During theory development, the full research team discussed the developing theory to help check for bias and to improve the robustness of the developed categories and emerging theory (Corbin and Strauss 2015).

3. Results

3.1. Overview of the grounded theory

Our grounded theory of *how consumers determine the veracity and trustworthiness of online user reviews within the context of genuine purchases online* comprises seven selective codes, capturing different aspects of how the specific cues to veracity are applied in the wider shopping process. 'Information sufficiency' lies at the centre of the theory, capturing how the other processes in our model revolve around participants seeking enough information to make a safe judgement about whether to incorporate a review into their purchase decision. The other codes indicated in the analysis were: cues (used to judge the veracity of reviews), decision process (how purchase and product decisions are made), motivation to look at reviews (reasons why consumers look at reviews), number of reviews looked at (how many reviews participants read), impact of reviews on decision process (impact of different reviews on purchase decisions), value (how consumers define value and decide how much research to conduct) and word of mouth (the types of reviews participants read). The data highlight that veracity judgements sit within the wider context of an overall purchase decision. However, our analysis prioritises the processes directly linked to veracity judgements. Therefore, we will only briefly describe the decision-making process to outline the contextual moderators that impact the final veracity judgement, before focusing on the process of detecting fake reviews.[6]

3.2. Decision-Making process

The interviews showed different approaches to purchasing a product online. Although not all participants agreed on the exact order of the steps taken during an online purchase, there was nevertheless a distinct pattern. We capture this pattern in a newly developed theoretical framework, displayed in Figure 1.

The shopping process starts with the product choice, deciding which product type is suitable for the intended purpose, and ends with the purchase decision – choosing a specific product within the product type for purchase. Within both processes, reviews are used to appraise the quality and suitability of the product for their intended use (additional information about the decision-making process can be found in the web appendix on the OSF website).

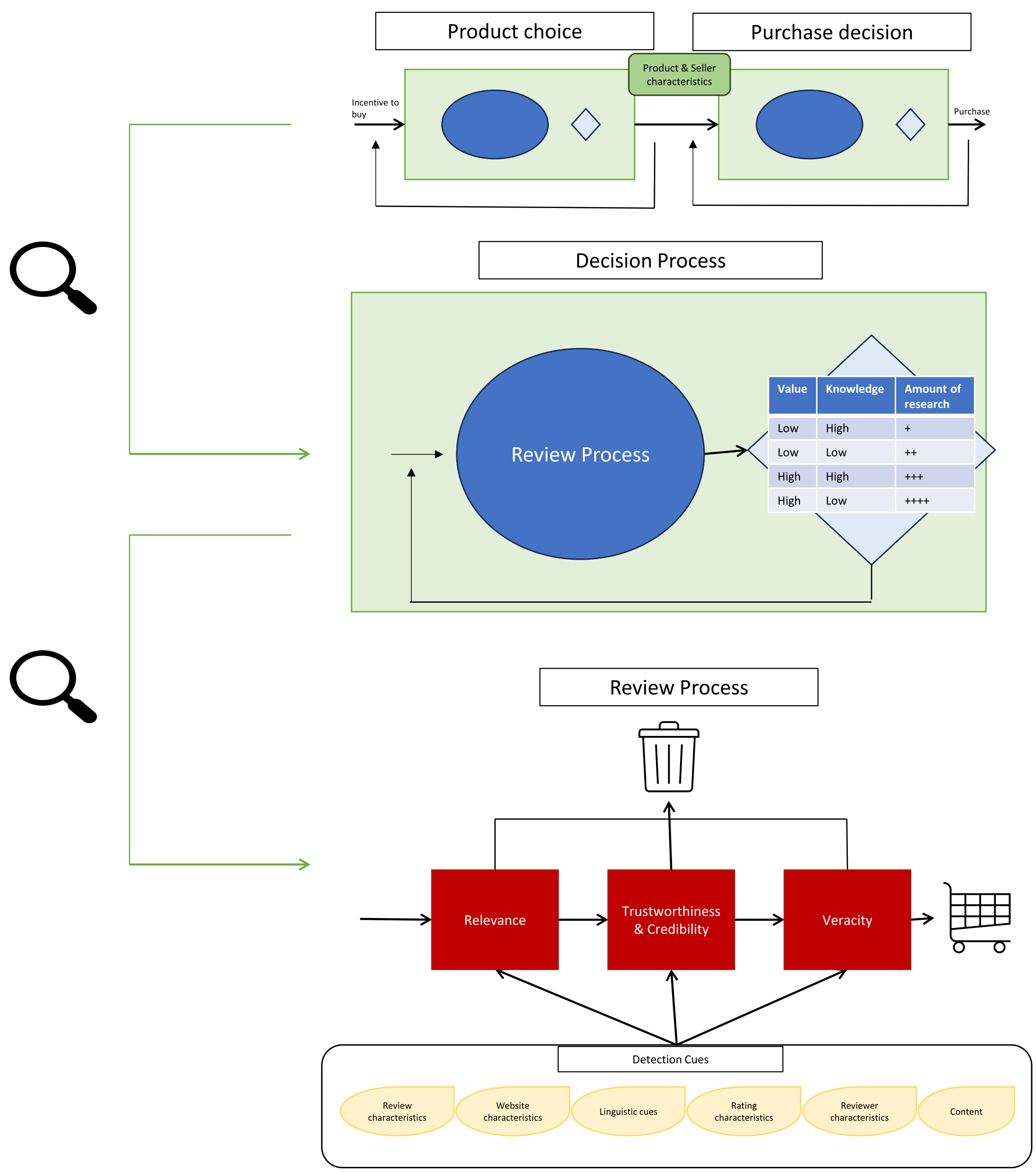

Figure 1. User-centred grounded theory of human detection of fake reviews.

3.3. Review process

Consumers analyse reviews based on relevance, trustworthiness, credibility, and veracity to determine its value. All participants indicated that this review process is often an automated process influenced by affect. Participants noted, for example, that they are not always aware of how they decide which reviews they perceive as relevant, trustworthy, credible, or true. They also indicated that more deliberate processing often only occurred when a relevant cue was perceived, for example, poor grammar, or the information contradicts verifiable information form another source. Additionally, while participants spoke of trustworthiness, credibility, and veracity interchangeably, their discussions about these concepts indicated these were different judgements. Therefore, at this point, a clear categorisation of cues that relate to specific judgements would decrease the validity of the analysis (Corbin and Strauss 2015). Nonetheless, there were cues that predominantly informed a specific type of judgement. We highlight these instances in our narrative synthesis.

We illustrate our theory in Figure 2, which shows that reviews are only incorporated into product choice and purchase decision if they are determined to be relevant, trustworthy, credible and true. If at any point the review is not deemed relevant, trustworthy, credible or true, the review is discarded.

Relevance of the Review. In the first step, participants indicated they skim through the content looking for relevance. To do so, they judge whether the content of the review fits to the product and is long enough to include relevant information. If the review is too short, for example comprising only three words or one sentence, most participants find it irrelevant as it provides insufficient information or detail. As participant 21 states ‘I *don't think I'm so much concerned with veracity right now, not with credibility, but if the review is helpful for me*’.

Subsequently, participants decide on relevance by reading the review carefully and comparing it to their own expectations and criteria, testing their assumptions. Therefore, at this stage, participants assess the product based on their assumptions, and the review tends to be evaluated fairly uncritically, primarily in terms of information provision and level of detail. Furthermore, as participant 16 explains, consumers search for information about ‘*the weighing, between the price difference [and] how much quality I need of the product*’. At this stage, most participants want information about the product, not the service from the seller. They only assess reviews addressing the quality of the seller at a later stage if they do not know the seller or have the same product from two different sellers. Photos of the products taken by the reviewer are considered useful, since people want to see what the product looks like.

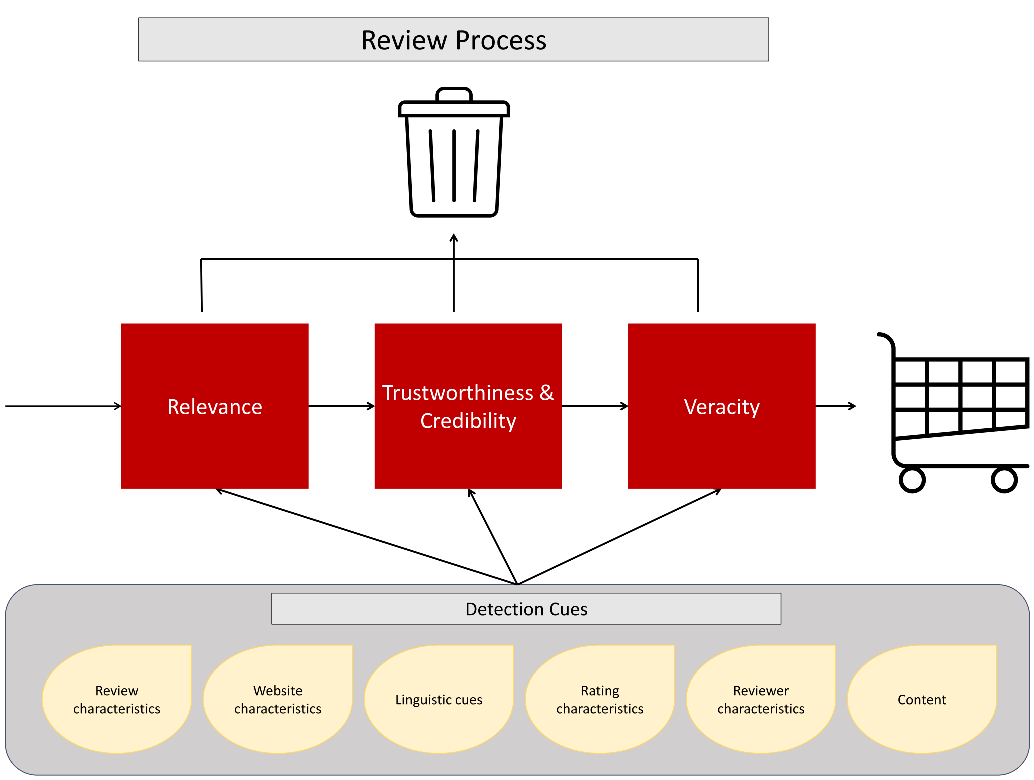

Figure 2. Enhanced review process as shown in Figure 1.

This is especially important when purchasing clothing. Participant 14 explains that if they '*post a picture of how [the clothes look] on them. And these are just normal women and not very thin models with inflated buttocks. So, then you can see what it looks like on a normal woman. So, I do always look at that*'.

3.4.1 Trustworthiness & credibility of a review

At the next stage – which can occur simultaneously or sequentially to the assessment of relevance – the trustworthiness and credibility of the review or reviewer are evaluated. This involves, for example, appraising the reviewer's motives and knowledge of the product. Participants' descriptions of these evaluations support a theoretical distinction between trustworthiness and credibility. We conceptualise trustworthiness as capturing whether reviewers are perceived as competent to evaluate the product, are motivated to help other consumers (rather than helping sellers, manufacturers, or other interested parties), and have integrity. The descriptions of trustworthiness thereby align well with Mayer's theoretical conceptualisation of trust (Mayer, Davis, and Schoorman 1995; Mayer and Davis 1999). In contrast, credibility assessments were primarily about whether reviews seemed to be based on genuine first-hand experience with a product. The participants mostly saw trustworthiness and credibility as overlapping concepts. A lack of direct experience with the product is associated with lower perceived competence, benevolence, and integrity.

To judge the trustworthiness and credibility of reviews and reviewers, participants considered different aspects of the reviews, such as argument structure, consistency in content between reviews, plausibility and whether multiple perspectives were considered. If the participant perceived the arguments to be implausible or the reviewer to be incapable or incompetent, the review was disregarded. As participant 24 states:

> *I assume, that one can identify which certain problems can occur per product, even with unspecific statements[-reviews]. So, you can tell from the [review]statement to what extent the user has prior training [experience with the product, or knowledge about technological products] or is perhaps part of the problem themselves [due to wrong use of the product].*

Veracity of the Review. In the last stage, participants consider the veracity of the review. While this can occur alongside judgements of relevance, trustworthiness, and credibility, it typically follows these judgements. Veracity differs from trustworthiness and credibility in that it is more specific: is the reviewer lying or telling the truth? Since most participants were aware of the occurrence of fake reviews, they developed different strategies to identify fake reviews. However, the larger the number of reviews, the more consumers believe the overall star-rating is truthful. As participant 12 comments '*when I see 728 [reviews], I tend to assume that most of them are truthful or authentic*'. If reviews are perceived as fake, these are disregarded from the judgement about the product.

3.4. Detection cues considered by consumers

Participants also described which cues they use to judge the relevance, trustworthiness, credibility, and veracity of consumer reviews. The detection cues can be categorised into the six overarching types of review features: review characteristics, reviewer-based, website characteristics, rating related, linguistic cues, and content (see Table 2).[7]

Review Characteristics. This category includes everything related to the reviews themselves, including but not limited to text structure, number of reviews, photos in the reviews and the date of the review. Participants indicated they prefer to buy products with multiple reviews because it reduces the influence of fake reviews on the overall rating. Furthermore, participants considered the date of the reviews, perceiving them as more realistic if there was a longer time gap between individual reviews. They assumed that if the reviews were paid or fake, the time gap between reviews would be shorter. Participant 16 states:

Table 2. Overview of sub-codes of the cues used to identify trustworthiness, credibility, and veracity of consumer reviews.

Sub-Code	Definition	Example
Review characteristics	Features that are related to the review outside of the content	Including but not limited to text structure, number of reviews, photos in the reviews and the date to the review
Reviewer based	Refers to characteristics of the reviewer	Profile pictures, the gender, the name of the reviewer, similarity with reviewer,
Website characteristics	This characteristic relates to where the review is placed	Business websites are less trustworthy than independent platforms; quality marks of the shop increase trust; Participants prefer to buy from Amazon, rather than third parties
Rating related	Features that are related to the rating of the review	Average rating, extreme ratings, and the valence of the review
Linguistic cues	Features that relate to the language, wording and subjective interpretation of the review	Grammar, spelling, Tone-of-voice, simplicity of wording/use of jargon, plausibility, objectivity, tangibility, validity, enthusiasm, reliability, and uniqueness of reviews.
Content	Features that are related to the content of the review	Photos of the products in the review, amount of detail, argument quality, and the depth of information

> "*Yes, sometimes the date. If there are a lot, for example, if the product [has] a lot of reviews of the same date in a row, but a long time ago or so, and has been online for a long time but then there was never another review. That so yes, that could also be bad sign.*"

Additionally, participants indicated that the number of sentences, sentence structure and sentence length influenced the perceived usefulness of the review. Participant 8 provides a clear example '*So the sentence structure in general, how the whole text is structured. So how argumentative, how argumentative they write, [if] it's more of a "everything pisses me off, full of shit", I don't take it seriously*'. Here, emotive language signals this review is probably unreliable as it presents only one side of the issue. Four participants also indicated that they trust a review more if they are marked as a verified purchase (on Amazon) or if it has been upvoted as helpful by other consumers.

Reviewer-based. This category refers to characteristics of the reviewer, such as profile pictures, gender, and reviewer name. Some participants argued that reviewer name influences how they perceive reviews. However, they also commented the name does not have to be the reviewer's real name. The participants preferred a nickname over names akin to random keystrokes. As Participant 4 states:

> *Whoever that is, is like "rdcgvhbjhuzft", because a cat ran over the keyboard. Then it's less trustworthy in my eyes than if it's called. LamasForever. It's not necessarily the best name, but for me it's more trustworthy than a cat running across the keyboard.*

Participants indicated that they tried to identify the reviewer's motive. For example, if a reviewer only expressed frustration without offering constructive criticism, such reviews were taken less seriously, as they rarely contained useful information. Participants also indicated that the similarity of the reviewer to themselves was important allowing them to relate more easily to the reviewer's perspective. For example, one similarity that made the reviewer more relatable was if the participant and the reviewer are both parents. Another important aspect was the perceived competence of the reviewer including their expertise and the realism of their expectations. This expertise and competence were mostly identified by how reviews were worded (see Linguistic Features). A last reviewer-based feature is that reviewer gender was used to evaluate expertise and experience for some products. Participant 8 stated:

> *Gender is also important to me, because when I look at a beauty product for [example] a face cream, I would rather have a woman's opinion. But when I read reviews about condoms, I'd rather hear the male opinion.*

Website Characteristics. Perceived trustworthiness of a review was influenced by where the review is placed, with some platforms trusted more than others. For example, participant 19 states:

> *And I think it also depends on the website because there are some websites who only allow you to put a review after you've bought it. So, if you have an account and they see from your account this product was purchased, only then you can put in a review of it. Then I trust the reviews a little bit more.*

Rating Related Features. This code captures evaluations of average ratings, extreme ratings, and review valence. According to the participants these features partially determine how seriously the review is taken. For example, the distribution of the star-ratings was an indicator of how many fake reviews were expected by the participants. Participants find reviews less plausible when there are only good reviews (3 or more stars), as often there is always someone that is not completely satisfied. Similarly, most participants prioritised reading negative reviews (2 or 1 star-rating) because these reviews were expected to give the most valuable information. Positive reviews are perceived as being more likely to be not genuine. Finally, participants who reported looking at all kinds of reviews – including positive, negative, and neutral – indicated that they consider the full spectrum of ratings. They also noted that they compare individual review ratings with the average rating and look for justification within the review to assess its usefulness.

Linguistic Features. This code includes features relating to the language, wording, and the subjective interpretation of the review. Specifically, language includes if the rules of the language are followed. For example, participant 25 clarifies how language can be used to assess competence '*If there are a lot of spelling mistakes, I don't believe them, then I don't care. Then I think you're stupid. I don't want any advice from you*'. The participants interpreted specific language cues as indicators of the reviewer's competence. These evaluations, in turn, influenced other judgments made by the participants: the higher the perceived competence of the reviewer, the more relevant, trustworthy, credible, and truthful the review was considered to be.

The wording of the review addresses individual preferences aside from the rules of written language. This includes types of words used, jargon, and tone. For example, if the tone is neutral, objective, and friendly or humorous, the review is perceived to be more honest. It was also articulated that wording can also contain clues to the problem-solving competence and the expertise of the reviewer, which then influenced the global evaluation of the review's trustworthiness. Additionally, to identify the veracity of the review, participants compared the use of words between reviews to identify duplicates, or reviews from the same reviewer under different accounts.

The subjective interpretation describes that people prefer reviews when the reviewer demonstrates some personality or humanness. However, this preference is tempered by reviewers who appear excessively emotional to an extent they may not be trusted to offer an objective product review. Participant 19 acknowledges limitations in how useful they find reviews, particularly when reviewers appear overly enthusiastic:

> *[…] like if someone says, "This shampoo changed my life", I'm like, I don't believe this. But if they say something like, okay, "I use the shampoo and my hair feels softer and more nourished", then I can believe it.*

Content. This category includes higher order evaluations of review contents. Specifically, rather than evaluation of specific features, this category addresses more global evaluation, for example, the amount of detail, the argument quality, and the depth of information. Participants said that the most important consideration was comparing content between reviews, so that claims within reviews can be directly tested against the evidence from other reviews. Another important aspect of content was the product specific details in a review. Participants indicated that the details mentioned in the reviews must be distinct to the product and not generic. Participant 6 describes it as:

> *For example, something that makes me wonder very quickly, if the whole [review] has a lot of adjectives that are meaningless. Like this. "Great. Wonderful." they say little about the product. If you can exchange [the adjectives] with each other product, where you have the feeling with a few key terms, you actually write that about everything, and it doesn't go into much detail about what the part actually is. If it doesn't go into much detail about the product.*

The more reviews contain distinctive details that are corroborated across multiple reviews, the more likely they are perceived to be trustworthy, credible, and genuine.

Impact of Reviews. Overall, we see that consumer reviews are used to support a decision to buy or not buy a product. As discussed, at each stage many factors influence the purchasing decision. However, even after the initial purchasing decision has been made, additional external factors can sway the choice, such as excessively long delivery dates. Either way, the review analysis is fed back into the holistic purchase decision as outlined in the overall model. Moreover, the way reviews are processed can be further biased by internal factors such as the individual's mood, knowledge, and perceived value of the product.

4. Discussion

This study develops a grounded theory of how consumers judge the veracity of reviews in the context of real-life purchases. We found that consumers follow certain protocols within a purchase decision, as illustrated in Figure 1. First, consumers determine which product type they want to buy, then decide which specific product is the best option. While gathering product information, consumers determine the usefulness of user reviews in a three-step process. First, the relevance of user reviews is determined. Only when a review is deemed (1) relevant, are its (2) trustworthiness, credibility, and (3) veracity evaluated. Therefore, as well as elaborating on 'how' consumers detect fake reviews, we also provide understanding of 'when' consumers consider review veracity, which is equally important in improving human fake review detection in naturalistic behaviour long-term.

4.1. Integration of theoretical framework & existing literature

While previous literature has prioritised identifying how consumers decide whether user reviews are genuine, our analysis shows that veracity is typically the last aspect of a review to be considered. Instead, consumers first consider review relevance. One reason people may prioritise relevance over veracity is to reduce cognitive load. When shopping online, consumers aim to achieve information sufficiency to reach a purchase decision. Especially, when the initial information about the product is limited, consumers actively seek a large amount of information from reviews. This continues until the consumer is satisfied, they are sufficiently informed to make a purchase. Since efficient decision-making is adaptive, people adapt their strategy to the task at hand (Lorch Jr and van den Broek 1997; Mézière et al. 2023). In our context we see this efficiency through shoppers evaluating consumer reviews in stages. Addressing relevance first filters out non-useful reviews before more effortful deliberation. In the event that the aforementioned criteria are not met, it is unnecessary to assess the veracity of the review.

Previous research on fake review detection has considered decision-making efficiency primarily through the Elaboration Likelihood Model (ELM) (Petty and Cacioppo 1986). However, our data indicates a decision-making process that conforms more closely to the heuristic-systematic model of risk processing (Chaiken 1980). The ELM prioritises consumer motivation and ability in provoking a shift from heuristic to more deliberative thinking (Petty and Cacioppo 1986). In contrast, the Heuristic – Systematic Model (HSM) better captures that the move from heuristic to systematic processing is motivated by consumer confidence in decision-making (Chaiken and Ledgerwood 2012). The lower the confidence an individual has in their decision, the greater the motive to engage in deeper processing and information seeking to achieve information sufficiency. Our grounded theory shows that the motivation to engage in cumulatively deeper levels of evaluation was directly motivated by participants' confidence in the sufficiency of the information they had to make a decision about whether to incorporate the review into their purchase decision or not.

An additional advantage of adopting a model with similarities to HSM over ELM is that the HSM makes explicit when heuristic and systematic processes can influence each other. Certainly, in our model, we see that cues processed more heuristically and more systematically can influence each other. Most of the identified cue categories appeal to the heuristic-mode of decision-making, and serve as short cut evaluations of relevance, trustworthiness, credibility and veracity. Only content cues are described as being processed in depth, and with deliberate assessment of corroboration between reviews and the verifiability of reported information. However, our participants did not indicate that the evaluations of these two different cue types were independent of each other, and both were used to come to an overall judgement.

Furthermore, our research findings align with the findings of previous studies, such as Ananthakrishnan, Li, and Smith (2020), who demonstrate the impact of the website on which a review is published and its perceived credibility. This is in accordance with credibility theory, which posits that consumers evaluate the trustworthiness of online reviews based on factors such as the source's expertise, objectivity, and the neutrality of the platform (Filieri 2016; Wathen and Burkell 2002). The present findings align with this theory, showing that consumers exhibit a greater tendency to place their trust in reviews when they perceive that the platform has an incentive to promote content that is authentic, as opposed to merely positive. The expressed sentiment is that independent platforms are generally held in higher esteem, yet it is also recognised that trust is bolstered when the platform's objectives are aligned with those of the consumer.

Our results add further nuance to existing research by integrating credibility theory and warranting theory (DeAndrea et al. 2018; Walther et al. 2009), which suggests information is seen as less valuable when heavily influenced by the sender. In the present study, consumers were observed to engage in a contemplative process of information filtration, evaluating not only the content of reviews, but also the motivations of both the reviewer and the platform. This behaviour underscores an awareness of the broader dynamics shaping review environments and reflects an effort to ensure the sufficiency and reliability of the information they use to make decisions. It contributes to the existing literature the importance of the broader shopping context in fake review detection.

If a review is considered relevant, it is then evaluated on two dimensions – trustworthiness and credibility – and only subsequently veracity. This new insight is important because it shows that participants are generally less concerned about the veracity of a review and more about the relevance and trustworthiness of a review. This could be because consumers are aware that fake reviews exist but also aware that it is hard to detect them (Cohen 2020; Plotkina, Munzel, and Pallud 2020). Therefore, consumers concentrate on information that is relevant and perceived as trustworthy and credible. This adds to existing literature by implicating that to improve the resilience of consumers against fake reviews and their confidence to detect them, the education of consumers might have to shift focus more towards fake review awareness and fact validation.

Participants in this study also indicated that (1) the less their knowledge about the product, the more reviews they consult, and (2) the less they know about the seller, the more they research the seller. This supports uncertainty reduction theory, which states that people aim to reduce uncertainty or increase the predictability of behaviour within interactions with other parties (Kramer 1999). Additionally, it again speaks to a heuristic-semantic processing of information. Participants sought additional information when the risk of making a wrong purchase decision is higher – either due to a higher value of or little knowledge about the product (Chaiken and Ledgerwood 2012). While Kusumasondjaja, Shanka, and Marchegiani (2012) and Racherla, Mandviwalla, and Connolly (2012) use uncertainty reduction theory to identify review detection cues (e.g. argument quality and reviewer identity), our theory adds that consumers also want to reduce uncertainty about the product and about the seller. They do this by searching for information about the product and the seller, to make an informed decision. Our theory thereby incorporates the benefits of uncertainty reduction theory, in that it classifies specific cues used to reduce uncertainty, while addressing that consumers rarely directly interact with review authors. Therefore, there can be no process of uncertainty reduction through interaction (Kramer 1999). Instead, our model more directly captures how consumers reduce the uncertainty by seeking information sufficiency from additional reviews and other cues on websites.

Consumers directly contrasting information contained in reviews with each other indicates the verifiability approach to lie detection may inform future interventions. The verifiability approach posits that, compared with deceivers, truth-tellers provide more details that can be confirmed (Nahari, Vrij, and Fisher 2014; Palena et al. 2021). We found that our participants naturally adopted this approach and reported using information that is theoretically verifiable to judge the credibility, trustworthiness, and veracity of a review. This also links to the HSM, because consumers only use more effortful verifiability with a deeper level of processing when they lack the expertise to easily verify information provided by the review, or when a specific fakeness cue is identified and triggers a shift from automatic to more deliberate thinking. Therefore, training consumers about what indicates a trustworthy and accurate review could foster the use of verifiability. This would encourage consumers to be less inclined to base trustworthiness decisions on affective appraisals and better equip them to make more systematic judgements for higher risk purchases (e.g. Dunn and Schweitzer 2005; Siegrist 2021). Additionally, the training would tap into behaviour that people are already performing, potentially leading to an increase in detecting verbal cues (Hauch et al. 2016). Moreover, compared to other methods of deception detection, the verifiability approach is comparatively simple to train (Palena et al. 2021). We add to the existing literature by confirming that participants assess the credibility, trustworthiness, and veracity of a review based on argument quality, the level of detail provided, and – most notably – details that suggest actual experience with the product.

4.2. Features identified as cues to deception

The features used to determine the trustworthiness, credibility, and veracity of a review, identified from participants in this study only partly overlap with those from earlier research (see Table 3).

Previous research has categorised detection cues into review characteristics, textual characteristics, reviewer characteristics, seller characteristics and characteristics of the platform where the review is displayed (Ansari, Gupta, and Dewangan 2018; Ansari and Gupta 2019; Filieri 2016; Walther et al. 2023). Seller characteristics are used by consumers to determine the value of the product and characteristics of the platform are used to determine the overall value of the reviews. However, both aspects are not within the scope of this research (both can be found in the web appendix). In this grounded theory, we also found that detection cues were categorised into review characteristics and reviewer-based characteristics. However there were differences elsewhere, rating related characteristics, website characteristics linguistic cues, and content. We contribute to previous research by differentiating between review characteristics and rating related characteristics. Participants sometimes combine information from the review text and star ratings – for example, when the content doesn't match the rating. More often, however, they consider these elements separately. For example, similar to Corbos, Bunea, and Breazu (2023), our participants reported sometimes not even looking at the content of reviews and only at overall star-ratings (e.g. with inexpensive products) and the number of reviews. Furthermore, participants differentiated between the content of the review and linguistic cues stating that the content of a review is independent of how the information is delivered. Similar to previous studies, participants indicated the reviewer, seller, and website characteristics are relevant to the purchase decision. However, they indicated that seller

Table 3. Categorisation of the different features from this study and existing literature.

Examples	Categories from Walther et al. (2023)	Categories from this study	Examples of features
Argument quality Message valence Quality of content Source quality	Review characteristics	Review characteristics Rating-related characteristics	Date of the review Number of reviews in total Average star-rating
Amount of detail Writing style Spelling grammar	Textual characteristics	Content Linguistic cues	Argument quality, sentence structure grammar and spelling, plausibility, tone
Identity-descriptive information	Reviewer characteristics	Reviewer based characteristics	Name, nationality
Seller reputation Trade record of the seller	Seller characteristics	Seller characteristics	Number of years active
Platform where the review appears	Characteristics of the platform where the review is displayed	Website characteristics	Secure website, website of manufacturer

characteristics are not used to determine the veracity of a review. Rather, seller characteristics are instead used to decide where to buy the product.

Therefore, we show that while existing research gives valuable insights into how fake reviews are detected, our grounded theory approach shows that in a realistic shopping situation, the veracity judgment process differs from laboratory settings. First, we show that veracity judgments are made at a later stage in the evaluation process. Secondly, we show that while cues manipulated in laboratory settings may effectively detect fake reviews, consumers do not necessarily rely on these cues when detecting fake reviews in real shopping situations.

4.3. Strengths, limitations and future research

A strength of our study is that we thoroughly sought disconfirming cases and a heterogeneous sample to increase the likelihood of a more generalisable theory (Creswell and Poth 2016; Marshall 1996; Mason 2010; Robinson 2014; Shaheen and Pradhan 2019). While the theory did not change after the 15th interview, as is typical in qualitative research (Guest, Bunce, and Johnson 2006; Morgan et al. 2002), to achieve heterogeneity in terms of age, and educational levels according to the target group, we expanded our search for a wider and more diverse range of different demographic backgrounds. We added ten more participants, reaching an appropriate sample size for grounded theory (Creswell and Poth 2016; Mason 2010). The first 15 participants that informed the initial theoretical framework were mostly aged between 20 and 30 years. Due to this homogeneity, we sampled more participants aged between 30 and 60 years to better represent the target population (Statista Research Department 2024b, 2024a). Reassuringly, the additional data supported the theoretical framework, indicating that while the socio-demographic characteristics differ, the shopping process does not. Thus, the theoretical framework is robust to socio-demographic variation (Mason 2010; Robinson 2014).

Another strength of our sample is that participants exhibited diverse shopping behaviours, which aligns with our aim to explore how reviews are used in authentic contexts. While we acknowledge the limitation that there may be nuances in specific types of purchases, we think that the diverse types of purchases discussed by our sample strengthens the theory. It is not limited to only specific products, but also includes repeat purchases, high and low value products, as well as gifts for third parties. A possible limitation is that the different layouts of the websites frequently used by the participants, including variations in how reviews and reviewer information are displayed – made it more difficult to identify common cues used to detect veracity. However, this also represents a strength as it shows that the heterogenous sampling was successful. The results are more likely to be generalisable, given that the findings are not limited to specific websites (Mason 2010; Robinson 2014).

While the interviewer used open-ended questions, and gave the participants time to think about their answers, the participants had to answer from memory and analysed their own automatic behaviour in retrospect (Bishop and Yardley 2017). While this is often the case for self-report interviews, there is room for bias and misremembering (Huber, Hill, and Lenz 2012). We attempted to minimise these limitations through using open-ended, non-leading questions, and asking for clarification during the interviews to elicit accurate and full accounts. We also sought disconfirming examples and used a constant comparative process so that analyses would not be unduly reliant on specific utterances without further verification (Huber and Power 1985; Lopez and Whitehead 2013).

Future research should triangulate the theoretical framework with realistic simulations of purchasing behaviour and zoom in on the processing of reviews in laboratory settings and field studies to strengthen the theory and allow for effective improvement of consumer resilience. For example, future research could use think-aloud methods while doing online shopping to gather more insights into the real-time thoughts and judgments of consumers either in the lab or when making genuine purchases.

5. Conclusion

We used a grounded-theory approach with 25 interviews to explore how consumers determine the veracity and trustworthiness of online reviews in a genuine purchase context. Our research resulted in a new framework. This framework shows that contrary to existing research, veracity judgments are not a priority for consumers during the shopping process. Our key insight is that consumers first assess the relevance of reviews. If relevant, they then evaluate trustworthiness, credibility, and finally veracity.

This shows that research focussing on veracity judgements should consider the wider shopping process to gather accurate results. Additionally, the results indicate that research focused on the impact of fake reviews on the purchase decision should consider whether a review is considered relevant to the purchase decision before asking about the perception of the review's trustworthiness or veracity. Importantly, further research is needed to test and refine our framework based on retrospective recall of cognitive processes, for example by using thinking aloud methodology or experiments.

Notes

1. Participants describe that the frequency of their online shopping behaviour is dependent on external factors, such as an increase during the pandemic due to the lockdown, and the product type, food and books for example are more frequently purchased than electronics or clothes, and that some product types have a tendency to be purchased recurrently, such as re-fills for shampoos or basics such as T-shirts.
2. Manufacturer CISO, Webex Last Version 43.6.0.26407, https://www.webex.com/de/index.html
3. Manufacturer VideoLAN, VLC Plus Player Last Version 3.0.18, https://www.videolan.org/vlc/index.de.html
4. Paid Transcription service https://www.amberscript.com/en/
5. Manufacturer VERBI GmbH, MAXQDA Last version: Release 2022.7, https://www.maxqda.de/download
6. Quotes in the results section were translated into English with the help from DeepL and checked by the interviewer to ensure that the content of the translations accurately reflected participants original intended meanings. DeepL does not store any data and the translation process does not violate the agreements made with participants when they consented to participate.
7. We also identified seller characteristics as a feature type, however, since these are used to inform the purchase decision and not the review judgement, it does not fall within the scope of this study. We describe this category in the web appendix.

Acknowledgements

The first author, Michelle Walther, wrote the majority of the article as part of her PhD project conducted at the University of Applied Sciences Bonn-Rhein-Sieg and the University of Twente. The idea of the design and research direction came from the first author. The supervisors, Steven James Watson, Alexander Boden and Marielle Stel were involved in the detailed design idea and throughout the whole project, the editing process and approval decision to publish the article. The first author, Michelle Walther, receives a scholarship from the University of Applied Sciences Bonn-Rhein-Sieg to work on her PhD part-time. Supplementing materials are included in the web appendix accompanying the online version of this article. We also want to thank the participants of this study and the internal and external reviewers that helped improve the manuscript.

Author contributions

CRediT: **Michelle Walther:** Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing; **Steven Watson:** Supervision, Writing – review & editing; **Alexander Boden:** Supervision, Writing – review & editing; **Marielle Stel:** Supervision, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Hochschule Bonn-Rhein-Sieg University of Applied Sciences.

Data availability statement

The data that support the findings of this study are openly available in [Open Science Framework https://osf.io/ju5x7/?view_only=63bd9c2d2b054f6a9dc6f5788e342a88].

ORCID

Michelle Walther http://orcid.org/0000-0003-1192-5351

References

Ananthakrishnan, U. M., B. Li, and M. D. Smith. 2020. “A Tangled Web: Should Online Review Portals Display Fraudulent Reviews?” *Information Systems Research* 31 (3): 950–971. https://doi.org/10.1287/isre.2020.0925.

Ansari, S., and S. Gupta. 2019. “Fake Reviews and Manipulation: Do Customer Reviews Matter?” *Research Papers*. https://aisel.aisnet.org/ecis2019_rp/143.

Ansari, S., S. Gupta, and J. Dewangan. 2018. “Do Customers Perceive Reviews as Manipulated?” *A Warranting Theory Perspective* 16.

Armat, M. R., A. Assarroudi, and M. Rad. 2018. “Inductive and Deductive: Ambiguous Labels in Qualitative Content Analysis.” *The Qualitative Report* 23 (1), http://eprints.medsab.ac.ir/300/.

Azimi, S., K. Chan, and A. Krasnikov. 2022. “How Fakes Make It Through: The Role of Review Features versus Consumer Characteristics.” *Journal of Consumer Marketing* 39 (5): 523–537. https://doi.org/10.1108/JCM-04-2021-4597.

Backman, K., and H. A. Kyngäs. 1999. “Challenges of the Grounded Theory Approach to a Novice Researcher.” *Nursing & Health Sciences* 1 (3): 147–153. https://doi.org/10.1046/j.1442-2018.1999.00019.x.

Bedos, C., P. Pluye, C. Loignon, and A. Levine. 2009. “Qualitative Research.” In *Statistical and Methodological Aspects of Oral Health Research*, edited by E. Lesaffre, J. Feine, B. Leroux, and D. Declerck, 113–130. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470744116.ch8

Beltzung, L., A. Lindley, O. Dinica, N. Hermann, and R. Lindner. 2020. “Real-Time Detection of Fake-Shops through Machine Learning.” *2020 IEEE International Conference on Big Data (Big Data)*: 2254–2263. https://doi.org/10.1109/BigData50022.2020.9378204.

Biaggi, C., and S. Wa-Mbaleka. 2018. “Grounded Theory: A Practical Overview of the Glaserian School | JPAIR

Multidisciplinary Research." *JPAIR Multidisciplinary Research* 32 (1): 1–29. https://doi.org/10.7719/jpair.v32i1.573.

Bishop, F., and L. Yardley. 2017. "Qualitative Assessment." In *Assessment in Health Psychology*, edited by Y. Benyamini, M. Johnston, and E. C. Karademas. Hogrefe Publishing GmbH.

Booth, A., C. Carroll, I. Ilott, L. L. Low, and K. Cooper. 2013. *Desperately Seeking Dissonance: Identifying the Disconfirming Case in Qualitative Evidence Synthesis.* https://doi.org/10.1177/1049732312466295

Brüggemann, P., and R. Olbrich. 2023. "The Impact of COVID-19 Pandemic Restrictions on Offline and Online Grocery Shopping: New Normal or old Habits?" *Electronic Commerce Research* 23 (4): 2051–2072. https://doi.org/10.1007/s10660-022-09658-1.

Burgoon, M., and G. Miller. 2018. "An Expectancy Interpretation of Language and Persuasion." In *Recent Advances in Language, Communication, and Social Psychology*, edited by H. Giles and R. N. S. Clair, 199–225. Routledge.

Chaiken, S. 1980. "Heuristic versus Systematic Information Processing and the use of Source versus Message Cues in Persuasion." *Journal of Personality and Social Psychology* 39 (5): 752–766. https://doi.org/10.1037/0022-3514.39.5.752.

Chaiken, S., and A. Ledgerwood. 2012. "A Theory of Heuristic and Systematic Information Processing." *Handbook of Theories of Social Psychology* 1:246–266. https://doi.org/10.4135/9781446249215.n13.

Chatterjee, S., R. Chaudhuri, A. Kumar, C. Lu Wang, and S. Gupta. 2023. "Impacts of Consumer Cognitive Process to Ascertain Online Fake Review: A Cognitive Dissonance Theory Approach." *Journal of Business Research* 154:113370. https://doi.org/10.1016/j.jbusres.2022.113370.

Chevalier, S. 2021. *Online Shopping Frauds UK* 2020. Statista. https://www.statista.com/statistics/1182221/online-shopping-fraud-incidents-uk/.

Clarke, V., and V. Braun. 2017. "Thematic Analysis." *The Journal of Positive Psychology* 12 (3): 297–298. https://doi.org/10.1080/17439760.2016.1262613.

Cohen, J. 2020, October 7. *How to Spot a Fake Review on Amazon*. PCMag UK. https://uk.pcmag.com/old-shopping/121470/how-to-spot-a-fake-review-on-amazon.

Corbin, J., and A. Strauss. 2015. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (4th ed.). Sage Publications, Inc.

Corbos, R.-A., O.-I. Bunea, and A. Breazu. 2023. "Influence of Online Consumer Reviews on the Sales of Large Household Appliances: A Survey in Romania." *Electronic Commerce Research*, https://doi.org/10.1007/s10660-023-09758-6.

Creswell, J. W., and C. N. Poth. 2016. *Qualitative Inquiry and Research Design: Choosing among Five Approaches*. SAGE Publications.

DeAndrea, D. C., B. Van Der Heide, M. A. Vendemia, and M. H. Vang. 2018. "How People Evaluate Online Reviews." *Communication Research* 45 (5): 719–736. https://doi.org/10.1177/0093650215573862.

Dunn, J. R., and M. E. Schweitzer. 2005. "Feeling and Believing: The Influence of Emotion on Trust." *Journal of Personality and Social Psychology* 88 (5): 736. https://doi.org/10.1037/0022-3514.88.5.736.

Elo, S., and H. Kyngäs. 2008. "The Qualitative Content Analysis Process." *Journal of Advanced Nursing* 62 (1): 107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x.

Filieri, R. 2016. "What Makes an Online Consumer Review Trustworthy?" *Annals of Tourism Research* 58:46–64. https://doi.org/10.1016/j.annals.2015.12.019.

Gilgun, J. F. 2013. "Grounded Theory, Deductive Qualitative Analysi, and Social Work Reserach and Practice." In *Qualitative Research in Social Work*, edited by A. E. Fortune, W. J. Reid, and R. Miller, 107–135. Columbia University Press.

Guest, G., A. Bunce, and L. Johnson. 2006. "How Many Interviews Are Enough? An Experiment with Data Saturation and Variability." *Field Methods* 18 (1): 59–82. https://doi.org/10.1177/1525822X05279903.

Hauch, V., S. L. Sporer, S. W. Michael, and C. A. Meissner. 2016. "Does Training Improve the Detection of Deception? A Meta-analysis." *Communication Research* 43 (3): 283–343. https://doi.org/10.1177/0093650214534974.

Hofmann, W., and L. Grigoryan. 2023. *The Social Psychology of Everyday Life*. Elsevier. https://www.sciencedirect.com/science/article/pii/S006526012300014X.

Hou, J., and A. Zhu. 2023. "Identification of Usefulness for Online Reviews Based on Grounded Theory and Multilayer Perceptron Neural Network." *Applied Sciences* 13 (9): Article 9. https://doi.org/10.3390/app13095321.

Huber, G. A., S. J. Hill, and G. S. Lenz. 2012. "Sources of Bias in Retrospective Decision Making: Experimental Evidence on Voters' Limitations in Controlling Incumbents." *American Political Science Review* 106 (4): 720–741. https://doi.org/10.1017/S0003055412000391.

Huber, G. P., and D. J. Power. 1985. "Retrospective Reports of Strategic-Level Managers: Guidelines for Increasing Their Accuracy." *Strategic Management Journal* 6 (2): 171–180. https://doi.org/10.1002/smj.4250060206.

Hussain, N., H. Turab Mirza, G. Rasool, I. Hussain, and M. Kaleem. 2019. "Spam Review Detection Techniques: A Systematic Literature Review." *Applied Sciences* 9 (5): Article 5. https://doi.org/10.3390/app9050987.

Jensen, M., J. Averbeck, Z. Zhang, and K. Wright. 2013. "Credibility of Anonymous Online Product Reviews: A Language Expectancy Perspective." *Journal of Management Information Systems* 30 (1): 293–324. https://doi.org/10.2753/MIS0742-1222300109.

Jindal, N., and B. Liu. 2008. "Opinion Spam and Analysis." Proceedings of the 2008 International Conference on Web Search and Data Mining: 219–230. https://doi.org/10.1145/1341531.1341560.

Karatsareas, P. 2022. "Semi-structured Interviews." *Research Methods in Language Attitudes*: 99–113. https://doi.org/10.1017/9781108867788.010.

Kim, J., J. Kang, S. Shin, and S.-H. Myaeng. 2021. "Can You Distinguish Truthful from Fake Reviews? User Analysis and Assistance Tool for Fake Review Detection." Proceedings of the First Workshop on Bridging Human–Computer Interaction and Natural Language Processing: 53–59. https://www.aclweb.org/anthology/2021.hcinlp-1.9.

Kramer, M. W. 1999. "Motivation to Reduce Uncertainty: A Reconceptualization of Uncertainty Reduction Theory." *Management Communication Quarterly* 13 (2): 305–316. https://doi.org/10.1177/0893318999132007.

Kronrod, A., I. Gordeliy, and J. K. Lee. 2023. "Been There, Done That: How Episodic and Semantic Memory Affects the Language of Authentic and Fictitious Reviews." *Journal of Consumer Research* 50 (2): 405–425. https://doi.org/10.1093/jcr/ucac056.

Kronrod, A., J. K. Lee, and I. Gordeliy. 2017. "Detecting Fictitious Consumer Reviews: A Theory-Driven Approach Combining Automated Text Analysis and Experimental Design." *Marketing Science Institute Working Papers Series*: 17–124.

Kusumasondjaja, S., T. Shanka, and C. Marchegiani. 2012. "Credibility of Online Reviews and Initial Trust: The Roles of Reviewer's Identity and Review Valence." *Journal of Vacation Marketing* 18 (3): Article 3. https://doi.org/10.1177/1356766712449365.

Lopez, V., and D. Whitehead. 2013. "Sampling Data and Data Collection in Qualitative Research." *Nursing & Midwifery Research: Methods and Appraisal for Evidence-Based Practice* 123:140.

Lorch Jr, R. F., and P. van den Broek. 1997. "Understanding Reading Comprehension: Current and Future Contributions of Cognitive Science." *Contemporary Educational Psychology* 22 (2): 213–246. https://doi.org/10.1006/ceps.1997.0931.

Luo, C., X. R. Luo, L. Schatzberg, and C. L. Sia. 2013. "Impact of Informational Factors on Online Recommendation Credibility: The Moderating Role of Source Credibility." *Decision Support Systems* 56:92–102. https://doi.org/10.1016/j.dss.2013.05.005.

Marshall, M. N. 1996. "Sampling for Qualitative Research." *Family Practice* 13 (6): 522–526. https://doi.org/10.1093/fampra/13.6.522.

Mason, M. 2010. "Sample Size and Saturation in PhD Studies Using Qualitative Interviews." *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 11 (3): Article 3. https://doi.org/10.17169/fqs-11.3.1428.

Mayer, R. C., and J. H. Davis. 1999. "The Effect of the Performance Appraisal System on Trust for Management: A Field Quasi-Experiment." *Journal of Applied Psychology* 84 (1): 123–136. https://doi.org/10.1037/0021-9010.84.1.123.

Mayer, R. C., J. H. Davis, and F. D. Schoorman. 1995. "An Integrative Model of Organizational Trust." *The Academy of Management Review* 20 (3): 709–734. https://doi.org/10.2307/258792.

Mézière, D. C., L. Yu, E. D. Reichle, T. von der Malsburg, and G. McArthur. 2023. "Using Eye-Tracking Measures to Predict Reading Comprehension." *Reading Research Quarterly* 58 (3): 425–449. https://doi.org/10.1002/rrq.498.

Miller, C. 2019. "Interviewing Strategies, Rapport, and Empathy." In *Diagnostic Interviewing*, edited by D. L. Segal, 29–53. Springer US.

Mohawesh, R., S. Xu, S. N. Tran, R. Ollington, M. Springer, Y. Jararweh, and S. Maqsood. 2021. "Fake Reviews Detection: A Survey." *IEEE Access* 9:65771–65802. https://doi.org/10.1109/ACCESS.2021.3075573.

Morgan, M., B. Fishhoff, A. Bostrom, and C. Atman. 2002. *Risk Communication: A Mental Models Approach*. Cambridge University Press. https://scholar.google.com/scholar_lookup?hl = en&publication_year = 2002&author = MG+Morgan&author = B+Fischhoff&author = A +Bostrom&author = CJ.+Atman&title = Risk +communication%3A+a+mental+models+approach.

Munzel, A. 2015. "Malicious Practice of Fake Reviews: Experimental Insight into the Potential of Contextual Indicators in Assisting Consumers to Detect Deceptive Opinion Spam." *Recherche et Applications En Marketing (English Edition)* 30 (4): 24–50. https://doi.org/10.1177/2051570715604155.

Munzel, A. 2016. "Assisting Consumers in Detecting Fake Reviews: The Role of Identity Information Disclosure and Consensus." *Journal of Retailing and Consumer Services* 32:96–108. https://doi.org/10.1016/j.jretconser.2016.06.002.

Nahari, G., A. Vrij, and R. P. Fisher. 2014. "The Verifiability Approach: Countermeasures Facilitate Its Ability to Discriminate between Truths and Lies." *Applied Cognitive Psychology* 28 (1): 122–128. https://doi.org/10.1002/acp.2974.

Ott, M., Y. Choi, C. Cardie, and J. T. Hancock. 2011. Finding Deceptive Opinion Spam by Any Stretch of the Imagination. *arXiv:1107.4557 [Cs]*. http://arxiv.org/abs/1107.4557.

Palena, N., L. Caso, A. Vrij, and G. Nahari. 2021. "The Verifiability Approach: A Meta-analysis." *Journal of Applied Research in Memory and Cognition* 10 (1): 155–166. https://doi.org/10.1016/j.jarmac.2020.09.001.

Paul, H., and A. Nikolaev. 2021. "Fake Review Detection on Online E-commerce Platforms: A Systematic Literature Review." *Data Mining and Knowledge Discovery* 35 (5): 1830–1881. https://doi.org/10.1007/s10618-021-00772-6.

Peng, L., G. Cui, M. Zhuang, and C. Li. 2016. "Consumer Perceptions of Online Review Deceptions: An Empirical Study in China." *Journal of Consumer Marketing* 33 (4): 269–280. https://doi.org/10.1108/JCM-01-2015-1281.

Petrescu, M., P. Kitchen, C. Dobre, S. Ben Mrad, A. Milovan-Ciuta, D. Goldring, and A. Fiedler. 2022. "Innocent until Proven Guilty: Suspicion of Deception in Online Reviews." *European Journal of Marketing* 56 (4): 1184–1209. https://doi.org/10.1108/EJM-10-2019-0776.

Petty, R. E., and J. T. Cacioppo. 1986. "The Elaboration Likelihood Model of Persuasion." In *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*, edited by R. E. Petty and J. T. Cacioppo, 1–24. Springer.. https://doi.org/10.1007/978-1-4612-4964-1_1

Plotkina, D., A. Munzel, and J. Pallud. 2020. "Illusions of Truth—Experimental Insights into Human and Algorithmic Detections of Fake Online Reviews." *Journal of Business Research* 109:511–523. https://doi.org/10.1016/j.jbusres.2018.12.009.

Racherla, P., M. Mandviwalla, and D. J. Connolly. 2012. "Factors Affecting Consumers' Trust in Online Product Reviews." *Journal of Consumer Behaviour* 11 (2): Article 2. https://doi.org/10.1002/cb.385.

Ren, Y., and D. Ji. 2019. "Learning to Detect Deceptive Opinion Spam: A Survey." *IEEE Access* 7:42934–42945. https://doi.org/10.1109/ACCESS.2019.2908495.

Robinson, O. C. 2014. "Sampling in Interview-Based Qualitative Research: A Theoretical and Practical Guide." *Qualitative Research in Psychology* 11 (1): 25–41. https://doi.org/10.1080/14780887.2013.801543.

Román, S., P. Riquelme, and D. Iacobucci. 2019. *Perceived Deception in Online Consumer Reviews: Antecedents, Consequences, and Moderators* (pp. 141–166). https://doi.org/10.1108/S1548-643520190000016010

Sánchez-Paniagua, M., E. Fidalgo, E. Alegre, and F. Jáñez-Martino. 2021. “Fraudulent E-Commerce Websites Detection through Machine Learning.” In *Hybrid Artificial Intelligent Systems*, edited by H. Sanjurjo González, I. Pastor López, P. García Bringas, H. Quintián, and E. Corchado, 267–279. Springer International Publishing. https://doi.org/10.1007/978-3-030-86271-8_23

Shaheen, M., and S. Pradhan. 2019. “Sampling in Qualitative Research.” In *Qualitative Techniques for Workplace Data Analysis*, edited by M. Gupta, M. Shaheen, and K. P. Reddy, 25–51. Hershey, PA: IGI Global. https://www.igi-global.com/chapter/sampling-in-qualitative-research/207789.

Siegrist, M. 2021. “Trust and Risk Perception: A Critical Review of the Literature.” *Risk Analysis* 41 (3): 480–490. https://doi.org/10.1111/risa.13325.

Spithoven, R. 2020. *Verbonden risico’s—Maatschappelijke veiligheid in de black box society*. Den Haag: Boom criminologie.

Statista Research Department. 2023. E-Commerce—Entwicklung des Umsatzes 2023. Statista. https://de.statista.com/statistik/daten/studie/3979/umfrage/e-commerce-umsatz-in-deutschland-seit-1999/.

Statista Research Department. 2024a. *Online Shops: Amazon Shoppers in Germany*. Statista. https://www.statista.com/study/68424/online-shops-amazon-shoppers-in-germany/.

Statista Research Department. 2024b. *Online Shops: Amazon Shoppers in the Netherlands*. Statista. https://www.statista.com/study/68419/online-shops-amazon-shoppers-in-the-netherlands/.

Trivedi, V. 2021, April 30. E-commerce Fraud to Surpass $20B in 2021, an 18% Jump Over Last Year, Report Finds. *Payments Dive*. https://www.paymentsdive.com/news/e-commerce-fraud-to-hit-20-billion-2021-an-18-jump-from-prior-year/599312/.

Urquhart, C. 2013. *Grounded Theory for Qualitative Research: A Practical Guide*. London: SAGE.

Walther, M., T. Jakobi, S. J. Watson, and G. Stevens. 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* 10:100278. https://doi.org/10.1016/j.chbr.2023.100278.

Walther, J. B., B. Van Der Heide, L. M. Hamel, and H. C. Shulman. 2009. “Self-generated versus Other-Generated Statements and Impressions in Computer-Mediated Communication: A Test of Warranting Theory Using Facebook.” *Communication Research* 36 (2): 229–253. https://doi.org/10.1177/0093650208330251.

Wang, X., A. R. Ashraf, N. Thongpapanl, and K.-Y. Wang. 2022. “Perceived Deception and Online Repurchase Intention: The Moderating Effect of Product Type and Consumer Regulatory Orientation.” *Journal of Consumer Behaviour* 21 (6): 1522–1539. https://doi.org/10.1002/cb.2109.

Wang, X., K. Liu, S. He, and J. Zhao. 2016. “Learning to Represent Review with Tensor Decomposition for Spam Detection.” In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, edited by Jian Su, Kevin Duh, and Xavier Carreras, 866–875. Austin, TX: Association for Computational Linguistics. https://doi.org/10.18653/v1/D16-1.

Wathen, C. N., and J. Burkell. 2002. “Believe It or Not: Factors Influencing Credibility on the Web.” *Journal of the American Society for Information Science and Technology* 53 (2): 134–144. https://doi.org/10.1002/asi.10016.

Winkler, R. L., and A. H. Murphy. 1973. “Experiments in the Laboratory and the Real World.” *Organizational Behavior and Human Performance* 10 (2): 252–270. https://doi.org/10.1016/0030-5073(73)90017-2.

Zhang, D., L. Zhou, J. L. Kehoe, and I. Y. Kilic. 2016. “What Online Reviewer Behaviors Really Matter? Effects of Verbal and Nonverbal Behaviors on Detection of Fake Online Reviews.” *Journal of Management Information Systems* 33 (2): 456–481. https://doi.org/10.1080/07421222.2016.1205907.