NBER WORKING PAPER SERIES

THE IMPACT OF FAKE REVIEWS ON DEMAND AND WELFARE

Jesper Akesson
Robert W. Hahn
Robert D. Metcalfe
Manuel Monti-Nussbaum

Working Paper 31836
http://www.nber.org/papers/w31836

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2023

We would like to thank the teams at Which? and The Behaviouralist for working with us to design, implement, and analyze the experiment. We also thank Iranzu Monreal, JingKai Ong and Senan Hogan-Hennessy for excellent research assistance. Sandro Ambühl, Rocio Concha, Chris Dellarocas, Chiara Farronato, Andrey Fradkin, Matt Gardner, Ginger Jin, Matt Kahn, Tai Lam, Jonathan Libgober, John List, Stephen McDonald, Dina Mayzlin, Davide Proserpio, Itzhak Rasooly, Chris Stanton, and Scott Wallsten provided excellent and helpful comments. Any opinions expressed in this paper are those of the authors and do not necessarily represent those of the institutions with which they are affiliated. This research was funded by Which?. Declarations of interest: none. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Impact of Fake Reviews on Demand and Welfare
Jesper Akesson, Robert W. Hahn, Robert D. Metcalfe, and Manuel Monti-Nussbaum
NBER Working Paper No. 31836
November 2023
JEL No. C90,D18,M30

ABSTRACT

Although fake online customer reviews have become prevalent on platforms such as Amazon, Google, and Facebook, little is known about how these reviews influence consumer behavior. This paper provides the first experimental estimates of the effects of fake reviews on individual demand and welfare. We conduct an incentive-compatible online experiment with a nationally representative sample of respondents from the United Kingdom (n = 10,000). Consumers are asked to choose a product category, browse a platform resembling Amazon, and select one of five equally priced products. One of the products is of inferior quality, one is of superior quality, and three are of average quality. We randomly allocate participants to variants of the platform: five treatment groups see positive fake reviews for an inferior product, and the control group does not see fake reviews. Moreover, some participants are randomly selected to receive an educational intervention that aims to mitigate the potential effects of fake reviews. Our analysis of the experimental data yields four findings. First, fake reviews make consumers more likely to choose lower-quality products. Second, we estimate that welfare losses from such reviews may be important—on the order of $.12 for each dollar spent in the setting we study. Third, we find that fake reviews have heterogeneous effects. For example, the effect of fake reviews is smaller for those who do not trust customer reviews. Fake reviews also have larger effects on those who shop online more frequently. Fourth, we show that the educational intervention reduces the adverse welfare impact of fake reviews by 44%.

Jesper Akesson
The Behavioralist
United Kingdom
jesper@thebehaviouralist.com

Robert W. Hahn
University of Oxford
and Carnegie-Mellon University
robert.hahn@smithschool.ox.ac.uk

Robert D. Metcalfe
Department of Economics
University of Southern California
Los Angeles, CA 90007
and NBER
robert.metcalfe@usc.edu

Manuel Monti-Nussbaum
The Behavioralist
United Kingdom
manuel@thebehaviouralist.com

A randomized controlled trials registry entry is available at AEARCTR-0008343

1 Introduction

E-commerce represented about 14% of all U.S. retail consumption in 2020–2021 (USCB, 2021), and online customer reviews play an important role in many of these purchasing decisions. Major retailers and search engines, such as Amazon, Apple, and Google routinely display reviews on their websites, and over 82% of American adults read customer reviews before purchasing products online for the first time (Smith et al., 2016). In theory, genuine online customer reviews can inform consumers about the quality of products and can reduce issues related to moral hazard and adverse selection (Bajari and Hortaçsu, 2004; Chevalier and Mayzlin, 2006; Dellarocas, 2006; Jin and Kato, 2006; Resnick et al., 2006; Cabral and Hortacsu, 2010; Anderson and Magruder, 2012; Klein et al., 2016; Luca, 2016; Belleflamme and Peitz, 2018; Acemoglu et al., 2019; Farronato and Zervas, 2019; Goldfarb and Tucker, 2019; Lewis and Zervas, 2019; Bonatti and Cisternas, 2020; Reimers and Waldfogel, 2021).

However, recent research shows that many reviews are *fake* and that they promote low quality—and sometimes even dangerous—products (Hu et al., 2011b; Anderson and Simester, 2014; Mayzlin et al., 2014; Luca and Zervas, 2016; KC and Mukherjee, 2016; Dwoskin and Timberg, 2018; Zinman and Zitzewitz, 2016; Which?, 2020c; He et al., 2021, 2022, 2023). For example, an examination of 720 million customer reviews on Amazon found that around 42% were fake or unreliable (Lee, 2020). Moreover, Lappas et al. (2016) show that fake reviews increase the visibility of products and services, which suggests that some companies may benefit from posting positive fake reviews.

Although fake reviews appear to be widespread, little is known about how they influence consumer demand and economic welfare. On the one hand, past research shows that (genuine) reviews can change consumer behavior and welfare (see, *e.g.*, Chevalier and Mayzlin (2006) or Reimers and Waldfogel (2021)), and the U.S. Federal Trade Commission argues that fake reviews reduce consumer welfare (FTC, 2021). On the other hand, it is also possible that consumers can distinguish between fake and genuine reviews, which may mean that fake reviews have small effects on consumer welfare.

This paper contributes to our understanding of how fake reviews influence consumers' demand and welfare by presenting the first incentive-compatible field experiment on fake reviews. We recruited a representative sample of 10,000 consumers in the United Kingdom (UK) and asked them to complete a common shopping task on an online platform resembling Amazon. The platform offers three types of products: dash-cams, headphones, or cordless vacuum cleaners.[1] Participants are first asked to choose a product category. Each participant

[1] These categories were selected because a leading UK consumer group, Which?, had documented the use of fake reviews for low quality products in these categories.

is then shown a 'search page' that displays five products in a particular category, such as five cordless vacuum cleaners, which we refer to as their 'consideration set' (Honka et al., 2019). They can then view one or more product pages before deciding which product to purchase. This type of shopping scenario is similar to how consumers use the Amazon platform (Forbes, 2019).[2] While all of the products have the same sale price, one has been classified as a *Don't Buy* product by the UK consumer protection organization Which?; one has been classified as a *Best Buy* product by Which?; and the remaining three received mediocre product ratings.[3]

In the field experiment, we examine the effects of the types of fake reviews that have been identified in the academic literature (see, *e.g.*, Kumar and Shah (2018)). For example, we make use of inflated star ratings. Inflating the star ratings of products is among the most common fake review strategies, and it has been documented in several instances, especially as product quality decreases (Hu et al., 2011b,a; Which?, 2020c). Another common feature that we study is overly positive reviews, which typically employ exaggerated language and repetitive phrases. Ott et al. (2011) and Anderson and Simester (2014) find that fake reviews are less descriptive of the actual product than real reviews, and Hu et al. (2012) shows that fake reviews are less precise in their description of the products and are generally more positive.[4]

A third set of features we incorporate is the use of 'sloppy' (but positive) fake written reviews, where the reviewer may admit that they were paid to write the review, or where they mistakenly reviewed the wrong product. Fake reviews with these characteristics have been found across several platforms (Which?, 2020c), and Dwoskin and Timberg (2018) shows that there is great variation in the sophistication in the language of fake reviews. Finally, we make use of a platform endorsement. Retail platforms, such as Amazon, often use algorithms to endorse products with good ratings (other platforms use the same approach to endorse workers). Some studies suggest that these algorithms may be beneficial (Adomavicius and

[2]The reason for developing a platform resembling a provider like Amazon is that commercial platforms were unlikely to agree to run a natural field experiment on this subject, and there was not sufficient exogenous variation present in observational data to answer our research questions. This online approach has also been used, for instance, to study how consumers respond to product taxes (Taubinsky and Rees-Jones, 2018). See section 5 for a discussion on the external validity of our experimental design.

[3]Which? is the largest consumer advocacy group in the UK. For more information see https://www.which.co.uk/. Consumers who use the actual Amazon platform do not see Which? product ratings, and we do not display these ratings to participants in our experiment (nor do we display the *Best Buy* or *Don't Buy* classifications). Participants do, however, have access to the same type of information that is displayed on Amazon's product pages, and can search online for more information if they wish.

[4]We conducted an assessment of the veracity of reviews when determining whether to include them in the experiment. More specifically, we classified a review as genuine if there was a strong positive correlation between the score from the product testing conducted by Which? and the content of the review. Reviews were deemed to be fake if there was no correlation or a negative correlation between the aforementioned factors, and if they exhibited typical signs of fakery. Of course, one can never be certain if a review that we collect online is fake or not, or if it is simply wrong or uninformed. Thus, our experiment estimates the effect of adding positive inaccurate reviews that are *very likely* to be fake. Ultimately, however, it does not really matter whether the reviews are truly fake or not, as long as they are similar to a fake review.

Tuzhilin, 2005; Panniello et al., 2014; Horton, 2017). For example, Gao et al. (2015) and Nosko and Tadelis (2015) show that promoting higher quality sellers leads to greater welfare, and Barach et al. (2020) find that cheap talk signals used by online platforms, such as "recommended", are effective in steering demand. We, however, wanted to determine if this type of endorsement might be welfare-reducing if applied to a low-quality product, which has been shown to occur in cases where a low-quality product receives many fake, overly positive reviews (Which?, 2020b).

We randomized consumers into six different groups. We then compared purchasing decisions across these groups. **Group 1**—the control condition—was only shown informative reviews (*i.e.*, reviews where the assessments are strongly and positively correlated with the quality of the product). **Group 2** was shown the same reviews as Group 1, but the *Don't Buy* product has inflated star ratings, distributed in a way that is typical for products with fake reviews (*i.e.*, mostly 5-star ratings). **Group 3** was shown the same information as Group 2, but with the addition of fake and overly positive written reviews on the product page of the *Don't Buy* product. **Group 4** was shown similar information to Group 3, with the main difference being that the fake written reviews are more easily identifiable as being fake (*e.g.*, stating the reviewer was paid for writing the review). **Group 5** was shown the same information as Group 3, with the addition of a platform endorsement for the low-quality product.[5]

Our final treatment group (**Group 6**) was exposed to the same information as Group 5, but with the addition of an educational intervention that warned participants that some reviews were false, and it provided them with some tips on spotting fake reviews. The intervention appeared at the top of the search page and did not target a particular product, meaning that it should be straightforward to implement in other settings. Moreover, this intervention may have desirable properties as it tries to reduce the adverse effects of fake reviews without needing to remove those reviews. Glazer et al. (2020) provides a theoretical argument for why it may be welfare-decreasing to remove fake reviews, and Yasui (2021) hypothesizes that an educational approach may be better than trying to remove them. Such an

[5]Groups 3-6 saw slightly modified versions of fake written reviews that actually existed for these products on the Amazon platform at the time of the experiment. We made minor changes to the written reviews to ensure consistency across product categories. The star ratings that participants were shown in Groups 2-6 were, however, inflated and exceeded the true star ratings found on Amazon. Those in Group 1 were not exposed to any fake material, and they were only shown the 'true' Amazon star rating as well as genuine written reviews that we found on the Amazon platform. In other words, when designing the experiment, we tried to reduce deception to the greatest extent possible while still being able to answer our research questions. It is also important to note that we never told participants that the product pages and reviews that they were shown were real. We debriefed everyone at the end of the survey experiment on its purpose, and everybody who participated in the experiment was better off financially. They all received a flat fee for participation, and they had a chance of receiving a product of their choice for free (the fake reviews could only influence their choice of the free product). In conclusion, we believe that our research design is appropriate, as it helps us answer several key research questions while directly making our participants better off.

educational approach has also been recommended as a way of countering fake *news* (Lazer et al., 2018), but we have no prior experimental evidence on whether this approach changes behavior with respect to fake reviews.[6]

To translate differences in purchasing decisions across groups into estimates of changes in economic welfare, we conducted a companion survey with a representative sample of 1,000 UK adults. In the survey, we presented participants with full and accurate information about the products available in the main experiment, and we elicited their willingness to pay (WTP) for these products using the incentive-compatible Becker–DeGroot–Marschak mechanism. We used the WTP estimates to create a new outcome variable for our experimental analysis– the dollar value of the decisions participants made.[7] In other words, if a participant picks the *Best Buy* headphones in the main experiment and if we predict that they would value these headphones at $30 (if they had full information), the participant is recorded as having gained $30 in welfare (because their price is zero). If fake reviews induce this participant to instead choose the *Don't Buy* headphones—which for the sake of argument they would value at $10 if they had full information—we conclude that fake reviews reduced this consumer's welfare by $20 (*i.e.*, $30-$10). This calculation captures changes in welfare in our setting as participants have to choose a product, and the price of all products is zero (in reality, all products used in the respective product categories have around the same sale price).[8] While this is, to the best of our knowledge, the first attempt to estimate the welfare loss from fake reviews, our estimates do not capture the full welfare costs of fake reviews across the economy, but rather the welfare cost of fake reviews in this particular setting.

Our empirical analysis yields four main findings. First, we find that fake reviews influence consumer choice. More specifically, the inflated star ratings make consumers 5.8 percentage points (standard error of 1.2) more likely to choose the *Don't Buy* product (a 55% increase relative to Group 1's demand for the *Don't Buy* product, the condition with no fake reviews). We found that a one-star increase in the rating of a *Don't Buy* product increases demand for that product by 38%, and that the elasticity of demand with respect to stars for the low quality product is 1.21.[9] Adding the fake written reviews (*i.e.*, comparing Group 2

[6]There is ongoing research that tries to understand the effect of warnings related to partisan political information that is fake or misleading (*e.g.*, Chan et al. (2017); Pennycook et al. (2020); Andı and Akesson (2020)).

[7]For our welfare calculations, pound values are converted to dollars using the exchange rate on 1 December 2021 (£1 = $1.33)

[8]While it is common for searches to return several products at or near the same price point on Amazon, our setting is somewhat special as all products have *exactly* the same price.

[9]We obtain the elasticity of demand with respect to stars by computing the average percent increase in stars that individuals in Group 2 are exposed to and by comparing that to the average percent change in demand that we observe relative to Group 1. Our findings and effect sizes corroborate previous work that found that higher star ratings increase demand (Jin and Leslie, 2003; Anderson and Magruder, 2012; Luca and Zervas, 2016; Lewis and Zervas, 2019), and that employee ratings influence the hiring of workers (Pallais, 2014; Stanton and Thomas, 2016; Benson et al., 2020).

to Group 3) further increases the share of participants that purchase *Don't Buy* products by around 6 percentage points (standard error of 1.3), and reduces the share of consumers that purchase *Best Buy* products by around 3.6 percentage points (standard error of 1.5). Adding the 'sloppy' fake reviews has similar effects on consumption.[10] All of these results are robust to adjusting the standard errors based on multiple hypothesis testing.

To understand the mechanisms underlying the effects we observe, we elicited the following information: how confident participants were that they made the right decision, their perceived ease of making an informed choice, what information they based their choice on, whether they read (and trusted) the customer reviews, and how much time they spent on the shopping task. We find that those who were allocated to the sloppy fake review group (Group 4) were 3.8 percentage points less likely to say that they read the reviews ($p = 0.007$) than those in the control group. We do find evidence, with lower power, suggesting that those allocated to the inflated stars group (Group 2) were 3.1 percentage points ($p = 0.069$) less likely to say that they based their choice on the product description and were 2.7 percentage points less likely to say that they read the written reviews ($p = 0.055$).[11]

Second, we estimate that being exposed to both inflated star ratings and fake text reviews (comparing Group 1 to Group 3) translates into a welfare loss of around $.12 per $1 that consumers spend on the platform. A one-star increase in the rating of the *Don't Buy* products reduces consumer surplus by $.03 per dollar spent on the platform. The presence of fake text reviews (comparing Groups 2 and 3) reduces consumer surplus by $.07 per dollar spent on the platform. We also find that fake reviews have negative effects for those who choose the dash-cams, cordless vacuum cleaners, and headphone product categories, suggesting that our results may generalize to many different products. However, fake reviews have slightly larger effects for cordless vacuum cleaners (the most expensive product), suggesting that people are not better at identifying fake reviews as the cost of misinterpreting the fake review increases. Taken together, these results imply that fake reviews influence consumer demand and could cause meaningful consumer harm.

Third, we find that fake reviews have heterogeneous effects. More specifically, those who do not trust reviews are less likely to be influenced by them. For example, the inflated star ratings do not appear to have economically meaningful effects on these participants, and the

[10] A concern one might have is that individuals think that all customer reviews are accurate in the experiment, while they actually think that a lower proportion of reviews are accurate in the "real world". If this is the case, we may overestimate the effects of fake reviews on consumption, as consumers would be overly gullible in our experimental setting. However, we asked participants whether they trust reviews on Amazon and whether they trusted the reviews in the experiment, and we find a striking similarity in their answers to both questions, suggesting that this is unlikely to be a concern. We discuss concerns related to external validity in Section 4.

[11] More specifically, if those who did not purchase a *Don't Buy* product ended up not viewing the *Don't Buy* product page, they would not have been exposed to any fake reviews (with the exception of the inflated star ratings that appeared on the search page).

fake written reviews (Group 3) only increase the share that purchase *Don't Buy* products by 4 percentage points (*i.e.*, the effect of fake written reviews is 50% lower for this group relative to the general population). The loss in consumer surplus from fake star and written reviews (*i.e.*, comparing Groups 1 and 3) for this segment is $0.08 per $1 spent on the platform (for comparison, it is $0.13 for those who trust reviews). These results hold even when controlling for a range of demographic variables and the frequency with which participants shop online.

In addition, we find that fake reviews have larger negative effects for those who use Amazon (and e-commerce) more frequently. In other words, experience does not seem to translate into sophistication when it comes to fake reviews. Those who are more experienced also spend less time when completing the shopping task. This suggests that the problems associated with fake reviews will not solve themselves as consumers become more accustomed to shopping online. Rather, the opposite is more likely to happen, which emphasizes the need for initiatives aimed at curbing the presence—and effects—of fake reviews. This finding runs contrary to the literature on fake news, which shows that less experienced users are more likely to believe, share, and disseminate false information (see, *e.g.*, Guess et al. (2020)). It is also a surprising finding in light of work by List (2003), which suggests that consumers become better at navigating the complexities of markets the more they engage with them.

Fourth, we show that the educational intervention reduces, but does not eliminate, the negative effects of the fake reviews. More specifically, the intervention reduces the share that buy the *Don't Buy* product by 23%, increases the share that buy a mediocre product by 12%, and does not influence the share that buy the *Best Buy* product. Further, on average, being exposed to the intervention increases consumer welfare by $4, which means that it reduces the negative impacts of fake reviews by around 44%. In other words, the intervention increases consumer welfare by $0.06 per dollar spent on the platform.[12] We also find that the educational warning is effective for all types of consumers (there are, for example, no heterogeneous differences by income and education). The lack of heterogeneous effects is encouraging, as past research in economics suggests that educational warnings work best for highly-educated people, meaning that they may exacerbate economic inequalities (Hastings et al., 2013).[13] Our findings indicate that there may be simple steps that retailers can take to design their online platform to curb some of the adverse effects of fake reviews and improve consumer welfare. These findings are in line with research suggesting that the design of the online platform can affect search and demand (even in the absence of physical search costs) (Dinerstein et al., 2018).

[12]This treatment effect is approximately equivalent to the difference between those who do and do not trust the reviews. In other words, the intervention has roughly the same (positive) effect on consumer welfare as going from being trusting of reviews to being skeptical about their informativeness.

[13]It is important to remember that the intervention reduces the effects of fake reviews, *even though* the product with the fake reviews has been endorsed by the very platform that is displaying the intervention.

To better understand why the intervention changed behavior, we also examine whether it influenced the strategies that participants adopted, their beliefs, and their attitudes. Our analysis shows that those in the education intervention condition (Group 6) were 4 percentage points less likely to base their product choice on star ratings ($p = 0.019$) and that they were 4 percentage points more likely to base their choice on the brand of the product ($p = 0.017$) than those in Group 5. These effects are consistent with the recommendations presented in the intervention, which amongst other things emphasized that consumers should "Inspect the comments (don't rely on star ratings alone)".

This paper contributes to the literature on fake reviews and search behavior in four main ways.[14] First, our paper reinforces the findings of a directly related quasi-experimental study by He et al. (2021), which suggests that fake reviews are likely to harm consumers. That study finds that fake reviews are primarily applied to low-quality products and that their application is associated with increases in sales. Our results are also in line with those of Zhuang et al. (2018), which finds that adding fake positive reviews to the description of a hotel may increase consumers' stated intention to stay at that hotel.

Second, to the best of our knowledge, this study is the first to causally demonstrate how to reduce the negative effects of fake reviews. The intervention that we test bears some resemblance to low-cost digital interventions used to tackle the negative effects of fake *news* (see, for example, Andı and Akesson (2020); Pennycook et al. (2020)). Moreover, the intervention provides an easy-to-implement complement to efforts aimed at detecting and removing fake reviews (see Zhang et al. (2016) and Yao et al. (2017) for examples of how to improve the detection of fake reviews).[15] Our paper differs from previous literature in that we show that our educational intervention works equally for all types of people–young and old, rich and poor, and experienced and inexperienced.

Third, we find that it may be in the short-term interest of firms to produce fake reviews to increase expected profits, supporting previous research showing the reasons why firms might engage in misleading marketing (Mayzlin, 2006; Bordalo et al., 2015; Piccolo et al., 2018; Aköz et al., 2020; Yasui, 2021; Knapp et al., 2021). This finding holds particularly true as we find that fake reviews have greater effects on demand for those who shop on online platforms, like Amazon, more frequently.

[14]Our paper is also related to the broader literature on customer reviews and advertising, which includes studies on the number of reviews and review inflation (Aral, 2014; Filippas et al., 2018; Fradkin and Holtz, 2023), customer reviews and discrimination (Botelho and Gertsberg, 2021; Cui et al., 2020), and how companies respond to reviews (Chevalier et al., 2018). Our paper is also related to List (2006) who finds that private markets can solve the lemons problem through third-party verification on the quality of the product and reviews.

[15]Ananthakrishnan et al. (2020) conduct an experiment ($n = 238$) and find that restaurants with fake reviews that have been labeled as being fake are less likely to be chosen than restaurants without any fake reviews. However, they do not compare the 'labeling' intervention to a control group with unlabeled fake reviews, so it is impossible to assess the efficacy of the intervention.

Fourth, our study relates to the broader literature on how online platforms influence consumer search and demand. Some studies have used structural and experimental models to estimate how changes in the design of online platforms affect consumer search and demand (Dinerstein et al., 2018; Ursu, 2018; Hodgson and Lewis, 2020; Gardete and Hunter, 2020; Sahni and Nair, 2020; Lam, 2021). Our paper is consistent with these studies as it also shows that small changes in platform design can lead to meaningful changes in consumer demand.

The remainder of this paper is structured as follows. Section 2 outlines the experimental design. Section 3 discusses the empirical analysis and section 4 discusses the external validity of our results. Finally, Section 5 concludes and reviews areas for future research.

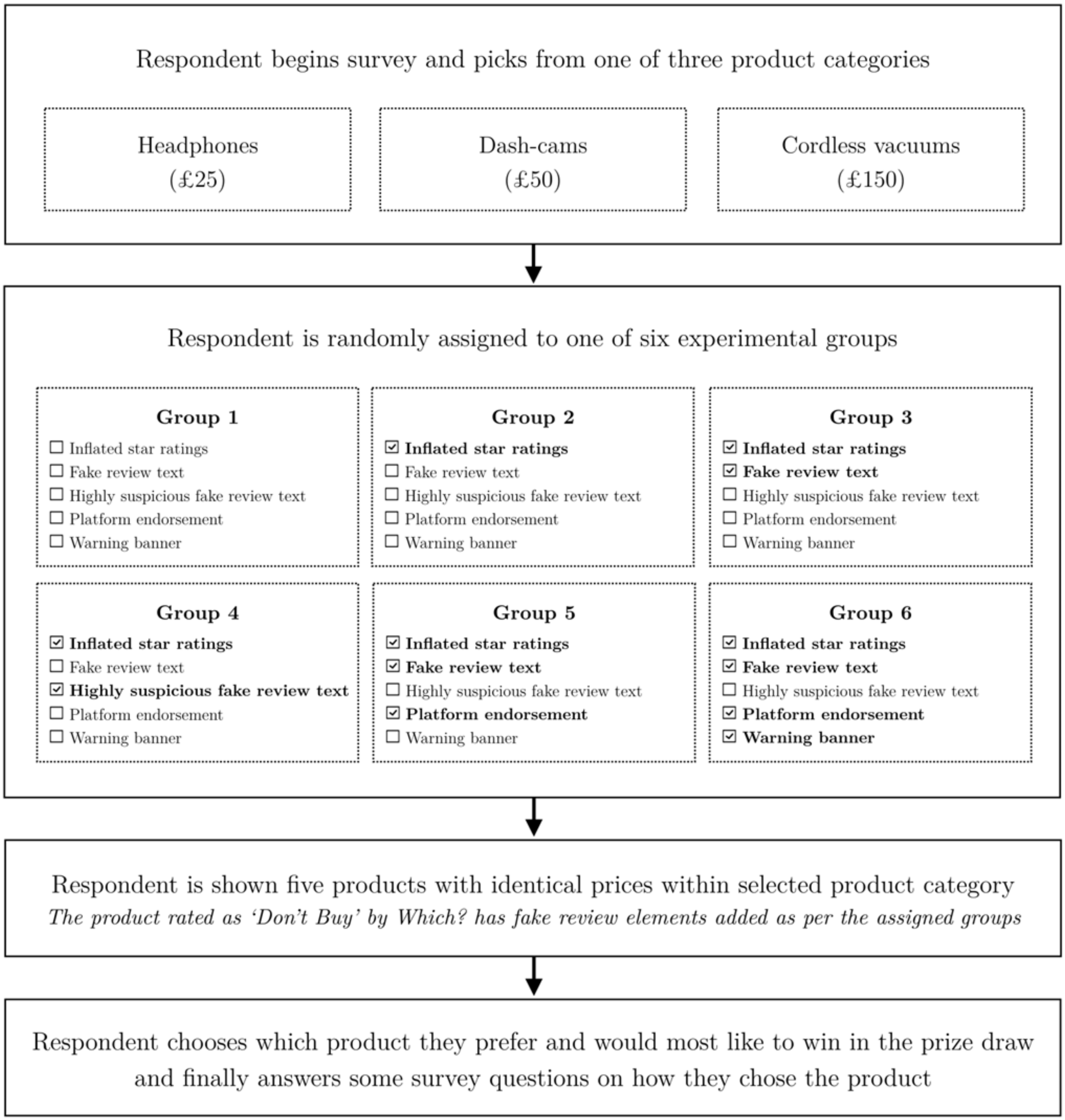
2 Experimental design

The field experiment was conducted in February 2020 and took place within an online survey coded using Qualtrics. We recruited a nationally representative sample of 10,000 UK adults to take part in the experiment. On average, the survey took 11 minutes to complete, and participants were paid £2 in exchange for their participation. Some randomly chosen participants also were sent the product that they choose. The sample was recruited via the panel provider Dynata, and respondents could complete the survey via desktop or mobile devices.

Participants began by responding to questions about their socio-demographic characteristics and online shopping behavior, and then completed a shopping task in an environment resembling the Amazon platform (please see the Appendix H for a full list of survey questions). After completing the shopping task, we asked participants questions about whether they searched for more information online when completing the task, whether they spent as much time and effort when choosing a product as they would have when normally shopping online, in addition to some questions about how they chose a product and their beliefs. We randomly allocated one sixth of the sample to a group that completed a version of the shopping task that did not include any fake reviews (Group 1). The remainder of the sample was allocated to one of five treatment groups, and these participants were exposed to fake reviews when completing the task (Groups 2-6). Those in Group 6 were also shown an educational intervention. The shopping task, and the randomization, are summarized in Figure 1 and the experimental conditions are described in Appendix E.

We also conducted a companion survey with a nationally representative sample of UK residents (n = 1,000) to measure consumers' willingness to pay for the products used in the experiment. Participants in this survey were provided with 'full information' about the products, and we then elicited their willingness to pay in an incentive-compatible way (using a

Figure 1: Summary of the experimental design



Notes. In this figure, we present the randomization and steps involved in the main experiment.

BDM mechanism). This survey was also coded using Qualtrics (please see the Appendix I the full survey used to elicit participants' willingness to pay).

In the remainder of this section, we describe the shopping task that those in Group 1 (the condition without fake reviews) completed, explain how the task varied for those in the treatment conditions, and discuss our data.[16] The section concludes with an overview of the methodology used in the willingness to pay survey.

2.1 Experimental groups

2.1.1 Group 1: No fake reviews

The shopping task began with participants choosing among three product categories: headphones, dash-cams, and cordless vacuum cleaners. These categories were selected because Which? had documented the use of fake reviews for low quality products in these categories. Once participants had chosen a product category, they were shown a search page that included five products from that category (these products constitute their consideration set). Participants were then asked to select the product that they preferred in that set and were told that they would be entered into a lottery to win the product of their choice. If they won, we would ship it to their home once they completed the survey.

Participants could gather information about the products by looking at the search page, as well as by clicking on and viewing the product pages for the five respective products in the consideration set. Figure 2 presents an example of what the search page looked like.[17]

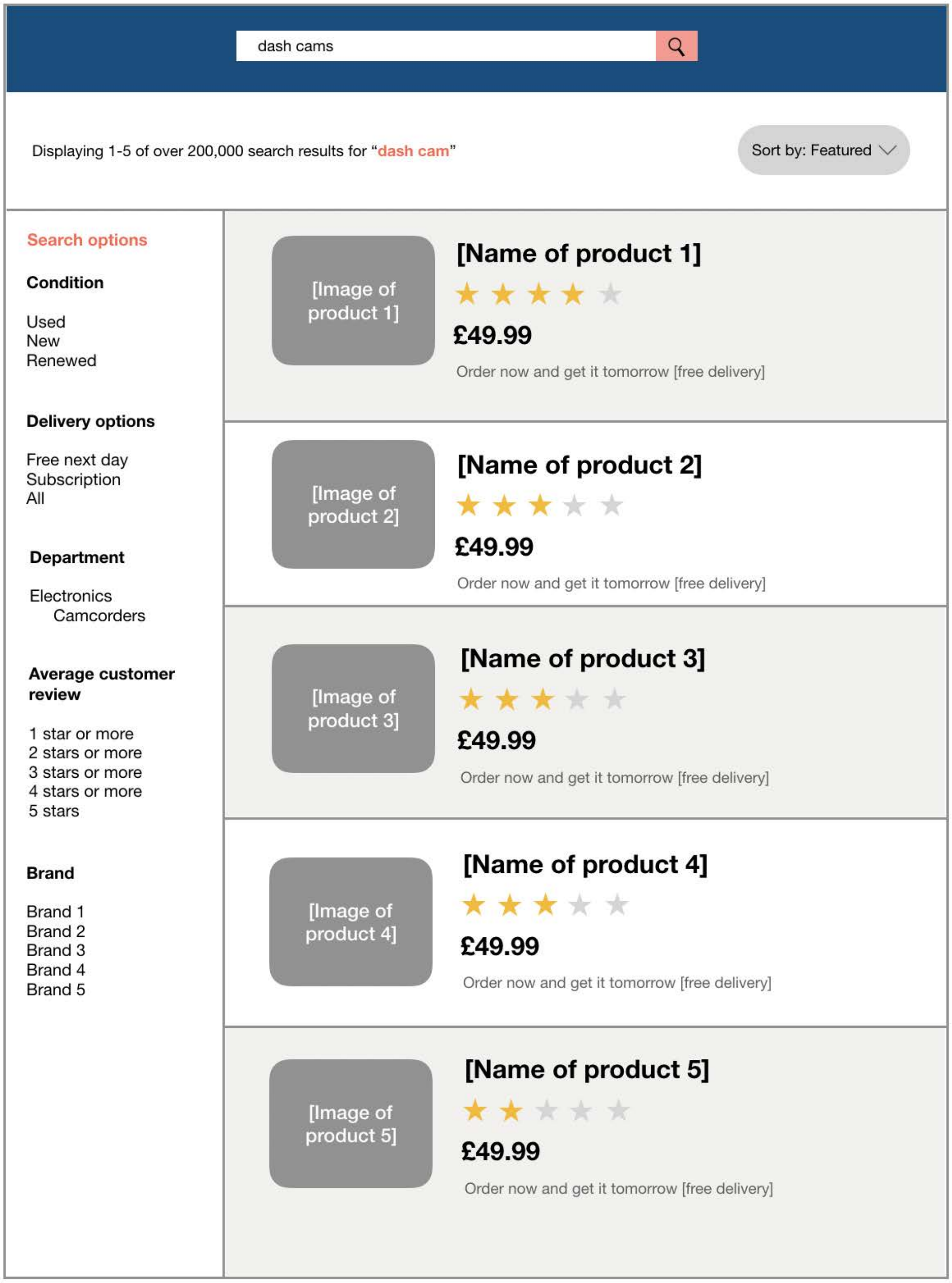
Participants did not have to view all the product pages. They did, however, need to navigate to a product page to select a product and complete the survey. The product pages included various types of information, such as seven written reviews, a star rating, and a product description. The ratings and written reviews were positively correlated with the true quality of the products.[18] Please see Figure 3 for an example of a shortened product page (the full page can be seen in Figure A2 in the Appendix).

[16]We refer to those allocated to Group 1 as "Group 1" throughout the paper for sake of simplicity.

[17]One drawback of our design is that we are unable to study if fake reviews and the educational intervention influence whether consumers purchase anything at all. While there are many situations where individuals have committed to buying something (*i.e.*, a situation like the one replicated in the experiment), it would nonetheless be interesting to conduct future experiments that allow for an "outside option".

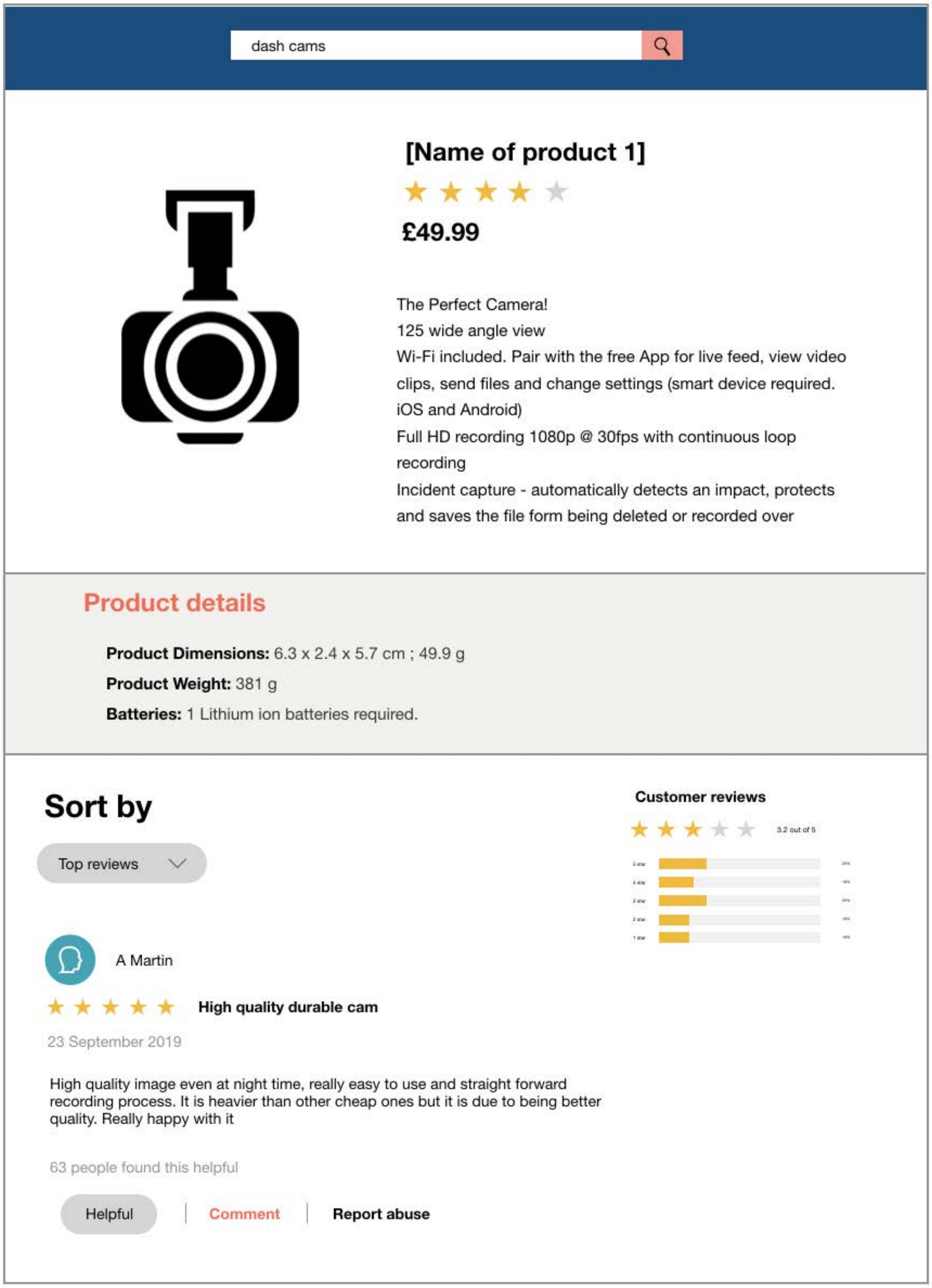
[18]The reviews displayed in the experiment were actual user reviews with real star ratings copied from the product pages on Amazon (these were collected in early February 2020). We only included written reviews that appeared to be real (*i.e.*, reviews that did not exhibit clear signs of being fake) and that aligned with Which?'s product ratings. The star ratings for these products were also positively correlated with the Which? product ratings.

Figure 2: Example of the search page



Notes. In this figure, we present a stylized example of the search page shown to participants.

Figure 3: Example of a product page



Notes. In this figure, we present a stylized example of a product page shown to participants, which they could navigate to from the search page.

The five products that were displayed on the search page had the same price in the experiment.[19] More specifically, the headphones were listed at £25, the dash-cams listed at £50, and the cordless vacuum cleaners all listed at £150 (or roughly $34, $68, $204 in dollar terms).[20] Further, within each product category, we included one product that Which? had classified as a *Best Buy* and one product that they had classified as a *Don't Buy*. The three remaining products always had product ratings in between the *Best Buy* and *Don't Buy* products. Participants were not shown any information about the Which? reviews or scores for any of the products, but they could have searched for this information online.[21] Please see Tables A17-A19 in the Appendix for a list of all products and product characteristics.

2.1.2 Group 2: Inflated star ratings

Group 2 was asked to complete the same shopping task as Group 1 (the condition with no fake reviews), with one main difference: the *Don't Buy* product had an inflated number of five star reviews, bringing its average rating to 4.8 out of 5 (an average increase of 1.43 stars). Those in Group 2 were shown artificially inflated star ratings for these products, and those in Group 1 saw the actual star ratings. We altered the distribution of the ratings for Group 2 by increasing the share of reviews with 5 stars and reducing the share of reviews with low ratings. More specifically, the *Don't Buy* dash-cam had a rating of 3.3 in Group 1, the headphones had a rating of 2.9, and the cordless vacuum cleaner had a rating of 3.8. All *Don't Buy* products had a rating of 4.8 in Group 2. In addition, the *Don't Buy* product was moved higher up in the search ranking, as this was sorted by star rating. Figure 4 compares the star ratings of the *Don't Buy* dash-cam between Group 1 and Group 2.

[19] The five products within each category were available at or close to the same price point online, which respondents could have discovered if they searched for the product names.

[20] Using the exchange rate between GBP and USD on 7 October 2021.

[21] Which? conducts rigorous testing, rates products, and publishes their recommendations and warnings on their website (Which?, 2020d). The very best products in their respective categories are awarded *Best Buy* status. Their tests are completely independent and based on benchmarks determined by impartial experts. Which? provides a warning about low-quality products by labeling them as *Don't Buy* products. Typically, the *Don't Buy* label is used for products with an overall test score of less than 40% or 45%, or when serious health and safety issues are detected by their tests.

Figure 4: Star ratings of the *Don't Buy* dash-cam in Group 1 and Group 2

Group 1: No fake reviews **Group 2:** Inflated star ratings

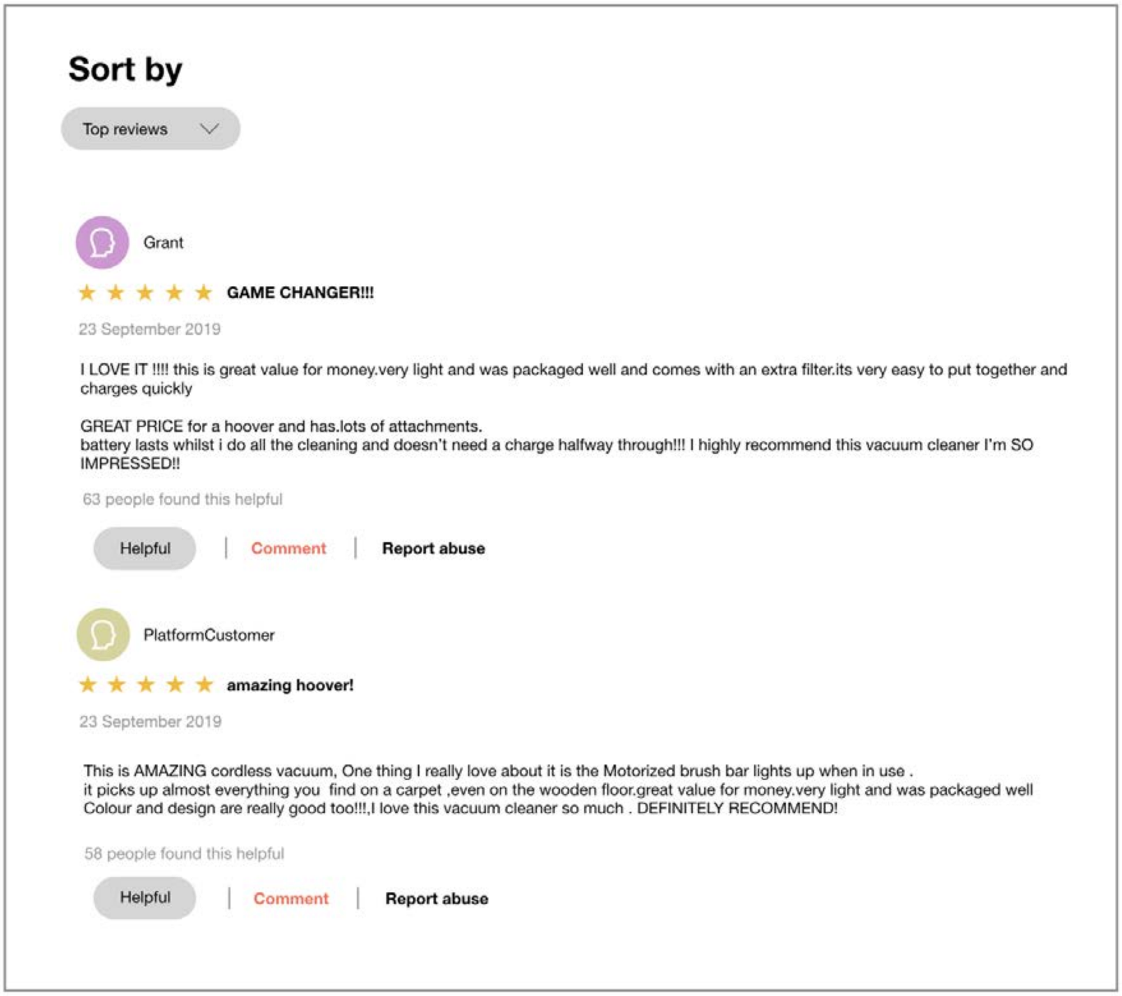
Notes. In this figure, we present the star ratings of the *Don't Buy* dash-cam product in two experimental groups. Those in Group 1 (the condition with no fake reviews) were shown the star ratings on the left-hand-side, and those in Group 2 (the inflated star ratings group) were shown the star ratings on the right-hand-side. The star ratings shown to Group 2 exhibit a distribution that could be suspicious, with a very large percentage of 5-star ratings and a very low percentage of 1-star to 3-star ratings.

Inflating the star ratings of products is among the most common fake review strategies, and it has been documented in several instances (Which?, 2020c; Hu et al., 2011b,a). While fake reviews could be applied to either high-quality or low-quality items, empirical evidence shows that the degree of fakery typically increases as product quality decreases (*i.e.*, the sellers of low-quality products seem to utilize fake reviews more than others) (Hu et al., 2011b). Intuitively, this makes sense as those selling high-quality products should have the smallest incentives to produce fake reviews.

2.1.3 Group 3: Fake review text

Group 3 was exposed to the same information as Group 2, with the addition of fake written customer reviews displayed on the *Don't Buy* product pages. More specifically, we replaced the written reviews that were displayed to Groups 1 and 2 with reviews that were highly favorable toward the *Don't Buy* product (as in Group 1, participants in Group 3 were also shown 7 written reviews per product). To avoid simply recreating the effects of positive reviews more generally, we deliberately included reviews with elements that might raise suspicion and that are common in actual fake reviews, including the following: exaggerated language, repetitive phrases and formatting, fewer 'verified purchase' reviews, several reviews left on the same date, the same reviewer leaving two reviews, a review left by someone called 'PlatformCustomer'. We also included one negative review contradicting the otherwise overwhelmingly positive feedback (this is common for products with many fake reviews) (Which?, 2020d).[22] Figure 5 presents an example of the fake review text used for Group 3.[23]

Figure 5: Example of fake written reviews in Group 3



Notes. In this figure, we present an example of two fake customer reviews that were shown to Group 3. These fake reviews contain various suspicious elements. More specifically, both reviews were left on the same date and contain the repetitive use of the exclamation point as well as improper and inconsistent use of punctuation and capitalization. Additionally, the second review was left by someone called 'PlatformCustomer'.

[22]Ott et al. (2011) and Anderson and Simester (2014) investigate reviews that cannot be real, finding that they are less descriptive of the real product than real reviews. Hu et al. (2012) also shows that fake reviews are less exact in their description of the products and are generally more positive (by textual sentiment).

[23]We did not include any other text, like a Q&A section, which has shown to be important for people's search (Banerjee et al., 2021).

2.1.4 Group 4: Highly suspicious fake review text

The shopping task for Group 4 was the same as the shopping task for Group 3, with the difference being that the written fake reviews were even more suspicious and exaggerated. More specifically, the written reviews for the *Don't Buy* products included: two five-star rated face cream reviews (even though the products were not face creams); positive reviews where the reviewer admitted to being offered money to leave positive reviews; and negative reviews claiming that the reviewer have been offered incentives to change their reviews. Fake reviews with these characteristics have been found across platforms (Which?, 2020c), and Dwoskin and Timberg (2018) shows that there is great variation in the sophistication of fake reviews. Figure 6 provides an example of the highly suspicious written fake review treatment. All of the written *Don't Buy* product reviews for the respective treatment conditions and product categories are displayed in Figures A7-A15.

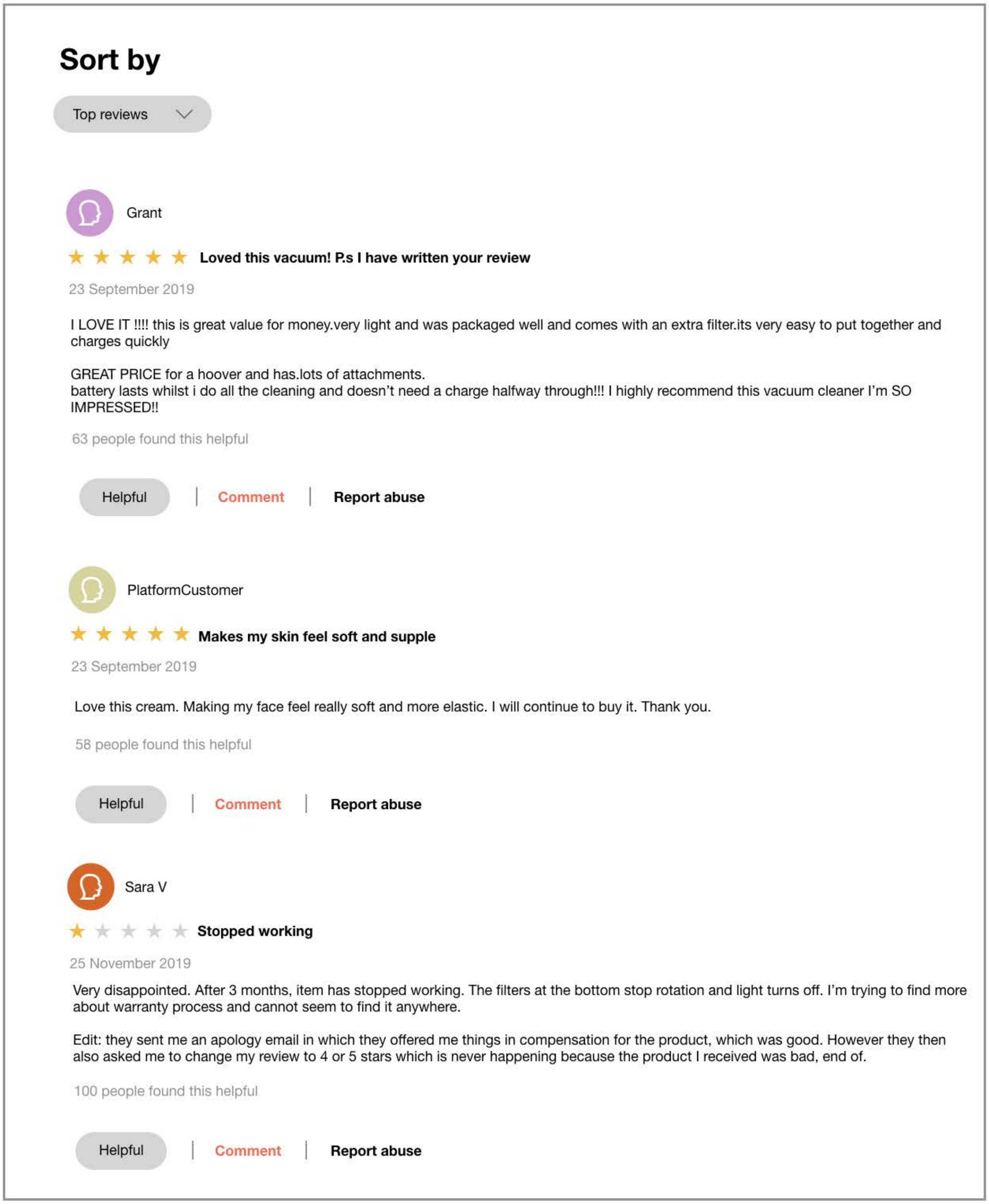
2.1.5 Group 5: Platform endorsement

Group 5 was asked to complete the same shopping task as Group 3, but was also exposed to a 'platform endorsement'. Some online review platforms provide endorsement labels for products or services that receive particularly good customer feedback, providing a potential route for those manipulating reviews to extend the influence of that manipulation. To investigate the impact of platform endorsements on consumer decision-making, we add an "Amazon's Choice"-style treatment for Group 5. While such endorsements are not an element of fake reviews *per se*, some of Which?'s investigative research has found instances where fake reviews had contributed to products receiving a platform endorsement (Which?, 2020a). The endorsement logos were placed next to the product on the search page and on the product page as well (please see Figures A3 and A4 for examples of how the endorsements were placed).

2.1.6 Group 6: Educational intervention

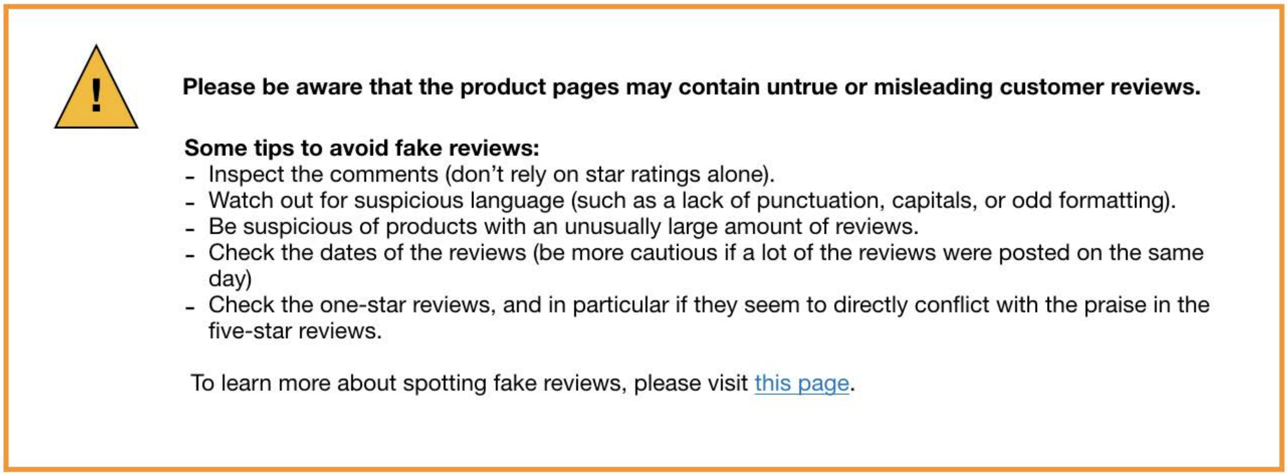
Group 6 received the same information as Group 5, but were also shown an educational intervention in the form of a warning banner. The banner, displayed in Figure 7, contained a warning about the possible presence of fake reviews, and offers advice to consumers on how to avoid being influenced by these reviews. It was added to the top of the search and product pages. The advice presented on the banner was taken from Which?'s guide on identifying fake reviews. Please see Figures A5 and A6 for examples of how the intervention was displayed on search and product pages.

Figure 6: Example of highly suspicious fake written reviews in Group 4



Notes. In this figure, we present an example of three fake written customer reviews that were shown to Group 4. These fake reviews contain some highly suspicious elements. More specifically, the first reviewer mentioned that he left the review as requested by the seller in exchange for a free gift while the second review is completely irrelevant as it is about face cream rather than any of the products included in the experiment. Additionally, the third review revealed that the reviewer was asked by the seller to change her negative review to a positive one.

Figure 7: Details of the warning banner



Notes. In this figure, we display the warning banner that was included at the top of the search and product pages for those in Group 6.

2.2 Data and outcomes of interest

All data for this experiment are collected through the online survey experiment and the companion willingness to pay survey. The main outcome of interest in this experiment is the product choice that the participant made. More specifically, we are interested in whether they chose a *Best Buy*, *Don't Buy*, or mediocre product. In addition to the outcomes of interest, we collected demographic data (*e.g.*, gender, age, education, income, and the region in the UK that they live in). We also collected data on participants' online shopping habits, whether they trust reviews on Amazon, the share of reviews on Amazon they think are fake, and whether they think it is easy to spot fake reviews on Amazon. These variables are used to conduct heterogeneity analyses. Moreover, we collected data on whether participants searched for more information online; whether they spend as much time when completing the task as they would have when normally shopping online; how confident they were in their decision, their perceived ease of making an informed choice; what information they based their choice on; whether they read customer reviews, and how much time they spent on the shopping task. Please see Appendix H for a full list of survey questions. The experimental groups are balanced on observables. We present balance tables and summary statistics for all variables in Appendix F.

2.3 Sample characteristics and statistics related to survey completion

The survey included 10,000 participants recruited via a reputable online survey panel (Dynata). The sample is representative of the UK population in terms of age, gender, and regional distribution of the population. Table A3 presents the demographic characteristics of the sample, and Table A4 presents key statistics about survey completion: respondents spent an average of 11 minutes on the entire survey; 51% answered the survey on a mobile phone; their average age is 43 years old; their average income is around £29,500 per year; and 51% are female. Most participants were experienced with online shopping: 97% report that they use Amazon, and over 80% shop online at least once a month. Moreover, 60% of participants claim that they spent as much time when completing the task as they would have when normally shopping online, and 18% said that they searched for more information online when completing the task (we do not find differences in these variables by treatment condition).

2.4 Relating participant decisions in the experiment to welfare

To understand the welfare implications of fake reviews, we also conducted a companion survey where we measure participants' willingness to pay for the products included in the experiment.

2.4.1 Methodology

We recruited a nationally representative sample of 1,000 UK adults to take part in the survey. The survey was coded using Qualtrics and participants were recruited via Prolific. The survey began by presenting users with instructions (see appendix for full instructions) explaining how we would be eliciting their willingness to pay and that it was in their best interest to report the true values.

Before eliciting their willingness to pay, participants were required to complete a short practice exercise and answer a few comprehension questions to make sure that they understood the instructions. In addition to ensuring that participants fully understood the exercise, the comprehension test (along with other variables such as the time spent reading instructions) allowed us to conduct robustness analyses that examine how the willingness to pay changes when those who do not understand the instructions are excluded.

We then proceeded with the survey by asking the participants to choose a product category, like in the main experiment. Based on the chosen product category, we showed the participants the five products included in the main experiment (in a random order) in a table. The table was designed to look like one typically seen on product comparison websites, which listed the product names, pictures of the products, and the following characteristics: brands, colors, weights or dimensions, Which? recommendations and test scores, and product-specific features.

We elicited participants' willingness to pay using the strategy proof Becker–DeGroot–Marschak (Becker et al., 1964) method. To ensure that participants truthfully reported their willingness to pay for all products, we explained that we would randomly pick one product and a random price point for that product, and then implement it (if they won the lottery). Those who won the lottery: 1) received the product and the difference between the lottery amount and the randomly chosen price point if their willingness to pay exceeded the randomly chosen price, or 2) the lottery amount if their willingness to pay was lower than the randomly chosen price.

Finally, after eliciting the willingness to pay, we asked the same set of demographic and background questions as in the main experiment (*e.g.*, gender, age, income, education, where participants live, how often participants shop online, and how often participants use Amazon).

The data collected from this companion survey allows us to verify whether the *Best Buy* product indeed dominates that *Don't Buy* product when consumers have full information.[24] It also provides us with an estimate of the magnitude of the benefit when participants switch from a *Don't Buy* to a *Best Buy* product.

Crucially, the data from this survey allow us to study the negative welfare impact of fake reviews and the positive welfare impact of the intervention. We do this by taking the 'full information' willingness-to-pay estimates from the survey and using these to create a new outcome variable that we use when analyzing the data from the experiment—the amount that participants would be willing to pay for the product that they chose if they had access to full, objective information about all relevant product characteristics. We calculate the value of this variable by conducting regressions that estimate the association between participants' characteristics and their willingness to pay for the respective products. We then use these predictive models to estimate participants' willingness to pay for the product they selected in the main experiment. We also use the simple average WTPs for the products in the survey, rather than a predictive model, as a robustness check.

To provide an illustrative example, imagine that all high-income individuals in the survey valued the *Best Buy* headphone at $30. Furthermore, we find that a high-income indi-

[24]We are, however, limited by the fact that the participants do not actually have the products in their hands and so they do not have *all* information. However, this choice situation is very similar to buying on Amazon where you cannot see and feel the product.

vidual in the Group 1 in the experiment (no fake reviews) picked the *Best Buy* headphones. We then conclude that their willingness to pay for the product they chose would have been $30 if they had access to full information. Now imagine that the same individual was placed in a fake review condition, and instead chose the *Don't Buy* headphones, which are assumed to be worth $10 to that individual, based on our survey. We would then conclude that fake reviews reduced this consumer's welfare by $20 ($30-$10). This calculation assumes that the products were the same price, which is the case in our experimental design (see Appendix D for a more formal derivation of the welfare model used here).[25]

2.4.2 Survey analysis

Table 1 presents the average willingness to pay (WTP) for the products tested in the experiment. As predicted, we find that the *Best Buy* products clearly dominate the *Don't Buy* products. Indeed, we also find that the *Best Buy* products dominate all mediocre products.

We conduct robustness analyses of participants' willingness to pay (please see the appendix), which involves dropping those who did not comprehend the instructions or who sped through the survey. These robustness checks do not meaningfully alter our results.

Finally, we conduct regressions predicting participants' willingness to pay for the products in the survey, which we use to estimate participants' willingness to pay in the main experiment (see Tables A7-A9 in the appendix). As can be seen in the appendix, the demographic characteristics significantly predict participants' WTP, and can thus be used when conducting multiple imputations.

[25]A limitation of our analysis is that the sample for the companion survey is not the same as the sample for the main experiment. Nonetheless, we believe that our approach can provide a reasonable estimate of the possible efficiency gains from interventions aimed at reducing the adverse effects of fake reviews because both surveys use a nationally representative sample. See the conclusion for a discussion of how these estimates might be refined.

Table 1: Average WTP for all products

Variable	Median	Mean (£)	Std Dev	N
		Headphones		
Best Buy	26	25.78	8.67	354
Don't Buy	6	6.49	6.96	354
Mediocre 1	14	15.63	9.63	354
Mediocre 2	14	15.37	8.33	354
Mediocre 3	18	17.20	9.25	354
		Dash cams		
Best Buy	59	53.67	19.27	288
Don't Buy	3	14.01	17.08	288
Mediocre 1	27	31.94	19.63	288
Mediocre 2	35	35.79	18.94	288
Mediocre 3	43	39.04	17.30	288
		Cordless vacuum cleaners		
Best Buy	129	131.88	54.18	356
Don't Buy	9	36.38	45.64	356
Mediocre 1	81	83.04	52.78	356
Mediocre 2	129	120.96	61.30	356
Mediocre 3	81	83.45	53.44	356

Notes. In this table, we present summary statistics relating to participants' willingness to pay for the fifteen products used in the main experiments (5 per product category). As in the main experiment, participants were first asked to choose a product category, after which we elicited their willingness to pay for the five products in their chosen category. Please see the appendix for the questions used to elicit participants' WTP.

3 Results

In this section, we present our analysis of the experimental data. We begin by examining the effects of fake reviews on participants' consumption choices. We then conduct a heterogeneity analysis that reveals if fake reviews have differential impacts for different population segments. Finally, we estimate the effect of the educational intervention on participant behavior. We conduct all analyses using linear probability models (LPM) and ordinary least squares (OLS). Three of our main outcomes are binary—whether participants chose a *Don't*

Buy, *Best Buy*, or a mediocre product.[26]

Our fourth (and final) outcome of interest is continuous—participants' willingness to pay for the product that they chose when completing the shopping task. We used two methods of coding this outcome. The first method involves conducting regressions associating participants' willingness to pay with their demographic characteristics in the companion survey, and then using these demographic characteristics to impute participants' predicted willingness to pay in the experiment. Using this method, the outcome of interest we use in our analysis below is the predicted willingness to pay for the product that participants ended up choosing, as this represents the value that participants obtained in the experiment (the products were free). The second method is similar, but instead involves imputing the average willingness to pay obtained in the companion survey (*i.e.*, we do not predict the WTP of individual participants, and instead simply use the average WTP that we measured for the product they chose–these values are listed in Table 1).

3.1 Fake reviews influence consumer behavior and reduce welfare

We begin by evaluating the effects of fake reviews on the products selected by participants and on consumer welfare. We do so by comparing the share of consumers in the respective experimental conditions that chose *Don't Buy*, *Best Buy*, and mediocre products, as well as their WTP for the product they chose. Note that the relevant comparisons differ for each group. Group 2 was shown the same information as Group 1, with the addition of inflated star ratings. This means that Group 1 is the relevant comparison group for Group 2. However, Group 3 was shown the same information as Group 2, with the addition of fake written reviews. This means that Group 2 is the relevant comparison group for Group 3 (if we want to learn about the incremental effect of adding fake written reviews). We can, of course, compare Group 3 to Group 1, but we then estimate the combined effect of both inflated star ratings and fake written reviews on consumer choice. Finally, the relevant comparison group for Group 4 is also Group 2, and the relevant comparison group for Group 5 is Group 3 (see Figure 1 for an overview of the relevant comparison groups for all conditions).

Table 2 shows that 10% of Group 1 (the condition with no fake reviews) chose the *Don't Buy* product (see the constant coefficient in column 1); in contrast 16% chose the *Don't Buy* product in Group 2 (adding that constant coefficient to the Group 2 coefficient in column 1). In contrast, Group 2 is significantly less likely than Group 1 to purchase a mediocre product, and is about as likely to purchase a *Best Buy* product. These findings demonstrate that inflated

[26] Effect estimates are rounded to the nearest integer value when we report our results. We also use GBP values throughout this paper because the experiment was conducted with UK participants.

Table 2: The effects of fake reviews

	(1) Chose *Don't Buy*	(2) Chose Mediocre	(3) Chose *Best Buy*	(4) WTP (predicted)	(5) WTP (average)
Group 2	0.058***	-0.041**	-0.016	-2.776	-2.931*
(G1 + fake stars)	(0.012)	(0.017)	(0.016)	(1.710)	(1.519)
	†††	††			†
Group 3	0.126***	-0.074***	-0.052***	-7.156***	-6.687***
(G2 + fake reviews)	(0.013)	(0.017)	(0.015)	(1.684)	(1.494)
	†††	†††	†††	†††	†††
Group 4	0.110***	-0.067***	-0.043***	-6.184***	-6.676***
(G2 + sloppy reviews)	(0.013)	(0.017)	(0.015)	(1.688)	(1.484)
	†††	†††	†††	†††	†††
Group 5	0.143***	-0.106***	-0.036**	-7.165***	-7.251***
(G3 + endorsement)	(0.013)	(0.017)	(0.015)	(1.668)	(1.478)
	†††	†††	††	†††	†††
Constant	0.105***	0.612***	0.283***	58.852***	57.137***
	(0.008)	(0.012)	(0.011)	(1.220)	(1.086)
Observations	8326	8326	8326	7423	8326
R^2	0.017	0.005	0.002	0.004	0.004

Notes. In this table, we present the effects of being randomly assigned to Groups 2-5 on the share of participant that chose *Don't Buy*, Mediocre, and *Best Buy* products, as well as participants' WTP for the product they chose. The regressions are conducted using LPM and OLS. The omitted group in each regression is Group 1 (the control group with no fake reviews). There are some missing values in Column (4) as we are missing demographic data for some participants in the main experiment. Standard errors in parentheses (* p<0.1 ** p<0.05 *** p<0.01). The significance levels of our coefficients do not change if we adjust for multiple hypothesis testing († p<0.1 †† p<0.05 ††† p<0.01 using the multiple hypothesis testing correction introduced by List et al. (2021)).

star ratings matter, even when applied to low-quality products.

Tables A14-A16 in the Appendix present a heterogeneity analysis that shows that fake reviews matter for all product categories (recall that participants could choose dash-cams, headphones, or cordless vacuum clears), suggesting that the effects that we record are generalizable. We do, however, find evidence suggesting that fake reviews have a larger impact on demand and welfare in the high-cost product category (cordless vacuum cleaners).[27]

Given that the average star rating of the low-quality product increased by 1.46 stars (a 46% increase relative to Group 1), and that demand increased for this product by 58%, we estimate that the elasticity of demand for the low-quality product with regard to stars is 1.21

[27] This result is consistent with Lewis and Zervas (2019) who find that ratings have a higher impact on demand for more expensive hotel rooms than for cheap hotel rooms. However, as only 1910 individuals picked the dash-cam product category, we do not have sufficient power to detect small differences in this sub-group. Nonetheless, our coefficient estimates are in the expected direction, and we find, for example, that those assigned to Group 5 are 8 percentage points more likely to choose the *Don't Buy* product.

(*i.e.*, 58/46). Moreover, assuming that star ratings have linear impacts on consumption, we can conclude that increasing the star rating by one star increases demand for the low-quality product by 38%. This, in turn, translates into a welfare loss of $.03 per dollar spent on the platform for every one-star increase in the rating of the low-quality product.

While we find that inflated star ratings influence consumer behavior, we find that fake written reviews have an even greater impact. Group 3 (fake written reviews) is 7 percentage points more likely to choose the *Don't Buy* product than Group 2 (inflated star ratings). We also find that highly suspicious fake reviews influence consumer decision-making, with Group 4 being 5 percentage points more likely to choose the *Don't Buy* product than Group 2 (inflated star ratings). Furthermore, unlike the inflated star ratings, we find that the written reviews make participants significantly less likely to purchase a *Best Buy* product, suggesting that written fake reviews cause consumer harm.

Strikingly, we find that fake written reviews and inflated star ratings have a very large combined effect on consumer decision-making. For example, Group 3 is 12.6 percentage points more likely to purchase a *Don't Buy* product than Group 1 (those who were shown no fake reviews). This corresponds to a 120% increase in the share of participants that purchase a low-quality product. Group 3 is also 7.4 percentage points less likely to purchase a mediocre product, and 5.2 percentage points less likely to purchase a *Best Buy* product. Taken together, this translates into around a £7 loss in welfare per participant in that group.

Finally, Table 2 shows that the platform endorsement matters. More specifically, 14% of Group 5 purchased the *Don't Buy* product, which is 2 percentage points more than Group 3. However, those in the platform endorsement group were also more likely than Group 3 to purchase a *Best Buy* product. Consequently, adding the endorsement only generated a modest loss in consumer welfare (around £1).

To understand the mechanisms underlying these results, we elicited people's confidence in their purchase, the ease of making an informed choice, what information they based their choice on, and whether they read the customer reviews. We find that those who were allocated to the sloppy fake review group (Group 4) were 3.8 percentage points less likely to say that they read the reviews ($p = 0.007$) than those in the control group. While there are no other statistically significant relationships between treatment assignment and these outcomes, we do find evidence suggesting that those allocated to the inflated stars group (Group 2) were 3.1 percentage points ($p = 0.069$) less likely to say that they based their choice on the product description and were 2.7 percentage points less likely to say that they read the written reviews

($p = 0.055$).[28] See Tables A18 and A19 for more information.

3.2 Not all consumers are equally gullible

Next, we examine whether fake reviews have a lower impact on consumers who think that these reviews are more likely to be fake. We do so by examining whether the effects of fake reviews are smaller for those who say that they do not trust reviews on Amazon.

Table 3 shows that those who trust and those who do not trust reviews in Group 1 (the condition with no fake reviews) are roughly as likely to purchase the *Don't Buy* product. However, the table also reveals that fake reviews have a dramatically larger effect for those who trust customer reviews. More specifically, those who do not trust reviews are not significantly influenced by inflated star ratings in this context. Furthermore, they are less susceptible to fake written reviews than their more trusting peers. Finally, those who do not trust reviews are not significantly influenced by the platform endorsement, unlike those who trust reviews.[29]

Next, we examine the extent to which experienced Amazon users are likely to fall for fake reviews. The motivation for conducting this analysis is that past work in economics (see, *e.g.*, List (2003)) shows that consumers often become more sophisticated the more they engage in a market. As Table A12 shows, we find that fake reviews have a larger negative impact on those who use the Amazon.co.uk platform more than once per month relative to those who use it less often. This result is perhaps surprising, and suggests that more experience does not translate into greater 'sophistication' when it comes to fake reviews. This result may, in part, be explained by the fact that those who use Amazon more often spent significantly less time when completing the shopping task (30 seconds, to be exact) than those who use it less frequently. Moreover, those who use Amazon more frequently were also more likely to say that they based their choice on star ratings, the number of reviews, and the content of reviews—factors which were manipulated in the fake review treatment groups.

Finally, we examine whether fake reviews have different effects for consumers who do not have a college degree, who have lower incomes, and those who use Amazon more frequently. This analysis can help inform the targeting of interventions aimed at addressing fake reviews. Tables A10 and A11 in the Appendix present the results of these heterogeneity analyses. For example, consumers without a college degree are more likely to be influenced

[28]It is also possible that we failed to detect statistically significant effects on these variables due to a lack of statistical power. More specifically, if those who did not purchase a *Don't Buy* product ended up not viewing the *Don't Buy* product page, they would not have been exposed to any fake reviews (with the exception of the inflated star ratings that appeared on the search page).

[29]These results hold even when controlling for a range of demographic factors. See Table A13.

Table 3: The effects of fake reviews for those who (dis)trust reviews

	(1) Chose *Don't Buy*	(2) Chose Mediocre	(3) Chose *Best Buy*	(4) WTP (predicted)	(5) WTP (average)
Does not trust reviews	-0.0267*	0.0312	-0.00447	1.330	3.172*
	(0.0147)	(0.0213)	(0.0191)	(2.048)	(1.876)
Group 2	0.0252*	-0.0326*	0.00741	-0.947	-0.741
(G1 + fake stars)	(0.0129)	(0.0170)	(0.0152)	(1.670)	(1.482)
Group 2 × no trust	-0.0418*	0.0426	-0.000810	-1.052	-1.706
	(0.0251)	(0.0369)	(0.0334)	(3.611)	(3.279)
Group 3	0.101***	-0.0691***	-0.0320**	-5.260***	-4.878***
(G2 + fake reviews)	(0.0142)	(0.0172)	(0.0148)	(1.643)	(1.447)
Group 3 × no trust	-0.0678**	0.0595	0.00827	-0.217	0.630
	(0.0278)	(0.0370)	(0.0324)	(3.507)	(3.221)
Group 4	0.0696***	-0.0577***	-0.0119	-3.364**	-3.459**
(G2 + sloppy reviews)	(0.0137)	(0.0171)	(0.0150)	(1.670)	(1.462)
Group 4 × no trust	0.00235	0.0461	-0.0484	-3.298	-4.685
	(0.0285)	(0.0366)	(0.0312)	(3.420)	(3.077)
Group 5	0.114***	-0.0994***	-0.0144	-6.078***	-6.071***
(G3 + endorsement)	(0.0141)	(0.0169)	(0.0148)	(1.601)	(1.414)
Group 5 × no trust	-0.0724**	0.0666*	0.00580	2.888	4.166
	(0.0284)	(0.0376)	(0.0332)	(3.492)	(3.246)
Constant	0.153***	0.584***	0.263***	56.38***	54.39***
	(0.00715)	(0.00980)	(0.00875)	(0.974)	(0.860)
Observations	9,643	9,643	9,643	8,677	9,643
R^2	0.015	0.007	0.001	0.003	0.004

Notes. In this table, we present the interaction between trusting reviews on Amazon and fake reviews. The LPM and OLS regressions are conducted by interacting the treatment terms with a dummy variable indicating whether participants trust reviews on Amazon. We do not control for any other factors in these regressions (see the Appendix for the results when controlling for other variables). The omitted group in these regressions is Group 1 (the control condition, which was not shown any fake reviews). * p<0.1 ** p<0.05 *** p<0.01.

by the platform endorsement (Group 4). We do not find any consistent associations between income and the effects of fake reviews.

3.3 Educational interventions can reduce the adverse effects of fake reviews

Table 4: The marginal effects of the education intervention

	(1) Chose *Don't Buy*	(2) Chose Mediocre	(3) Chose *Best Buy*	(4) WTP (predicted)	(5) WTP (average)
Group 6	-0.058***	0.065***	-0.007	3.012*	3.064**
(G5 + education)	(0.014)	(0.017)	(0.015)	(1.634)	(1.028)
Constant	0.247***	0.506***	0.246***	51.687***	49.886***
	(0.010)	(0.012)	(0.010)	(1.137)	(1.086)
Observations	3355	3355	3355	3008	3355
R^2	0.005	0.004	0.000	0.001	0.001

Notes. In this table, we present the effects of being assigned to Groups 6 (the educational intervention) on the share of participant that chose *Don't Buy*, mediocre, and *Best Buy* products, and on their willingness to pay for the product they chose. The regressions are conducted using Linear Probability Models. The comparison group in each regression is Group 5 (the endorsement group). Standard errors in parenthesis. * p<0.1 ** p<0.05 *** p<0.01. The significance levels of our coefficients do not change if we adjust for multiple hypothesis testing using List et al. (2021).

In this sub-section we evaluate the effects of the educational intervention. We do so by comparing consumer behavior between those in the educational intervention group (Group 6) and those in the platform endorsement group (Group 5), as those in Group 5 and 6 were shown the same information (with the exception of the educational intervention). Table 4 shows that the intervention makes participants 6 percentage points less likely to purchase a *Don't Buy* product, and it makes them around 7 percentage points more likely to purchase a mediocre product. The intervention does not, however, make participants significantly more likely to purchase a *Best Buy* product. Nonetheless, we find that the intervention increases average consumer welfare by £3, as the intervention increases the predicted value of the product that participants chose by that magnitude. This is roughly half of the welfare loss incurred by the consumers in the the group with inflated star ratings, fake text reviews, and a platform endorsement of the *Don't Buy* product (Group 5), suggesting that educational interventions can meaningfully address the negative effects of fake reviews.

We also examine the heterogeneous treatment effects of the educational intervention in Table A17. We show that the educational intervention works equally well for college-educated and non-college educated, and those with low- and above-median incomes. It also has statistically indistinguishable effects for those who trust and do not trust reviews on Amazon and for those who use Amazon more or less frequently.

Finally, we examine *why* the educational intervention changed participant behavior. To do so, we look at whether those in the intervention condition adopted different strategies, held different beliefs, or took different amounts of time relative to those in Group 5. Our analysis shows that those in their intervention condition (Group 6) were 4 percentage points less likely to base their product choice on star ratings ($p = 0.019$) and that they were 4 percentage points more likely to base their choice on the brand of the product ($p = 0.017$) than those in Group 5 (see Tables A20 and A21 for more information). These effects are consistent with the recommendations presented in the intervention, which among other things emphasized that consumers should "Inspect the comments (don't rely on star ratings alone)".

4 External validity

To help address the external validity of our results, we use the Selection, Attrition, Naturalness, and Scaling (SANS) framework of List (2020). First, to address *Selection*, we sampled a large nationally representative sample of UK adults. Moreover, our sample includes many individuals who use online platforms, and who may be influenced by fake reviews in their everyday lives. One common issue related to recruiting participants from online survey panels is that they exclude those who do not use the internet. However, that is not a problem in our case, as we are only interested in individuals who use the internet. Second, we have no *Attrition* in our experiment.

Third, we believe that we have achieved a high degree of *Naturalness* in the choice task, setting, and time-frame for our experiment. The choice task closely mimics how consumers interact with the Amazon platform. People searched for a product category, could visit one or more product pages, and then selected the product they wished to purchase. Further, participants were incentivized to choose the product that they liked the most. One key difference to the 'real world' is that participants were spending the experimenters' money rather than their own (as we were paying for the product).

Participants in our experiment had to buy a product (*i.e.*, they could not choose to buy nothing and to pocket the cost of the product). Moreover, they had to buy the product on our platform (they could not shop around for better deals with other online retailers). This set-up is, of course, not reflective of *all* online shopping journeys. That said, many individuals do begin a shopping journey determined to purchase a product, and many Amazon users do not look for products on other platforms before making a purchase. In fact, a recent survey of online shopping behavior found 77% of respondents go directly to Amazon when they are ready to buy a product (without checking other websites) (Forbes, 2019). We thus believe that the choice task is representative of important types of real-world shopping behavior.

The setting of the task was also realistic. People were on their computer, laptop, or on their smartphone when completing the task in our experiment, just as they would have been when using Amazon in their everyday lives. Moreover, the feel and design of the platform closely resembled the Amazon platform, and we presented the same type of information as exists on Amazon's product and search pages. Similar to actual shopping on online marketplaces, participants could search for products online when using our platform (if they, for example, wanted more information about the products)—and around 18% did. The time-frame of choice was also identical to the real-world decision of buying on a platform (participants could spend as long as they liked, just as when shopping in other settings).

With regards to the issue of *Scaling*, it is clearly the case that fake reviews 'scale' (as they exist across a range of platforms), and we find that they affect purchasing decisions for different population segments and product categories. We also believe that our educational intervention could be taken to scale, as it is simple to implement and does not require any targeting (either at an individual or product level). The main factor that may inhibit the scaling of the educational intervention is that platforms would have to dedicate space to it in a prominent place on their websites (and this is expensive real estate for many online marketplaces).

Overall, we believe that our results likely have a high degree of external validity based on these criteria. While we are assuming that the treatments work in the same way for those who did and did not sign up for the study, we do not see why treatment effects would differ for these groups. Moreover, there is no attrition, and the choice task closely resembles how people shop online in their everyday lives.

5 Conclusion

Online commerce provides large benefits to consumers (Dolfen et al., 2019). One of the key innovations associated with online commerce is the introduction of reviews that allow consumers to become informed about different aspects of a product at a relatively low cost. Genuine online customer reviews can help consumers navigate this rich—and sometimes challenging—marketplace. Research has shown, however, that many reviews are fake, and that these fake reviews could be harmful to both consumers and honest producers.

Recognizing the potential economic importance of this issue, several government agencies and companies have announced their intent to address fake reviews. The U.S. Federal Trade Commission proposed possible rules that would allow it to impose a fine of up to $50,000 for each fake review (Fowler, 2023). In 2022, Meta announced its intent to fight fake reviews and filed a lawsuit against a company because of fake reviews (Hutchinson, 2022). More recently, Amazon, Expedia, Tripadvisor, Glassdoor, and Trustpilot formed a global "Coalition for Trusted Reviews." Among other things, the companies hope to develop common standards for addressing the problem, share best practices, and share information on how fake review producers operate (Amazon, 2023).

While we are encouraged that key economic actors are highlighting the importance of fake reviews, researchers still know very little about the extent to which such reviews affect consumer behavior and welfare and how to minimize their adverse effects. This paper attempts to address this research gap. To the best of our knowledge, this study provides the first experimental estimates of the effects of fake reviews on individual consumption choices and welfare.

The analysis of the experiment produced four main findings. First, fake reviews make consumers more likely to choose inferior products. We found that these reviews have similar effects for those who choose the dash-cams, cordless vacuum cleaners, and headphone product categories, suggesting that our results may generalize. Second, we estimate that welfare losses from such reviews may be important—on the order of $.12 for each dollar spent on our platform. Taken together with the fact that fake reviews are typically applied to lower-quality products in the 'real world' (see, *e.g.*, (He et al., 2021)) these findings suggest that such reviews can generate substantial consumer harm. Furthermore, the results imply that the impacts of fake reviews should be taken into account when assessing the overall welfare impacts of online platforms with rating systems (see, *e.g.*,, Wu et al. (2015), Lewis and Zervas (2019), and Reimers and Waldfogel (2021) for examples of studies that assess the welfare impacts of reviews in online platforms).

Third, the effect of fake reviews is smaller for skeptical consumers. Moreover, those who use Amazon more frequently are more likely to changes their purchasing decision in response to fake reviews. Fourth, educational interventions have the potential to dramatically reduce the effects that fake reviews have on consumption choices. More specifically, the intervention reduces the negative welfare impact of fake reviews by around 44% in this context.

We suggest three areas of future research that will help us further understand the welfare consequences of fake reviews. The first is to consider doing natural field experiments on how fake reviews affect consumer behavior on platforms that sell directly to consumers. Natural field experiments should be conducted because fake reviews may have effects on actual platforms that we were unable to capture in this experiment. For example, trust is important for the functioning of online markets (Einav et al., 2016; Tadelis, 2016), but fake reviews could undermine consumer trust in reviews in general. This loss in trust could induce consumers to rely on other sources of information that may be less informative and more costly to access than honest reviews, and they might choose other platforms. Consumers who fall for fake reviews may also post negative reviews if they end up disliking the product they purchased, which may dampen the long-run effects of fake reviews.[30] That said, we believe that our field experiment is the first step in the 'low cost wind tunnel' of generating replicable evidence (List, 2022).

The second suggestion is to understand the relationship between the extent of fake reviews and consumer demand and welfare. For example, it would be useful to understand how many fake reviews are required to influence consumer demand.[31] It would also be useful to disentangle the marginal impact of fake review text on demand from the marginal impact of fake star ratings on demand. Further, it would be interesting to record *how* consumers interact with platforms in the presence of fake reviews, as this would help us better understand the mechanisms underlying the observed impact on demand.[32]

The third suggestion is to measure the economic impact on both consumers and producers of various interventions aimed at curbing fake reviews. While we found that our educational intervention reduces the impact of fake reviews, more interventions should be tested, and their impact on economic welfare should be assessed. This stream of work would build on recent studies focused on estimating the welfare effects of nudges and other similar interventions (Allcott and Kessler, 2019; Butera et al., 2022). Given that the cost of supplying "human-like" fake reviews has dropped through using large language models (such as ChatGPT) (Salminen et al., 2022; Crothers et al., 2023; Sadasivan et al., 2023), we think it is ever more important that the welfare considerations of fake reviews are examined.

[30]Note, however, that this dynamic depends on 1) consumers' ability to assess product quality, and on 2) their willingness to go through the hassle of posting a review online.

[31]A recent paper by Fradkin and Holtz (2023) found that the causal impact of one additional review (from zero to one) on Airbnb led to no overall changes in demand.

[32]Other questions that would be interesting to explore include whether it is better to have platforms with: 1) no reviews at all or 2) fake and real reviews, and under which circumstances either of the two options dominate. This type of question could be explored using the design presented in this study, if one was to add a group that was shown no reviews at all.

References

Acemoglu, D., Makhdoumi, A., Malekian, A., and Ozdaglar, A. (2019). Learning from reviews: The selection effect and the speed of learning. *URL https://economics. mit. edu/files/17178, unpublished*.

Adomavicius, G. and Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering*, 17(6):734–749.

Aköz, K. K., Arbatli, C. E., and Celik, L. (2020). Manipulation through biased product reviews. *Journal of Industrial Economics*, 68(4):591–639.

Allcott, H. and Kessler, J. B. (2019). The welfare effects of nudges: A case study of energy use social comparisons. *American Economic Journal: Applied Economics*, 11(1):236–76.

Amazon (2023). Amazon, booking.com, expedia group, glassdoor, tripadvisor, and trustpilot launch first global coalition for trusted reviews. Technical report.

Ananthakrishnan, U. M., Li, B., and Smith, M. D. (2020). A tangled web: Should online review portals display fraudulent reviews? *Information Systems Research*, 31(3):950–971.

Anderson, E. T. and Simester, D. I. (2014). Reviews without a purchase: Low ratings, loyal customers, and deception. *Journal of Marketing Research*, 51(3):249–269.

Anderson, M. and Magruder, J. (2012). Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. *Economic Journal*, 122(563):957–989.

Andı, S. and Akesson, J. (2020). Nudging away false news: Evidence from a social norms experiment. *Digital Journalism*, 9(1):106–125.

Aral, S. (2014). The problem with online ratings. *MIT Sloan Management Review*, 55:47–52.

Augenblick, N. and Rabin, M. (2019). Belief movement, uncertainty reduction, and rational updating.

Bajari, P. and Hortaçsu, A. (2004). Economic insights from internet auctions. *Journal of Economic Literature*, 42(2):457–486.

Banerjee, S., Dellarocas, C., and Zervas, G. (2021). Interacting user-generated content technologies: How questions and answers affect consumer reviews. *Journal of Marketing Research*, 58(4):742–761.

Barach, M. A., Golden, J. M., and Horton, J. J. (2020). Steering in online markets: the role of platform incentives and credibility. *Management Science*, 66(9):4047–4070.

Becker, G. M., DeGroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral science*, 9(3):226–232.

Belleflamme, P. and Peitz, M. (2018). Inside the engine room of digital platforms: Reviews, ratings, and recommendations.

Benson, A., Sojourner, A., and Umyarov, A. (2020). Can reputation discipline the gig economy? experimental evidence from an online labor market. *Management Science*, 66(5):1802–1825.

Bonatti, A. and Cisternas, G. (2020). Consumer scores and price discrimination. *Review of Economic Studies*, 87(2):750–791.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2015). Salience theory of judicial decisions. *Journal of Legal Studies*, 44(S1):S7–S33.

Botelho, T. L. and Gertsberg, M. (2021). The disciplining effect of status: Evaluator status awards and observed gender bias in evaluations. *Management Science*.

Butera, L., Metcalfe, R., Morrison, W., and Taubinsky, D. (2022). Measuring the welfare effects of shame and pride. *American Economic Review*, 112(1):122–68.

Cabral, L. and Hortacsu, A. (2010). The dynamics of seller reputation: Evidence from ebay. *Journal of Industrial Economics*, 58(1):54–78.

Chan, M.-p. S., Jones, C. R., Hall Jamieson, K., and Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological science*, 28(11):1531–1546.

Chevalier, J. A., Dover, Y., and Mayzlin, D. (2018). Channels of impact: User reviews when quality is dynamic and managers respond. *Marketing Science*, 37(5):688–709.

Chevalier, J. A. and Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, 43(3):345–354.

Crothers, E., Japkowicz, N., and Viktor, H. L. (2023). Machine-generated text: A comprehensive survey of threat models and detection methods. *IEEE Access*.

Cui, R., Li, J., and Zhang, D. J. (2020). Reducing discrimination with reviews in the sharing economy: Evidence from field experiments on airbnb. *Management Science*, 66(3):1071–1094.

Dellarocas, C. (2006). Strategic manipulation of internet opinion forums: Implications for consumers and firms. *Management Science*, 52(10):1577–1593.

Dinerstein, M., Einav, L., Levin, J., and Sundaresan, N. (2018). Consumer price search and platform design in internet commerce. *American Economic Review*, 108(7):1820–59.

Dolfen, P., Einav, L., Klenow, P. J., Klopack, B., Levin, J. D., Levin, L., and Best, W. (2019). Assessing the gains from e-commerce. Technical report, National Bureau of Economic Research.

Dwoskin, E. and Timberg, C. (2018). How merchants use facebook to flood amazon with fake reviews. Technical report.

Einav, L., Farronato, C., and Levin, J. (2016). Peer-to-peer markets. *Annual Review of Economics*, 8:615–635.

Farronato, C. and Zervas, G. (2019). Consumer reviews and regulation: evidence from nyc restaurants. *Working Paper*.

Filippas, A., Horton, J. J., and Golden, J. (2018). Reputation inflation. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 483–484.

Forbes (2019). *89 percent Of Consumers Are More Likely To Buy Products From Amazon Than Other E-Commerce Sites: Study*. (accessed December 20, 2021).

Fowler, G. A. (2023). Those 10,000 5-star reviews are fake. now they'll also be illegal. Technical report.

Fradkin, A. and Holtz, D. (2023). Do incentives to review help the market? evidence from a field experiment on airbnb. *Marketing Science*.

FTC (2021). *FTC Puts Hundreds of Businesses on Notice about Fake Reviews and Other Misleading Endorsements*. (accessed December 21, 2021).

Gao, G., Greenwood, B. N., Agarwal, R., and McCullough, J. S. (2015). Vocal minority and silent majority. *MIS quarterly*, 39(3):565–590.

Gardete, P. and Hunter, M. (2020). Guiding consumers through lemons and peaches: An analysis of the effects of search design activities.

Glazer, J., Herrera, H., and Perry, M. (2020). Fake reviews. *Economic Journal*, 131(636):1772–1787.

Goldfarb, A. and Tucker, C. (2019). Digital economics. *Journal of Economic Literature*, 57(1):3–43.

Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., and Sircar, N. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the united states and india. *Proceedings of the National Academy of Sciences*, 117(27):15536–15545.

Hastings, J. S., Madrian, B. C., and Skimmyhorn, W. L. (2013). Financial literacy, financial education, and economic outcomes. *Annu. Rev. Econ.*, 5(1):347–373.

He, S., Hollenbeck, B., Overgoor, G., Proserpio, D., and Tosyali, A. (2022). Detecting fake-review buyers using network structure: Direct evidence from amazon. *Proceedings of the National Academy of Sciences*, 119(47):e2211932119.

He, S., Hollenbeck, B., and Proserpio, D. (2021). The market for fake reviews.

He, S., Hollenbeck, B., and Proserpio, D. (2023). Leveraging social media to buy fake reviews. *Communications of the ACM*, 66(10):98–105.

Hodgson, C. and Lewis, G. (2020). You can lead a horse to water: Spatial learning and path dependence in consumer search.

Honka, E., Hortaçsu, A., and Wildenbeest, M. (2019). Empirical search and consideration sets. In *Handbook of the Economics of Marketing*, volume 1, pages 193–257. Elsevier.

Horton, J. J. (2017). The effects of algorithmic labor market recommendations: Evidence from a field experiment. *Journal of Labor Economics*, 35(2):345–385.

Hu, N., Bose, I., Gao, Y., and Liu, L. (2011a). Manipulation in digital word-of-mouth: A reality check for book reviews. *Decision Support Systems*, 50(3):627–635.

Hu, N., Bose, I., Koh, N. S., and Liu, L. (2012). Manipulation of online reviews: An analysis of ratings, readability, and sentiments. *Decision Support Systems*, 52(3):674–684.

Hu, N., Liu, L., and Sambamurthy, V. (2011b). Fraud detection in online consumer reviews. *Decision Support Systems*, 50(3):614–626.

Hutchinson, A. (2022). Meta files new lawsuit over the sale of fake customer reviews on facebook. Technical report.

Jin, G. Z. and Kato, A. (2006). Price, quality, and reputation: Evidence from an online field experiment. *RAND Journal of Economics*, 37(4):983–1005.

Jin, G. Z. and Leslie, P. (2003). The effect of information on product quality: Evidence from restaurant hygiene grade cards. *Quarterly Journal of Economics*, 118(2):409–451.

KC, S. and Mukherjee, A. (2016). On the temporal dynamics of opinion spamming: Case studies on yelp. Technical report, WWW '16: Proceedings of the 25th International Conference on World Wide Web.

Klein, T. J., Lambertz, C., and Stahl, K. O. (2016). Market transparency, adverse selection, and moral hazard. *Journal of Political Economy*, 124(6):1677–1713.

Knapp, B. et al. (2021). *Fake Reviews and Naive Consumers*. Universität Wien, Department of Economics, University of Vienna.

Kumar, S. and Shah, N. (2018). False information on web and social media: A survey.

Lam, H. T. (2021). Platform search design and market power.

Lappas, T., Sabnis, G., and Valkanas, G. (2016). The impact of fake reviews on online visibility: A vulnerability assessment of the hotel industry. *Information Systems Research*, 27(4):940–961.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., et al. (2018). The science of fake news. *Science*, 359(6380):1094–1096.

Lee, I. (2020). Amazon fake reviews reach holiday season levels during pandemic. *Bloomberg*.

Lewis, G. and Zervas, G. (2019). The supply and demand effects of review platforms. Working paper.

List, J. A. (2003). Does market experience eliminate market anomalies? *Quarterly Journal of Economics*, 118(1):41–71.

List, J. A. (2006). The behavioralist meets the market: Measuring social preferences and reputation effects in actual transactions. *Journal of Political Economy*, 114(1):1–37.

List, J. A. (2020). Non est disputandum de generalizability? a glimpse into the external validity trial. Technical report, National Bureau of Economic Research.

List, J. A. (2022). *The voltage effect: How to make good ideas great and great ideas scale*. Currency.

List, J. A., Shaikh, A. M., Vayalinkal, A., et al. (2021). Multiple testing with covariate adjustment in experimental economics. Technical report, The Field Experiments Website.

Luca, M. (2016). Reviews, reputation, and revenue: The case of yelp. com. *Com (March 15, 2016). Harvard Business School NOM Unit Working Paper*, (12-016).

Luca, M. and Zervas, G. (2016). Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science*, 62(12):3412–3427.

Mayzlin, D. (2006). Promotional chat on the internet. *Marketing Science*, 25(2):155–163.

Mayzlin, D., Dover, Y., and Chevalier, J. (2014). Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review*, 104(8):2421–55.

Nosko, C. and Tadelis, S. (2015). The limits of reputation in platform markets: An empirical analysis and field experiment. Working Paper 20830, National Bureau of Economic Research.

Ott, M., Choi, Y., Cardie, C., and Hancock, J. T. (2011). Finding deceptive opinion spam by any stretch of the imagination. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 309–319, Portland, Oregon, USA. Association for Computational Linguistics.

Pallais, A. (2014). Inefficient hiring in entry-level labor markets. *American Economic Review*, 104(11):3565–99.

Panniello, U., Gorgoglione, M., Hill, S., and Hosanagar, K. (2014). Incorporating profit margins into recommender systems: A randomized field experiment of purchasing behavior and consumer trust.

Pennycook, G., Bear, A., Collins, E. T., and Rand, D. G. (2020). The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management Science*, 66(11):4944–4957.

Piccolo, S., Tedeschi, P., and Ursino, G. (2018). Deceptive advertising with rational buyers. *Management Science*, 64(3):1291–1310.

Reimers, I. and Waldfogel, J. (2021). Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings. *American Economic Review*, 111(6):1944–71.

Resnick, P., Zeckhauser, R., Swanson, J., and Lockwood, K. (2006). The value of reputation on ebay: A controlled experiment. *Experimental Economics*, 9(2):79–101.

Sadasivan, V. S., Kumar, A., Balasubramanian, S., Wang, W., and Feizi, S. (2023). Can ai-generated text be reliably detected? *arXiv preprint arXiv:2303.11156*.

Sahni, N. S. and Nair, H. S. (2020). Does advertising serve as a signal? evidence from a field experiment in mobile search. *Review of Economic Studies*, 87(3):1529–1564.

Salminen, J., Kandpal, C., Kamel, A. M., Jung, S.-g., and Jansen, B. J. (2022). Creating and detecting fake reviews of online products. *Journal of Retailing and Consumer Services*, 64:102771.

Smith, A., Anderson, M., and Page, D. (2016). Online shopping and e-commerce. *Pew Research Center*, (202.419.4372).

Stanton, C. T. and Thomas, C. (2016). Landing the first job: The value of intermediaries in online hiring. *Review of Economic Studies*, 83(2):810–854.

Tadelis, S. (2016). Reputation and feedback systems in online platform markets. *Annual Review of Economics*, 8:321–340.

Taubinsky, D. and Rees-Jones, A. (2018). Attention variation and welfare: theory and evidence from a tax salience experiment. *Review of Economic Studies*, 85(4):2462–2496.

Ursu, R. M. (2018). The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science*, 37(4):530–552.

USCB (2021). Quarterly retail e-commerce sales.

Which? (2020a). *Amazon 'betraying trust' of millions of consumers with flawed Amazon's Choice endorsement*. (accessed May 18, 2021).

Which? (2020b). The real impact of fake reviews: a behavioural experiment on how fake reviews influence consumer choices. Research paper, Which? and the Behaviouralist.

Which? (2020c). *Which? Online Fake Reviews*. (accessed June 30, 2020).

Which? (2020d). *Which? symbols, logos and ratings*. (accessed September 8, 2020).

Wu, C., Che, H., Chan, T. Y., and Lu, X. (2015). The economic value of online reviews. *Marketing Science*, 34(5):739–754.

Yao, Y., Viswanath, B., Cryan, J., Zheng, H., and Zhao, B. Y. (2017). Automated crowdturfing attacks and defenses in online review systems. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, CCS '17, page 1143–1158, New York, NY, USA. Association for Computing Machinery.

Yasui, Y. (2021). Controlling fake reviews. Working paper, University of California, Los Angeles (UCLA).

Zhang, D., Zhou, L., Kehoe, J. L., and Kilic, I. Y. (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.

Zhuang, M., Cui, G., and Peng, L. (2018). Manufactured opinions: The effect of manipulating online product reviews. *Journal of Business Research*, 87:24–35.

Zinman, J. and Zitzewitz, E. (2016). Wintertime for deceptive advertising? *American Economic Journal: Applied Economics*, 8(1):177–92.

Appendices

1. We begin by presenting a simple model for how fake reviews influence consumers' decisions, along with proofs for the model in Appendix B and a parametric example in Appendix and C.
2. Appendix D includes a theory formalizing the welfare argument presented in Section 2.4.1.
3. In Appendix E, we present the information that participants were shown when navigating through the experiment.
4. Appendix F includes descriptive statistics and a balance table.
5. Appendix G includes the heterogeneity analyses described in Section 3.2.
6. Finally, Appendix H lists the survey questions used in the experiment and Appendix I lists the questions used in the companion survey.

A A model of how fake reviews influence consumers' assessment of product quality

In this section, we provide a simple model of the effect of fake reviews on the beliefs of a Bayesian consumer. We show how fake reviews could positively affect a consumer's assessment of product quality, how greater uncertainty about product quality could increase the impact of a fake review, and how skepticism regarding the genuineness of a review will reduce its impact on consumer behavior. We abstract from many of the details of actual online reviews for the sake of clarity. For example, the consumer in our model simply views a binary review X whose content may be either positive ($X = 1$) or negative ($X = 0$). Nonetheless, many of the key results of our model can be generalized to more complex environments. Moreover, as we are primarily interested in explaining how consumers may be affected by fake reviews when making a one-off purchase, we do not provide a full general equilibrium analysis, as in Glazer et al. (2020), nor a dynamic analysis of learning, as in Acemoglu et al. (2019).[33]

We consider a decision-maker (DM) who is initially uncertain about the quality of a product. We will measure the product's quality using the proportion $\theta \in [0,1]$ of genuine reviewers who would give the product a good rating. In other words, θ is the probability that a (genuine) review would be positive if such a review were randomly sampled from the population of genuine reviewers.

We formalize the DM's initial uncertainty about θ using a prior $p\colon [0,1] \to \mathbb{R}^+$. Importantly, we make almost no assumptions about the shape of the DM's prior, requiring only

[33] In a dynamic environment, it is possible that consumers learn over time, eventually becoming better at distinguishing between fake and genuine reviews. However, as can be seen in our heterogeneity analysis in Section 3.2, it does not seem like consumers become better at distinguishing between fake and real reviews over time. Indeed, the opposite seems more likely.

that $p(\theta) > 0$ for all $\theta \in [0,1]$ (*i.e.*, the prior has full support). We also assume that the DM is principally interested in the expected value of θ (which provides a natural summary statistic for their beliefs about product quality).

The DM believes the reviews are generated in the following way. With probability $q \in (0,1)$, the review is genuine, in which case it is positive with probability θ. Otherwise, the review is fake, in which case it is definitely positive about the product.[34] Thus, q is the DM's subjective probability that the review that they see is genuine. Furthermore, fixing any $\theta \in [0,1]$, the DM believes that there is a (prior) probability $q\theta + 1 - q$ that the review they will see is positive.

We begin by examining whether fake reviews 'work'. More formally, given an arbitrary prior p and subjective probability q, we want to know whether viewing a positive review (which the DM knows may be fake) increases their expected value of θ. The next proposition verifies that this is indeed the case.

Proposition 1. $\mathbb{E}[\theta|X = 1] > \mathbb{E}[\theta]$.

Proposition 1 is intuitive. When the DM receives a positive review, there are two possible explanations: (1) the review is fake; or (2) the review is genuine and happens to be positive. In the first case, the DM has learned nothing: fake reviews are always positive, so the fact that the fake review is positive cannot provide novel information. In the second case, however, the DM has learned something: the fact that a genuine review is positive suggests a high θ. Thus, as long as the DM thinks there is some chance that a review is genuine, no matter how small, a fake review will increase her assessment of the product's quality (which presumably also increases the likelihood that she purchases the product).

While we have proven Proposition 1 in the simplest of contexts, the basic idea extends to more complex environments. For example, suppose that the DM observes not one but $n \geq 1$ positive (but fake) reviews. Assuming that the reviews are 'exchangeable' (*e.g.*, because they are i.i.d.), it is equivalent to consider the DM updating her prior n times, once for each of the reviews. But then one can apply the logic of Proposition 1 iteratively, allowing one to see that the total effect of the reviews is increasing in n.[35] For example, while one fake (and positive) review inflates the DM's expectation of θ, two fake (positive) reviews do so by even more.

We thus conclude that fake reviews 'work' even if the DM believes that they are very likely to be fake ($q \approx 0$). It is natural to think, however, that skepticism about the honesty of reviews should attenuate their effect. In other words, we might think that the lower the DM deems q, the smaller the effect that fake reviews have. The next proposition verifies that this is also the case.

Proposition 2. *The effect of a fake review* $\mathbb{E}[\theta|X = 1] - \mathbb{E}[\theta]$ *is strictly increasing in both q and $Var(\theta)$.*

Proposition 2 is also intuitive. The higher q, the more likely the DM is to think that the review is genuine. But the possibility that the review is genuine is precisely the reason why

[34] In this paper, we focus on fake reviews that are positive. Nonetheless, this type of model can be expanded to analyze the effects of fake reviews that are negative.

[35] To update in light of the kth review, one needs to interpret the 'prior' as the posterior conditional on the previous $k - 1$ reviews.

it contains useful information. Given this, it is not surprising that the effect of the review should be larger the less likely the DM is to think that it is fake.

Proposition 2 also shows that the effect of fake reviews is increasing in the variance of their prior. One way of understanding this is to view the variance as a measure of the DM's initial uncertainty (as in, *e.g.*, Augenblick and Rabin (2019)). On this interpretation, Proposition 2 tells us the effect of fake reviews is larger on consumers who are initially more uncertain about the quality of the product—these are, after all, the consumers who have the most to learn from the review. We provide proofs in Appendix B and an illustrative parametric example of our model in Appendix C.

B Proofs

Proof of Proposition 1. By Bayes theorem,

$$p(\theta|X=1) = \frac{p(X=1|\theta)p(\theta)}{p(X=1)} = \frac{(q\theta+1-q)p(\theta)}{\int_0^1 (q\theta+1-q)p(\theta)d\theta}. \tag{1}$$

Thus,

$$\begin{aligned}
\mathbb{E}[\theta|X=1] &= \int_0^1 p(\theta|X=1)\theta d\theta \\
&= \frac{\int_0^1 (q\theta+1-q)p(\theta)\theta d\theta}{\int_0^1 (q\theta+1-q)p(\theta)d\theta} \\
&= \frac{q\int_0^1 \theta^2 p(\theta)d\theta + (1-q)\int_0^1 \theta p(\theta)d\theta}{q\int_0^1 \theta p(\theta)d\theta + (1-q)\int_0^1 p(\theta)d\theta} \\
&= \frac{q\mathbb{E}[\theta^2]+(1-q)\mathbb{E}[\theta]}{q\mathbb{E}[\theta]+1-q}.
\end{aligned} \tag{2}$$

As a result,

$$\begin{aligned}
\mathbb{E}[\theta|X=1]-\mathbb{E}[\theta] &= \frac{q\mathbb{E}[\theta^2]+(1-q)\mathbb{E}[\theta]}{q\mathbb{E}[\theta]+1-q} - \mathbb{E}[\theta] \\
&= \frac{q\mathbb{E}[\theta^2]+(1-q)\mathbb{E}[\theta]-\mathbb{E}[\theta](q\mathbb{E}[\theta]+1-q)}{q\mathbb{E}[\theta]+1-q} \\
&= \frac{q\left(\mathbb{E}[\theta^2]-\mathbb{E}[\theta]^2\right)}{q\mathbb{E}[\theta]+1-q} > 0
\end{aligned} \tag{3}$$

where the inequality holds since $\mathbb{E}[\theta^2]-\mathbb{E}[\theta]^2 = \mathrm{Var}(\theta) > 0$, $q \in (0,1)$ and $\mathbb{E}[\theta] > 0$. We thus conclude that $\mathbb{E}[\theta|X=1] > \mathbb{E}[\theta]$. □

Proof of Proposition 2. From (3),

$$\mathbb{E}[\theta|X=1]-\mathbb{E}[\theta] = \frac{q\mathrm{Var}(\theta)}{1-q(1-\mathbb{E}[\theta])}. \tag{4}$$

Since the numerator is increasing in q but the denominator is decreasing in q, it follows immediately that $\mathbb{E}[\theta|X=1]-\mathbb{E}[\theta]$ is increasing in q. Since both numerator and denominator are positive, it is also strictly increasing in $\mathrm{Var}(\theta)$. □

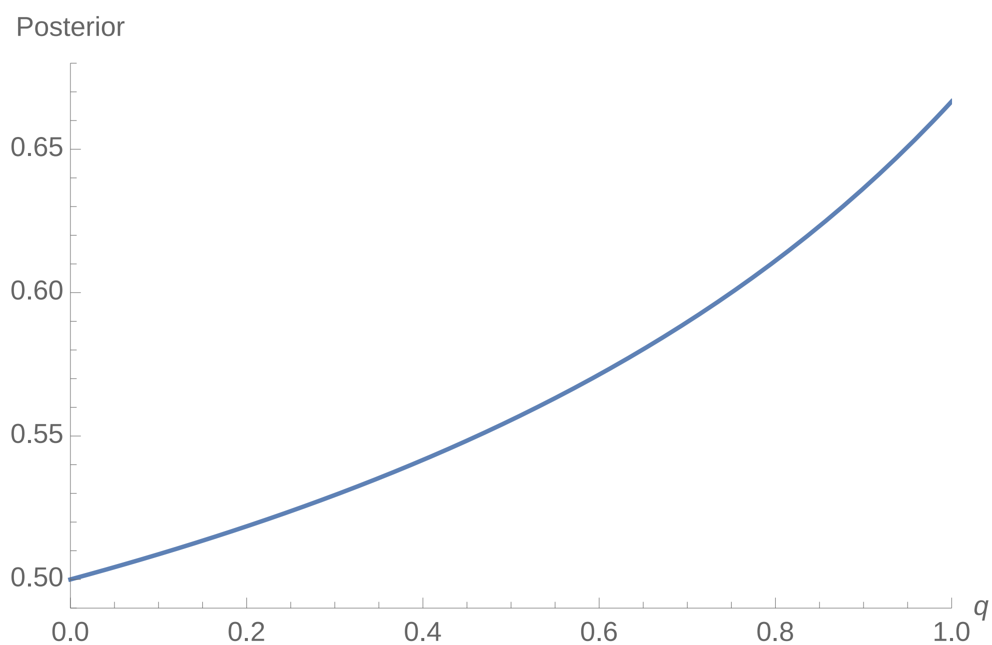
C Parameteric example

In this section we briefly illustrate our findings with a parametric example. To do so, assume that the DM's prior takes the form of a (standard) uniform distribution. By integrating, one finds that the expected value of the posterior is given by

$$\mathbb{E}[\theta|X=1] = \frac{3-q}{3(2-q)}.$$

Figure A1 illustrates. As can be seen, the expected value of the posterior always lies above the expected value of the prior (1/2), indicating that the fake review increases the DM's assessment of the product's quality (as in Proposition 1). However, the curve is also upward sloping, indicating that the effect of fake reviews is smaller the lower the subjective probability that the review is genuine (as in Proposition 2). Indeed, when $q=0$ (all reviews are thought to be fake), the prior is the same as the posterior, so the review has no effect whatsoever on the DM's beliefs. (At the other extreme of $q=1$, when all reviews are thought to be genuine, the review increases the DM's expectation from 1/2 to 2/3).

Figure A1: The effect of q on the expected value of the posterior



D A model for computing the change in welfare from the fake reviews

In this sub-section, we formalize the intuition presented above regarding how to tie participants' decisions to welfare outcomes. In other words, in this sub-section, we derive a measure of the change in welfare from fake reviews and being exposed to the educational intervention. Both fake reviews and the intervention may yield private benefits or harm by influencing their purchasing decisions and altering the utility associated with the product that they buy.

Basic model. We will assume we can measure the true utility (measured as willingness to pay) for any purchase that a consumer makes. A consumer will make a purchase if the perceived utility (which may differ from the true utility) exceeds the price. The perceived utility can be affected by a specific treatment (which may be the "intervention").

We begin with a single consumer with fixed preferences. We will then generalize this finding to treatment groups with different numbers of individuals and preferences. We will model the consumer as deciding whether to buy a single unit of a product in addition to her expenditure on x, the numeraire good that has a price of one. The consumer can choose one unit, z_i from a vector of product elements $z_i \in z = (z_1, \ldots, z_n)$. Each z_i has a single price, p. For example, an element of z can be thought of as a specific make and model of a vacuum cleaner.

We will define a vector $\delta = (\delta_1, \ldots, \delta_n)$ that will allow the consumer to choose only one unit of a product. The consumer's perceived utility over the choice of the product is given by $\hat{f}(\delta_1 z_1, \ldots, \delta_n z_n, t_i)$, where α is a taste parameter, and t_i represents the ith treatment arm in the experiment. The consumer is constrained to choose no more than one unit of a product element (*e.g.*, one type of vacuum cleaner). The consumer has an income of y, which is assumed to exceed the price of one unit of the product, p, and a lump sum tax, T, paid to the government.[36] Thus, she can afford to purchase the product if she wants to. We can now write the consumer's constrained maximization problem as follows:

$$\underset{\delta}{\text{Maximize}}\ x + \tilde{f}(\delta_1 z_1, \ldots, \delta_n z_n, t_i) \tag{5}$$

Subject to:

$$\begin{aligned} &\delta_i = 0 \text{ or } 1 \\ &0 \leq \sum \delta_i \leq 1 \\ &y = x + p \sum \delta_i + T \end{aligned}$$

The consumer maximizes the sum of utility she gets from the numeraire good, x, plus the perceived utility she gets if she decides to purchase one unit from the vector z (*i.e.*, utility is quasilinear). The first constraint says she can either purchase 0 or 1 units of each element of z. The second constraint says she can buy at most one element, say z_i of z. Together these constraints imply that she either does or does not purchase 1 unit of the product (*e.g.*, a vacuum cleaner). The budget constraint says that she spends her income entirely on the numeraire good, x, or she spends $y - p - T$ on the numeraire good and p on $z_i \in z$.

[36] The lump sum tax would cover the cost of the experiment and production costs. Note that the lump sum tax only reduces expenditure on the numeraire good.

Computing the welfare change for a typical consumer. We wish to compare the utility for the consumer under two different treatments, call them t_1 and t_2. Define U^{1*} as the utility from treatment 1; x^{1*} as the amount spent on x; and f^{1*} as the true utility from the purchase if she made one (and similarly for treatment 2). Then, let $U^{1*} = x^{1*} + f^{1*}$ be the true level of welfare with treatment 1 ; and

$U^{2*} = x^{2*} + f^{2*}$ be the true level of welfare with treatment 2. The welfare change, or change in net benefits, ΔNB, is the utility from treatment 2 less the utility from treatment 1. Formally:

$$\Delta NB = U^{2*} - U^{1*} = \left(x^{2*} + f^{2*}\right) - \left(x^{1*} + f^{1*}\right) \tag{6}$$

We consider four cases related to whether the consumer purchases a unit, z_i under treatment 1 or treatment 2.

Case 1: An element of good z is purchased under both treatments. In this case $x^{1*} = x^{2*} = y - p - T$, and the change in utility is $f^{2*} - f^{1*}$

Case 2: An element of good z is not purchased under both treatments. In this case $x^{1*} = x^{2*} = y - T$, and the change in utility is $f^{2*} - f^{1*}$. But we assume $f^{2*} = f^{1*} = 0$, so there is no change in actual utility from purchasing a product when the consumer does not purchase the product.

Case 3: An element of good z is purchased under t_2 but not under t_1. $x^{1*} = y - T; x^{2*} = y - p - T; f^{1*} = 0$ because the good is not purchased.

$$U^{2*} - U^{1*} = \left(x^{2*} + f^{2*}\right) - \left(x^{1*} + f^{1*}\right) = \left(y - p + f^{2*}\right) - (y + 0) = f^{2*} - p$$

In words, this says that the net gain in consumer surplus is given by the WTP - the price from the purchase under t_2.

Case 4: An element of good z is purchased under t_1 but not under t_2. In this case $x^{1*} = y - p - T; x^{2*} = y - T$; and we assume $f^{2*} = 0$ because the good is not purchased.

$$U^{2*} - U^{1*} = \left(x^{2*} + f^{2*}\right) - \left(x^{1*} + f^{1*}\right) = (y + 0) - \left(y - p + f^{1*}\right) = p - f^{1*}$$

This is the same as Case 3, except we are subtracting the loss in net surplus from the purchase of the product in t_1.

A general formula for estimating the welfare change associated with the experimental treatment. The preceding analysis was for one individual considering the purchase of one product under two different treatments. We will now extend this idea to a finite number of product categories and people in each treatment, who are not necessarily identical.[37]

We wish to compute the average change in welfare between treatments, labeled "1" and "2" below. We introduce the following notation:

Let i represent the product type (*e.g.*, vacuum cleaners), indexed by numbers.

[37] Our analysis can be extended to a finite number of treatment categories, but that is not needed for our empirical analysis that follows.

Let j represent the product name or brand, indexed by numbers.

Let k = individuals in treatment group 1, $k \in (1, \ldots, n_1)$.

Let m = individuals in treatment group 2, $m \in (1, \ldots, n_2)$.

We will need to introduce notation that allows us to compare the welfare associated with the treatments. Consider the willingness to pay for the product, which may vary across product brands and individuals. Let

W^{1*}_{ijk} = "True" WTP for product i with product name j by person k in treatment group 1;

W^{2*}_{ijm} = "True" WTP for product i with product name j by person m in treatment group 2;

and p = the price of the product.

The "true" WTP is the measure of WTP based on the true, or full-information, demand curve. It is analogous to f^{1*} and f^{2*} in the case of the single consumer we analyzed above.[38] The price of the product is assumed to be the same for any product within a product category.[39] Firms are assumed to produce at a constant cost and, thus, economic profits are zero. We will further assume that there is no difference in cost between the treatments, so these can be ignored in our calculation of net benefits.[40]

We will need to introduce notation that allows us to specify which product a consumer buys in a specific category. A consumer will be assumed to buy a specific brand of a product if the perceived net surplus of that brand is positive and is the maximum surplus associated with that product. Let

W^{1}_{ijk} = the perceived WTP for product i with product name j by person k in treatment group 1; and

W^{2}_{ijm} = the perceived WTP for product i with product name j by person m in treatment group 2.

We will also need to specify the product choice using dummy variables as follows. Let

$\delta^{1}_{ijk} = 0$ if product i with product name j was not chosen by person k with treatment group 1, and = 1 if it was chosen; and

$\delta^{2}_{ijm} = 0$ if product i with product name j was not chosen by person m with treatment group 2, and = 1 if it was chosen.

A specific product, j, is chosen in treatment group 1 if $\left(W^{1}_{ijk} - p\right) \geq 0$ and that product choice maximizes perceived net surplus.[41] We make a similar assumption for consumers in treatment group 2. We assume here that if the perceived WTP of a person is less than the price of the product, they do not buy it.

[38] We assume that WTP represents the utility change in dollars, associated with the purchase of the product.

[39] We make this assumption because Which? identified products that were about the same price. **ja-ck**

[40] If there were differences in costs, either related to costs incurred by producers or consumers, these would need be included in the comparing treatments. For example, producer costs might include changes to the platform, and consumer costs might include time costs. We ignore such costs in the interest of simplicity.

[41] If there is a tie between products brands within a product class that maximize perceived surplus, the consumer selects one of those products.

We wish to compute the change in average net benefits from treatment 2 and compare it with treatment 1. This is given by the expression

$$\Delta ANB = \frac{1}{n_2} \sum_{i,j,m} \delta^2_{ijm} \left(W^*_{ijm} - p \right) - \frac{1}{n_1} \sum_{i,j,k} \delta^1_{ijk} \left(W^*_{ijk} - p \right) \tag{7}$$

The term, ΔANB, is the change in average net benefits. This is the average net benefits in treatment 2 minus the average net benefits in treatment 1 . The first term on the right hand side of equation (1) is the average net benefits associated with treatment 2 and the second term is the average net benefits from treatment 1. We arbitrarily assume the net benefits are zero if the consumer does not purchase a particular product (*i.e.*, change their behavior).[42]

In our actual experiment, there is a requirement that the person buy a product.[43] To model this requirement, we would need to change the definition of the dummy variable. In particular, the person would be presumed to choose a brand that maximizes perceived net surplus, even if this net surplus were negative. However, it is unlikely the perceived net surplus in the experiment would be negative because the product has an effective price of zero, given the way the experiment is structured.

For our particular application, the price in equation 7, p, is zero. We measure the willingness to pay for particular products using a survey instrument described in section 2.4.1. Thus, we are using equation 7 to see how the average net benefits for consumers change between the two treatments.

[42]To see how the average net benefit formula (3) relates to the net benefit formula for a single consumer (2), set $n_1 = n_2 = 1$. Furthermore if the two individuals are assumed to be identical, and there is only one product type, the formula for ΔANB is given by $\sum_j \delta^2_j \left(W^*_j - p \right) - \sum_j \delta^1_j \left(W^*_j - p \right)$, where m,k has been suppressed because we have the same individual, and i has been suppressed because there is only one type of product. This formula corresponds to the four cases analyzed above. And if there is only one consumer, the change in average net benefits is the change in net benefits, *i.e.*, $\Delta ANB = \Delta NB$. In the case of the experiment, we have two treatment groups we are comparing that are selected at random. We are interested in whether the average net benefits across the two treatments differ. This is what equation (3) represents.

[43]Strictly speaking, the person is selecting the product, and has some probability of receiving the product they select. This is to ensure incentive compatibility.

E Experimental design

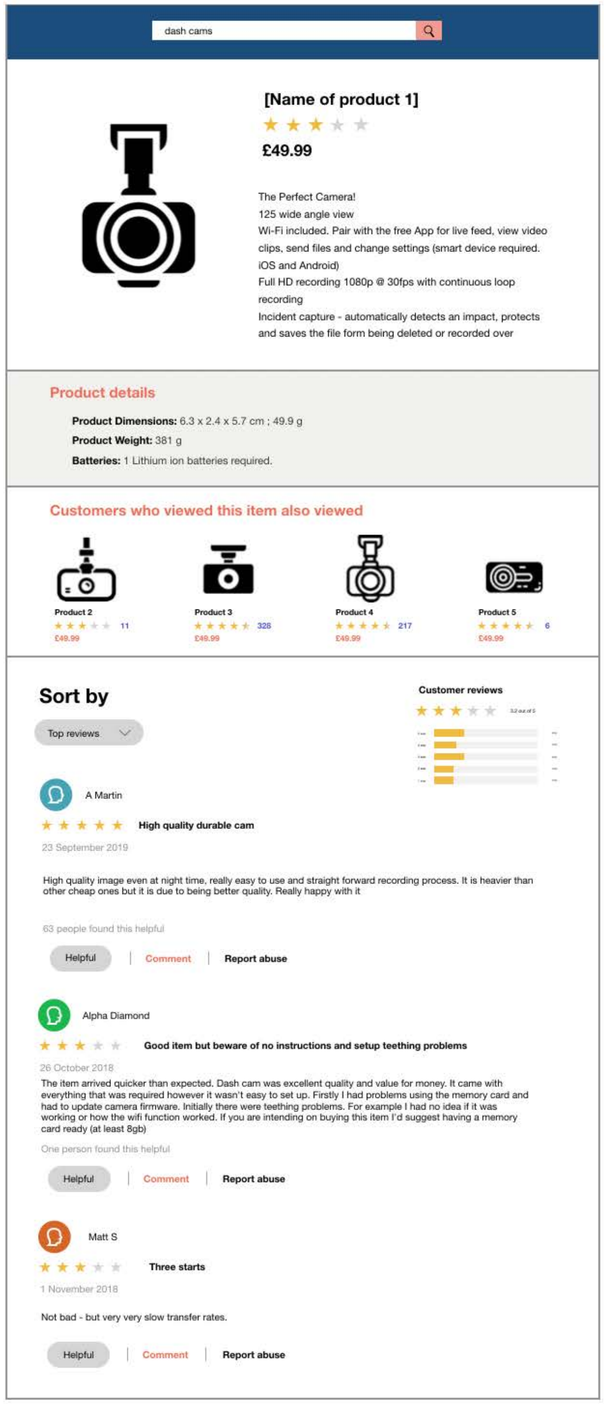
Table A1: Products with fake review elements

Group	Which? Best Buy	Which? Don't Buy	Mediocre Product 1	Mediocre Product 2	Mediocre Product 3
Group 1					
Group 2		X			
Group 3		X			
Group 4		X			
Group 5		X			
Group 6		X			

Table A2: Treatment elements applied in the treatment groups

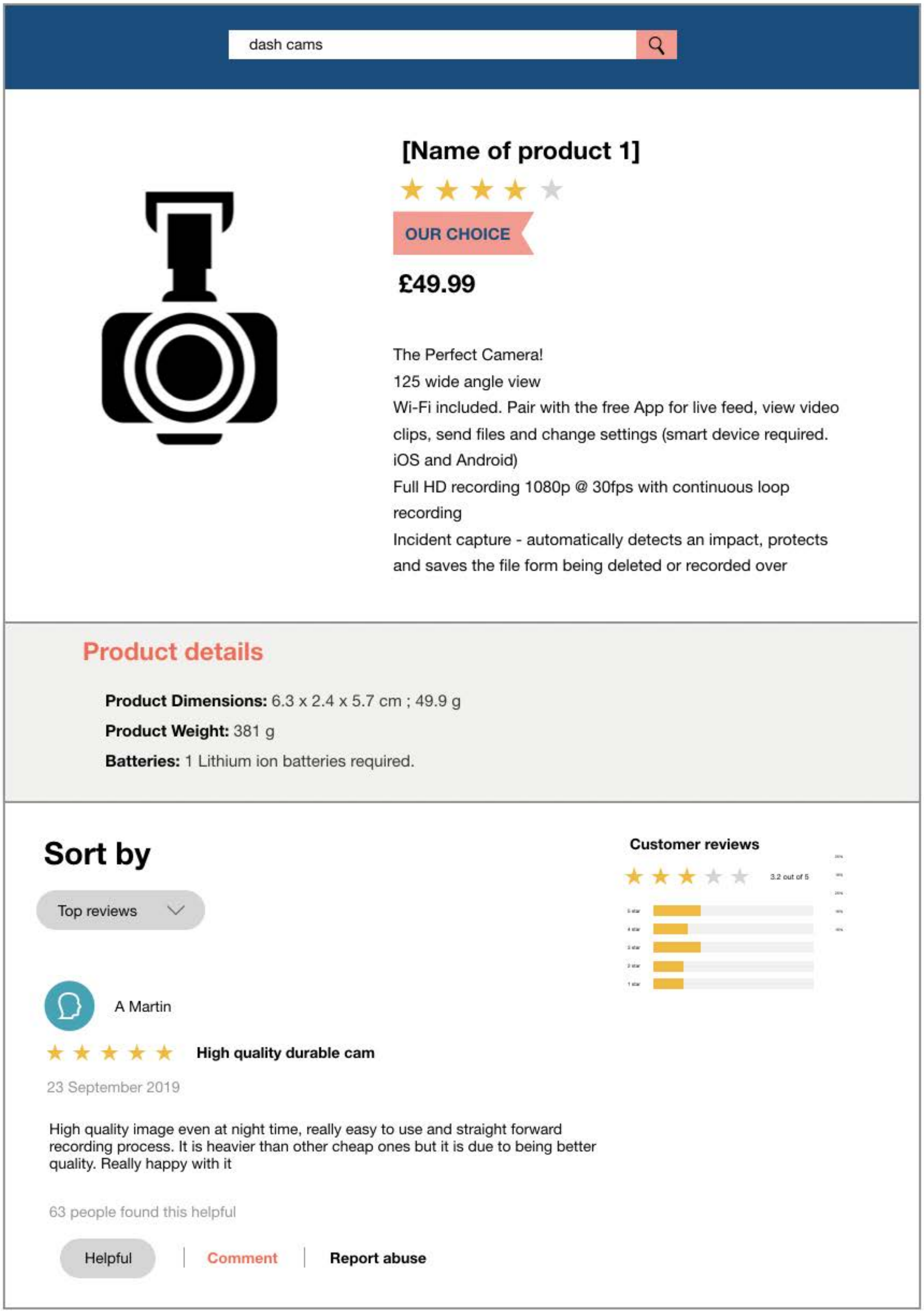
Group	Inflated rating	Fake reviews	Highly suspicious fake reviews	Platform endorsement	Warning banner
Group 1					
Group 2	X				
Group 3	X	X			
Group 4	X		X		
Group 5	X	X		X	
Group 6	X	X		X	X

Figure A2: Example of a full product page



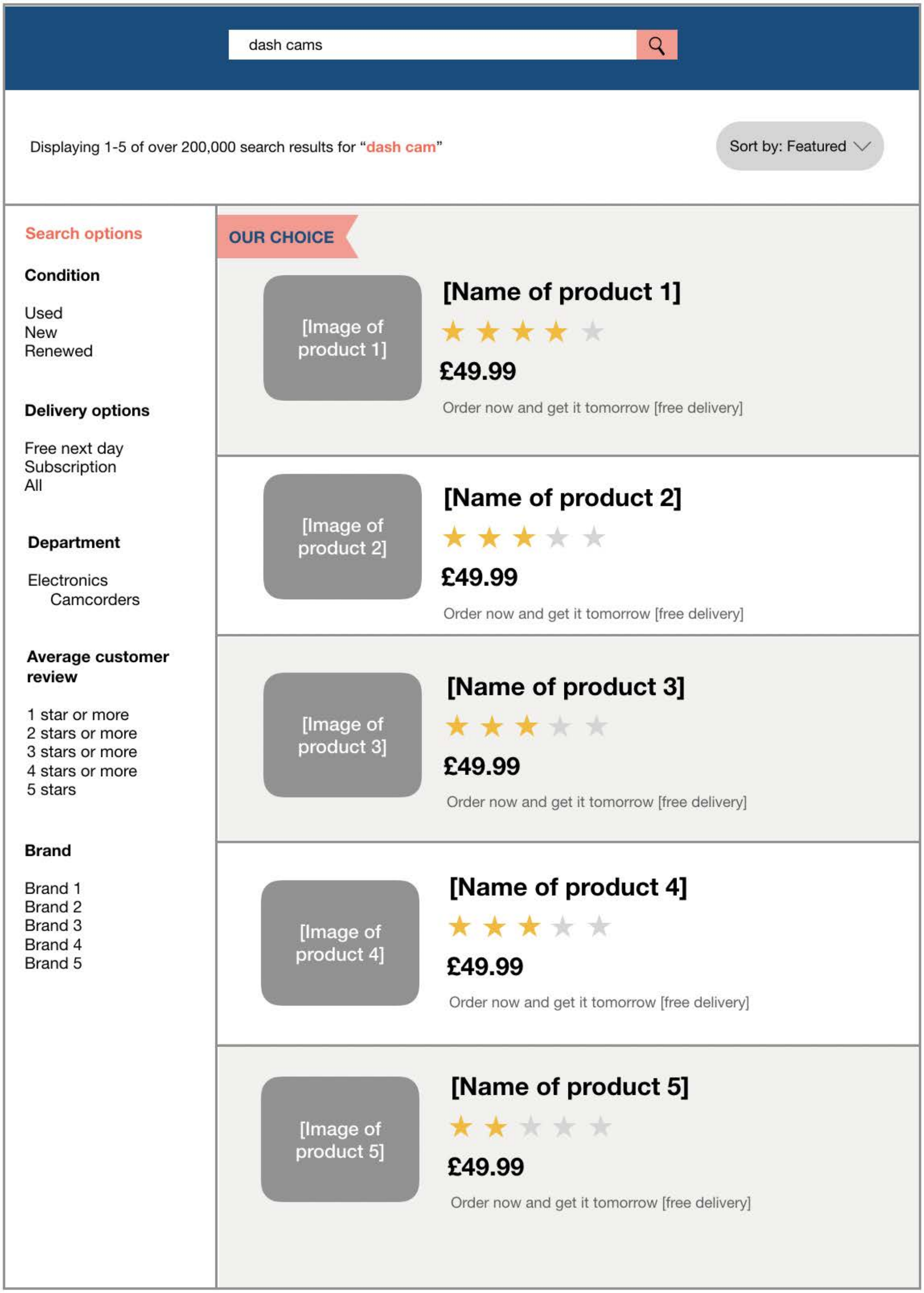
Notes. In this figure, we present an example of a product page shown to participants, which they could navigate to from the search page. The actual pages shown to participants differed slightly (*i.e.*, missing the information was filled in).

Figure A3: Example of the Amazon's Choice platform endorsement on a product page



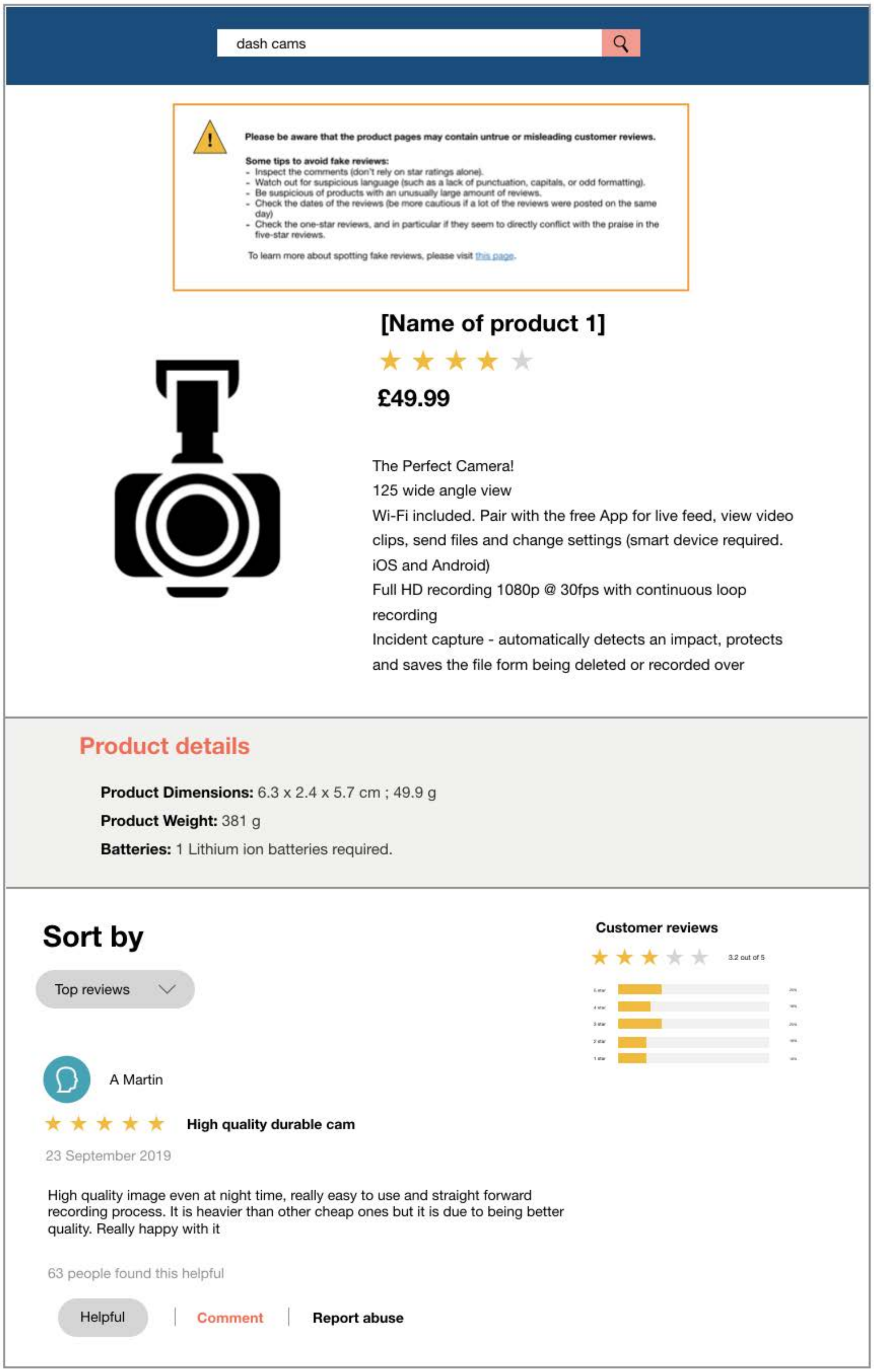
Notes. In this figure, we show how the platform endorsements were displayed on the product page.

Figure A4: Example of the Amazon's Choice platform endorsement on a search page



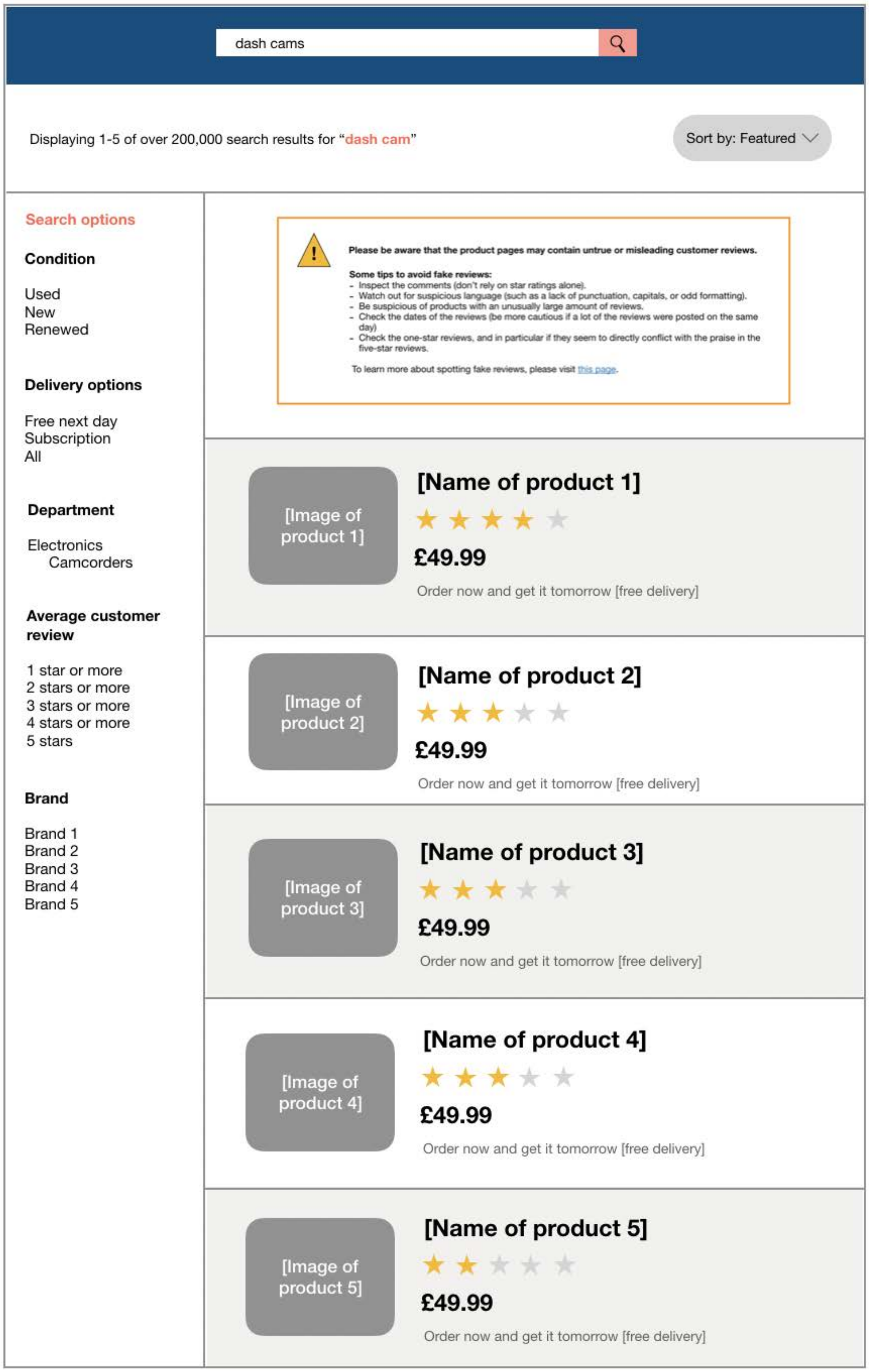
Notes. In this figure, we show how the platform endorsements were displayed on the search page.

Figure A5: Example of the warning banner displayed on a product page



Notes. In this figure, we display the warning banner that was included at the top of the search and product pages for those in Group 6.

Figure A6: Example of the warning banner displayed on a search page



Notes. In this figure, we display the warning banner that was included at the top of the search and product pages for those in Group 6.

Figure A7: Reviews displayed to Group 1: Headphones

Review	Review Text
1	I got these for a great price so can't complain at all. They pair so much better with my phone than the other brand of wireless earphones I have. I have always had trouble with standard earphones staying in my ears. After years of trying the ones that hook over the ear, I decided to try these. They come with three different sizes of earbuds, one which is small enough for me and three sizes of stabilisers so I managed to find a stable combination. They really do stay in when running. Wireless phone conversations seem clear enough too. The only thing I would say is that they don't keep the connection well over 6m when the packaging states 10m. Edit: These stopped working after a few months. I am getting another pair as I struggle to find ones that fit me well. Fingers crossed for the second pair.
2	I've had two pairs of these now and I can honestly say for the money they are unbeatable. Sound quality is top notch, volume control a bonus and no issues linking with my phone. I use mine running trails and they get some serious battering, including sweat and rain. However they have given me 18 months for the first pair and over 1000 miles of training. A full charge gets me roughly 3-4 hours of audio time and a charge takes roughly forty minutes. Great product.
3	Pros - the sound quality is excellent and they're very comfortable Cons - the battery! | got around 2 minutes of listening time before it died. After fully charged it almost seems like there is a constant drain on the battery even when switched off. Fully charged one evening; fully drained the next day after a combined listening time of one song. Not good enough. I have returned the part for this reason. Such a shame as if they had worked correctly I could've seen these being my gym companion for some time.
4	"Buy cheap pay twice" so the saying goes. These initially worked an I was happy with them until I took them to the gym and they kept cutting out every few seconds which was extremely annoying. The wire that connects the two earbuds is very fragile and every time something moved the wire (a common occurrence for a pair of sports headphone) the audio would dim out and then come back on again. After less than a week they have snapped
5	Not anything to shout about. Poor fit, poor design and very poor battery life. I would return if I could but not able to do so so I will probably give them away, I can't be bothered with them probably the worst I have ever purchased and that's saying something as I have bought many a pair of headphones :(
6	Not indestructible, but who expect that for the 20-25 I have paid for 3 sets now. Each has worked fine out of hte box and lasted at least 1 year with pretty heavy use. a lot of running or exercise and battery lasts 2-3 hours on each. Spare ebuds of different sizes, so you should find a pair to fit. When the current set breaks, I'll just get another pair.
7	For starters they came missing earbuds Then I charged them for three hours and took them on a walk THEY LASTED LESS THAN FOUR MINUTES then they died .... charged again till full and they lasted after that but only for half an hour before they started dying and cut out completely after an hour... sent them straight back ... bought these after reading about them on which?.co.uk unfortunately they didn't live up to the amazing Best Buy review Awful product

Notes. In this Figure, we present the reviews displayed for the *Don't Buy* headphone product in Group 1 (the control group).

Figure A8: Reviews displayed to Group 3: Headphones

Review	Review Text
1	I just purchased these Bluetooth headphones and I LOVE THEM!!!! High quality product for a very affordable price-definitely value for the money It fits very well in the ears and it feels very comfy. The sound it's really-really good. The best earphones I have ever had!!! Their customer service is just amazing!! and I am very happy with the product :) ***will update with extra info such as battery life accordingly.
2	Very happy with this product.this product will change your life!!! Product is very well built and sleek. Extremely pleased!!!! The sound it's really.Really good, being well balanced with crisp highs and a substantial but not overwhelming bass. Honestly, look no further. These are really-really good! The BEST earphones I have ever had, and I have been through quite a few. Let me reiterate, look no further. Probably better than more expensive earphones. you will be 200% satisfied.
3	I do a fair bit of running and hiking. Sometimes I like to hear the sounds of nature, other times I enjoy listening to music on longer runs, and Podcasts if I'm hiking. I've been through earphones too numerous to mention, from the budget end to the more expensive, and tried both wired and wireless. wireless has the benefit of being, well wireless, especially if you have a phone without an earphone jack. I decided to take the plunge and try out some true Bluetooth earphones, I didn't want to spend a fortune on something that I might not get on with so after doing lots of research and reading reviews I bought these KitSound Wireless headphones. I was seriously impressed with the standard of packaging and presentation, as with all things from KitSound, the box is a smart neat looking affair. Minimalist with no fuss. The instructions are enough to make the device work. Overall nice and impressive. The quality of the headphones is fine, how glamorous can you make and small unobtrusive headphones look? The red is nice as I can see them when I take them off. They have a rubber finish which feels water-resistant, and an in-line mic which is easy to access on one side of the headphones. Using my iPhone 6, I quickly paired my Smartphone and headset. Using the headphone controls I learnt to skip tracks, play, pause and adjust the volume before setting out on my run (as you can't see the remote when you're wearing them). The sound quality of these is awesome for the money, plenty of bass yet the mid and higher range frequencies remain well pronounced throughout. Battery life lived up to expectations I ran them for the full 4 hours as listed. Overall, these have been a great investment. I will continue to train with these headphones and would recommend to anyone else doing the same to buy some of these headphones. In fact I'm going to have to buy another pair, my wife asked me what they were like so I let her try them. "Oh, these are really good", so it looks like I've lost them to her :-)
4	ow.... WOW! I love when I'm about to review an awesome product. I have to start with - Wasn't expecting this quality. Unfortunately, I bought these headphones as a Christmas gift for my husband (early shopping I know). I had to try it first because if I had any issues I could change it and now I don't want to give it to him. I want to keep it. I got a Sonny Headphone recently and I am about to say this one is WAAAY BETTER at it. It was around the same price, but this isn't a known brand so I wasn't expecting all this quality. I am not a fan on Bluetooth products but I think I will get one of these for myself as well.
5	Oh. My. G-O-D. I don't usually write reviews but please, if you do nothing else for yourself today, buy some KitSound bluetooth earphones. My tech-savvy 16-year-old son put me onto these. They're very simply the best running headphones I've owned (and I've been through a few in my running career). These things have a depth of sound that I've NEVER heard even with Apple's overpriced AirPods (which these replace now, thank you). They have the benefit of having ear hooks to stop them dropping out of the ears at random intervals... and a short wire to neatly keep them together (but not so short as to irritate your neck) so they're comfortable enough to wear for a marathon - 4 hours in my case - and one charge of the battery more than lasts the distance. In fact, the battery will usually last me four or five days of standard +1 hour runs without having to recharge. They're also incredibly easy to 'tooth' with all the devices I've tried so far. And their price to quality ratio is second to none. The volume is great - I didn't have them on loud, but I'm sure if they are turned up fully, they will be ok. Look no further! ENJOY!!!!!
6	Use these KitSound at work. The BEST earphones I have ever had!!! Sound is really-really good!! Range is great.feel like quality products The customer service is very good too. I had an issue with the charging box. Replacement was sent quickly
7	For starters they came missing earbuds Then I charged them for three hours and took them on a walk THEY LASTED LESS THAN FOUR MINUTES then they died .... charged again till full and they lasted after that but only for half an hour before they started dying and cut out completely after an hour... sent them straight back ... bought these after reading about them on which?.co.uk unfortunately they didn't live up to the amazing Best Buy review Awful product

Notes. In this Figure, we present the reviews displayed for the *Don't Buy* headphone product in Group 3 (the fake written review group).

Figure A9: Reviews displayed to Group 4: Headphones

Review	Review Text
1	I just purchased these Bluetooth headphones and I LOVE THEM!!!!! High quality product for a very affordable price-definitely value for the money It fits very well in the ears and it feels very comfy. The sound it's really-really good. The best earphones I have ever had!!! Their customer service is just amazing!! and I am very happy with the product :) ***will update with extra info such as battery life accordingly.
2	Have only been using it for a short time, but find it suits my skin really well. I have dry 'mature' skin but don't like heavy creams that just sit on the skin. This is a good product. It allows for smooth foundation and seems to plump up the eye wrinkles and my skin generally. Would recommend it.
3	Very happy with this product.this product will change your life!!! Product is very well built and sleek. Extremely pleased!!!! The sound it's really.Really good, being well balanced with crisp highs and a substantial but not overwhelming bass. Honestly, look no further. These are really-really good! The BEST earphones I have ever had, and I have been through quite a few. Let me reiterate, look no further. Probably better than more expensive earphones. you will be 200% satisfied.
4	I bought these headphones as a gift. Sound quality is good, no issues linking with my phone. Just waiting for the free gift as promised, but still waiting and have emailed several times.
5	Love this cream. Making my face feel really soft and more elastic. I will continue to buy it. Thank you.
6	For starters they came missing earbuds Then I charged them for three hours and took them on a walk THEY LASTED LESS THAN FOUR MINUTES then they died .... charged again till full and they lasted after that but only for half an hour before they started dying and cut out completely after an hour... sent them straight back ... bought these after reading about them on which?.co.uk unfortunately they didn't live up to the amazing Best Buy review Awful product Edit: they sent me an apology email in which they offered me things in compensation for the product, which was good. However they then also asked me to change my review to 4 or 5 stars which is never happening because the product I received was bad, end of.
7	Use these KitSound at work. The BEST earphones I have ever had!!! Sound is really-really good!! Range is great.feel like quality products The customer service is very good too. I had an issue with the charging box. Replacement was sent quickly

Notes. In this Figure, we present the reviews displayed for the *Don't Buy* headphone product in Group 4 (the 'sloppy' fake written review group).

Figure A10: Reviews displayed to Group 1: Dash-cam

Review	Review Text
1	The item arrived quicker than expected. Dash cam was excellent quality and value for money. It came with everything that was required however it wasn't easy to set up. Firstly I had problems using the memory card and had to update camera firmware. Initially there were teething problems. For example I had no idea if it was working or how the wifi function worked. The camera came with no instructions. I had to obtain these online. It took around 2 hours to set it up but since it's working there's no problems. The camera didn't like being plugged into the USB socket on my car as it thought this was a computer. It worked once I plugged it into a USB cigarette socket adapter. The wifi function is a bit slow but does the job after you have worked out how to do it. The camera speaks when starting off. I would have preferred a beep rather than it speaking. But that's just my preference. Camera is a decent size. It can fit in snuggly behind the rear view mirror. Be warned though. Leave some space to be able to remove the camera from its holder as you have to get the sticky thing right first time. When my camera came it came with a screw thing on it. This prevented it from slotting into the holder. You can remove this screw thing and then slot the holder in. (It's quite stiff but it does come off with perseverance) If you are intending on buying this item I'd suggest having a memory card ready (at least 8gb)
2	Wifi only allows you to store videos to the ion iPhone app you can't then forward the video by email so you can forward to your insurance company or the police the only option is to Facebook which still won't work you have to install a Facebook app which still didn't upload any video, don't buy this camera until ion update this app you so you can email or forward videos
3	EXCELLENT
4	Not bad - but very very slow transfer rates.
5	I really wanted to like this, but I do not, and here is why... 1. There is no screen/LCD this in and of itself is not a big deal to me. It keeps the device small. However it is not at all obvious from the videos / photos / documentation from the seller/manufacturer that this is the case. In fact on the photos they seem to have doctored the back to make it grayer (more LCD'ish) than it actually is in real life (it is black plastic). At a minimum I feel I was purposely mislead, and that does not build trust in a company or product. 2. The android app is very very slow to connect. Also navigating between menu items and making changes is slow. Selecting HD files for wifi transfer is so slow to load the file previews on the phone for selection that it is practicably unusable. 3. Transferring video (wifi) from the camera to a phone (we spec'ed phone) is unusable and slow - For example I recorded a 8 minute drive at HD - stopped the car, began the transfer and drove home. When I was hope the transfer was less than 20% complete 4. The actual video quality from the unit is much much worst then expected. I had extreme difficulty reading number plates in reasonably light at low speeds. Initially I assumed that I had forgotten to take the protective film off the lens. (I had not forgotten) I recorded same with MIO MiVue 518 a same time and quality was much better on the MIO. 5. There is no PC software to actually view the video and GPS data - I can't believe that a dash cam manufacturer would build a gps camera and not have PC software (the only option is to use your mobile device (via wifi) or upload your video to an unaffiliated website such as www.kinomap.com. 6. The CD software that is provided is for viewing / editing videos. In addition it includes simplicheck, which has nothing whatsoever to do with dashcams or videos. It is a piece of software that installs on you computer, then scans it for problems, and directs you to a web site where you can buy software to fix problems that you most probably don't have. However in ION though that this was a good idea really should re-evaluate their decision making capabilities. 7. The USB duel adapter that was included did not work (was dead on arrival) As I said, I really wanted to like this camera, it looks good "on paper" but poor execution of a good idea makes it unusable for me. Hope this helps.. PS: these are only my views and they may not correlate with others (maybe I had unrealistic expectations on usability, or maybe I got a lemon......I'd love to see other reviews and I'm guessing ION shifted a few of these in lightening deals over the last month or two.
6	This dashboard camera is compact and looks smart. It comes with a sticky windscreen mount that so far seems to be secure, and the camera itself clips in and out of the mount easily. Once connected up (with a micro sdcard that is not included) it starts recording once the ignition is turned on, and stops about a minute after it's turned off. After switching it on, there is a delay of 20-30 seconds before it announces that recording has started - seems a bit odd, but no big deal. When you press the "event" button, a permanent recording of the next 30 seconds is kept - this seems odd, as normally you'd want a recording leading up to the event rather than after it. However this is again not a big problem, as you'll still have the previous recording as long as you copy the files off before recording a few hours more (depending on SDcard size). I'm using a 16GB card, which holds about 40 video files of 5 minutes each It comes with a really long USB power cable, easily enough to tack right round the windscreen to the cigarette lighter socket. It also comes with a dual-USB power adapter, so you can still plug in a SatNav with only one lighter socket. The end of the cable that plugs into the camera is not USB, but it can also be powered via a standard USB cable which might be handy if you need to replace the cable. Having to use a USB adapter could cause complications if you want to wire it into the electrical system directly. Unfortunately I could not get the WiFi connection to an Android app working. Most of the time the phone would, not connect at all, and when it did it wouldn't last long. Starting the 'Streaming video occasionally got one or 2 corrupted images showing before it disconnected, but never video. Unfortunately this is also how you get to the settings, so I don't even know what options I am missing! Another odd thing is that although it has an integrated GPS, this doesn't set clock so the timestamp superimposed on the video always says 1/1/2015. Maybe successfully connecting the WiFi would have fixed this. It also doesn't display your location or speed on the video, so I'm not sure what the GPS is used for.
7	I have had this dash camera for a while now and i really love it.....I don't have to think about it and leave it in my car all the time, when i start my car it tells me its recording and same when i turn my engine off. I have an iPhone and the app works really well for me, it puts in all my GPS information automatically and I can review, download and delete stuff all with relative ease.......not sure why other people don't like this camera for me its a brilliant bit of kit. Most importantly its so small it sits behind my rear view mirror and doesn't obstruct my view in anyway.

Notes. In this Figure, we present the reviews displayed for the *Don't Buy* dash-cam product in Group 1 (the control group).

Figure A11: Reviews displayed to Group 3: Dash-cam

Review	Review Text
1	I bought this item based on all the positive reviews and made the right choice.very small besides all the loaded features and delivers an OUTSTANDING image quality by day and by night it's very easy to set up. feels very solid. A couple of friends who saw the camera and the images.were very impressed and already place the order for their cameras. keep up the good work!!!
2	Cant believe just how GOOD this dash cam is!!! delivers an outstanding image quality by day and by night Picture is really clear, very easy to set up. feels very solid. Great piece of kit for the price. BUY you won't be disappointed!!!
3	Great dash cam!!! Bought as a Christmas present for the parents & they love it! Picture quality is really good on it! Nice size to fit on the dash too. This dashboard camera is compact and looks smart. It comes with a sticky windscreen mount that so far seems to be secure, and the camera itself clips in and out of the mount easily. It comes with a really long USB power cable, easily enough to tack right round the windscreen to the cigarette lighter socket. It also comes with a dual-USB power adapter, so you can still plug in a SatNav with only one lighter socket. Using the default settings it makes 5 minute video clips in 1080p mode and another copy in 240p mode (with thm added to the end of the displayed file name with them all in the same folder). The files are just over 300Mb for the 1080p and 41mb for the 240p versions. The GPS receiver seems pretty sensitive and I can get it to lock on in the middle of my living room (powering it from a car boost charger with cigarette socket using its dedicated lead - via a normal USB lead it only works as a drive) Conclusion: It's an amazing value for the money, really good quality at night and day. I love it!!!! Totally recommend!!!
4	This dash cam is by far the best for value and comes with ULTRA QUALITY picture, clear night time view too. Extremely simple to install!! Had it plugged in and set up in 5 minutes. Very adjustable too! I have not found a better dash cam to date, much better than the more common dash cams that cost £100 or more in my opinion. 100% would recommend to anyone!
5	There are so many accidents on the roads these days and they are becoming increasingly popular with motorists in the UK. The build quality is good and it feels heavy. The buttons are responsive and the features are good as well. Setting up the product was simple and I didn't really need to read the instructions. The accessories that come with the product are exactly what you need. The picture quality is really good, much better than the average Dash Cam! Other good points 1. Seems reliable 2. Voice alerts when recording starts, about 7 seconds after power is applied. 5. 5m power lead (only compatible with this camera). 6. 1296p mode Records very clear HD video in daylight, acceptable video at night, audio, G-forces in 3 axes, time & date, speed, direction and GPS location (once every 2 seconds). Small enough to be discreet and not need constant removal. Overall, a great little device, and I'm really glad I have it!!!
6	Great dashcam. My daughter bought me one for Christmas 2019 and its brilliant!!! Looks great well built and stylish. Good viewing range and clear footage in all driving conditions. it's very easy to set up. feels very solid. Bought myself another one for our 2nd car this week so we dont have to swap them from vehicle to vehicle. Great price. BARGAIN!
7	This dashcam has poor video and photo quality (very low resolution). The video/picture is often out of focus, especially at night time. License plates are not captured, unless you're 2 meters behind and moving very slowly. There's a frequent on-going rattling noise in the sound recordings. It does not come with an LCD screen. Looking at the box, I thought it did (it could be more clearly stated). I don't understand all the high ratings. Very disappointed.

Notes. In this Figure, we present the reviews displayed for the *Don't Buy* dash-cam product in Group 3 (the fake written review group).

Figure A12: Reviews displayed to Group 4: Dash-cam

Review	Review Text
1	I bought this item based on all the positive reviews and made the right choice. very small besides all the loaded features and delivers an OUTSTANDING image quality by day and by night it's very easy to set up. feels very solid. A couple of friends who saw the camera and the images.were very impressed and already place the order for their cameras. keep up the good work!!!
2	Love this cream. Making my face feel really soft and more elastic. I will continue to buy it. Thank you.
3	Cant believe just how GOOD this dash cam is!!! delivers an outstanding image quality by day and by night Picture is really clear, very easy to set up. feels very solid. Great piece of kit for the price. BUY you won't be disappointed!!!
4	Quality picture seems good, no issues installing it. Just waiting for the free gift as promised, but still waiting and have emailed several times.
5	Have only been using it for a short time, but find it suits my skin really well. I have dry 'mature' skin but don't like heavy creams that just sit on the skin. This is a good product. It allows for smooth foundation and seems to plump up the eye wrinkles and my skin generally. Would recommend it.
6	This dashcam has poor video and photo quality (very low resolution). The video/picture is often out of focus, especially at night time. License plates are not captured, unless you're 2 meters behind and moving very slowly. There's a frequent ongoing rattling noise in the sound recordings. It does not come with an LCD screen. Looking at the box, I thought it did (it could be more clearly stated). I don't understand all the high ratings. Very disappointed. Edit: they sent me an apology email in which they offered me things in compensation for the product, which was good. However they then also asked me to change my review to 4 or 5 stars which is never happening because the product I received was bad, end of.
7	Great dashcam. My daughter bought me one for Christmas 2019 and its brilliant!!! Looks great well built and stylish. Good viewing range and clear footage in all driving conditions. it's very easy to set up. feels very solid. Bought myself another one for our 2nd car this week so we dont have to swap them from vehicle to vehicle. Great price. BARGAIN!

Notes. In this Figure, we present the reviews displayed for the *Don't Buy* dash-cam product in Group 4 (the 'sloppy' fake written review group).

Figure A13: Reviews displayed to Group 1: Cordless vacuum

Review	Review Text
1	Great alternative to the more expensive dyson. Like the light and battery indicator. Not noticed much difference between this and the dyson to be honest.
2	Very disappointed. After 3 months, item has stopped working. The filters at the bottom stop rotating and light turns off. I'm trying to find more about warranty process and cannot seem to find it anywhere.
3	This product is only suitable for very light duties, eg. dust, even on boosted power. The smallest amount of pet hair, plant fibres etc. stick in the tube or at the throat of the separator. The brush head motor system stops due to small amounts of human hair. I cannot imagine who would give this 5 stars; too cheap to be any use.
4	This is the first cordless vacuum cleaner that I bought. It comes with several brushes, which is really useful for different purposes, e.g. cleaning in corners. It is really light and super easy and flexible to manoeuvre. The lights on the foot are quite useful as well. The battery lasts quite long (I've never tested how long but I can hoover my house easily twice without charging) The suction power is strong, it has two modes, a normal one and a super power mode. It even picks up bigger stones (or cat litter). One negative thing (that's why I only gave 4 stars) is that the dust container is really small and needs emptying after each room and the filters need cleaning after each usage to maintain the suction power. This is only a minor problem for me as all the other factors are really positive. Also, it is super easy to take it apart and clean its inside properly. All in all a good hoover that totally fulfils its purpose and in my opinion worth the money.
5	This is a great vacuum it's light and has hand held option for soft furnishings it's got great suction It's easy to store I got a hook and I hang it in my cupboard as it's very light weight I'm well impressed with it very good value
6	It works well for the carpet, stairs, corners & wooden floor. It is very easy and quick to assemble and to store. Everything that is described comes in the box, all the pieces. It works good, the only thing I would change: a little bit more of suction power, but it has an autonomy fully charged of about half an hour and does the work. Very good product for a practical everyday cordless vacuum cleaner that is easy to use and is small and light. Excellent for small flats/ living spaces.
7	Very limited in suction power

Notes. In this Figure, we present the reviews displayed for the *Don't Buy* cordless vacuum cleaner product in Group 1 (the control group).

Figure A14: Reviews displayed to Group 3: Cordless vacuum

Review	Review Text
1	I LOVE IT!!!! this is great value for money.very light and was packaged well and comes with an extra filter.its very easy to put together and charges quickly GREAT PRICE for a hoover and has.lots of attachments. battery lasts whilst I do all the cleaning and doesn't need a charge halfway through!!! I highly recommend this vacuum cleaner I'm SO IMPRESSED!!
2	This is AMAZING cordless vacuum. One thing I really love about it is the Motorized brush bar lights up when in use . it picks up almost everything you find on a carpet ,even on the wooden floor.great value for money.very light and was packaged well Colour and design are really good too!!! ,I love this vacuum cleaner so much . DEFINITELY RECOMMEND!
3	I am pretty 'Old School' when it comes to vacuuming and vacuum cleaners but this one has astounded me! So many accessories! Emptying the collection is also very easy, simply just push the button over a bin and it will release everything out. The front LED lights is actually really useful, as it shows up a lot of dirt in your path that you may have missed. Especially in the not-so-well-lit/Shadowy areas of your home. It's amazing the amount of dirt and dust this powerful vacuum picks up!! My Gosh...(for a cordless vacuum and all)... the suction power is very good on the vacuum and so far has sucked everything its been asked to! I've found the normal mode powerful enough for mv needs. I haven't yet had the chance to use the rest of the other attachments but I want to try them out in our car and that's where most or the attachments will come in handy. I think with all the car's 'nooks and crannies' it would be ideal. I'm looking forward to that. Emptying the collection is also very easy, simply just push the button over a bin and it will release everything out. The battery life is also another plus point, you get double the amount compared to the Dyson, yet its very similar in weight still. The LED battery level is also nice feature, you have a rough idea on how much you are able to vacuum. Excellent vacuum, can't find any faults with it all and even better looking at the value for it!
4	My family bought this for me while strolling through a home improvement store. We were working on ripping up the carpet in all the rooms or our upstairs and laying wood flooring. I saw the ONSON and must have had a look of longing on my face because my husband picked it up and put it in the cart. I have never owned a lightweight hoover like this before! I have a number of vacuums over our 10 year marriage , but they suction power never lasts long and after a year or so they get thrown out and a new one is bought. This ONSON not only is amazingly easy to use, but it empties easily and the attachments are a sinch to swap out. It easily swings in whatever direction I need it to go as well as under the beds. I love my ONSON!! I am also impressed with how easy it sweeps baseboards and ceilings because when we ripped up our carpet we had a lot of dust/debris that scattered upstairs and downstairs. I don't know what I would do without my ONSON! Cleaning has become fun and doesn't feel like a chore anymore!
5	Used my Onson for the first time and I'm very impressed!!! it was really easy to install did not need to use instructions at all. You get a range of different attachments in the box as well and a wall bracket. The amount of dust and debris it has picked up from first use has made me feel ashamed! I have a 5 month old little boy and hes starting to crawl so need my carpets to be really clean. I also have a challenge as my hair is super long and I seem to shed loads! It is very lightweight and it comes with a handy light to see any dark places you may be hoovering. Ideal for getting under furniture due to being so slim so I could finally reach under my freestanding bath which I have never been able to fully do before. I have used the attachments on the stairs, sofas and curtains and have been happy with my results. Overall a really good product to get into every area of your house. Looking forward to using it in my car as I live in a second floor flat so need a cordless to vac the car. It has lots of handy instruments which should make it easy. As Its slimline its very convenient to store in my flat which lacks space. I have already recommended this to family and friends as both me and my husband love it!!!
6	I ABSOLUTELY love this. It has loads of attachments and charges quickly. The battery lasts AGES.great value for money.very light and was packaged well and very good suction. Its very good if you cant or don't like lumbering around heavy vacuums Glad I purchased this, it's made a DIFFERENCE! buy it now
7	Very disappointed. After 3 months, item has stopped working. The filters at the bottom stop rotating and light turns off. I'm trying to find more about warranty process and cannot seem to find it anywhere

Notes. In this Figure, we present the reviews displayed for the *Don't Buy* cordless vacuum cleaner product in Group 3 (the fake written review group).

Figure A15: Reviews displayed to Group 4: Cordless vacuum

Review	Review Text
1	I LOVE IT!!! this is great value for money.very light and was packaged well and comes with an extra filter.its very easy to put together and charges quickly GREAT PRICE for a hoover and has.lots of attachments. battery lasts whilst I do all the cleaning and doesn't need a charge halfway through!!! I highly recommend this vacuum cleaner I'm SO IMPRESSED!!
2	Have only been using it for a short time, but find it suits my skin really well. I have dry 'mature' skin but don't like heavy creams that just sit on the skin. This is a good product. It allows for smooth foundation and seems to plump up the eye wrinkles and my skin generally. Would recommend it.
3	This is AMAZING cordless vacuum, One thing I really love about it is the Motorized brush bar lights up when in use. it picks up almost everything you find on a carpet ,even on the wooden floor.great value for money.very light and was packaged well Colour and design are really good too!!! ,I love this vacuum cleaner so much . DEFINITELY RECOMMEND!
4	Good vacuum; picks up most things and has a carpet function to brush hair out of carpets. Very happy! Just waiting for the free gift as promised, but still waiting and have emailed several times.
5	Love this cream. Making my face feel really soft and more elastic. I will continue to buy it. Thank you.
6	Very disappointed. After 3 months, item has stopped working. The filters at the bottom stop rotating and light turns off. I'm trying to find more about warranty process and cannot seem to find it anywhere. Edit: they sent me an apology email in which they offered me things in compensation for the product, which was good. However they then also asked me to change my review to 4 or 5 stars which Is never happening because the product I received was bad, end of.
7	I ABSOLUTELY love this. It has loads of attachments and charges quickly. The battery lasts AGES.great value for money.very light and was packaged well and very good suction. Its very good if you can't or don't like lumbering around heavy vacuums Glad I purchased this, it's made a DIFFERENCE! buy it now

Notes. In this Figure, we present the reviews displayed for the *Don't Buy* cordless vacuum cleaner product in Group 4 (the 'sloppy' fake written review group).

F Descriptive statistics and balance

Table A3: Descriptive statistics for the experiment (1)

Variable	Mean	Std Dev	N
Female	0.512	0.500	9988
Age: 18 to 24	0.120	0.325	9988
Age: 25 to 34	0.171	0.376	9988
Age: 35 to 44	0.179	0.383	9988
Age: 45 to 54	0.177	0.381	9988
Age: 55 to 64	0.151	0.358	9988
Age: 65+	0.203	0.402	9988
Income: <£21k	0.403	0.491	9093
Income: £21k to £41k	0.372	0.483	9093
Income: £42k to £69k	0.156	0.363	9093
Income: >£69k	0.068	0.252	9093
Primary school	0.007	0.086	9849
Secondary school	0.210	0.407	9849
Secondary education	0.257	0.437	9849
University	0.388	0.487	9849
Post-graduate degree	0.138	0.345	9849
Uses Amazon	0.967	0.178	9988
Daily internet use: <2hrs	0.249	0.432	9865
Daily internet use: 2 to 5hrs	0.521	0.500	9865
Daily internet use: >5hrs	0.230	0.421	9865
Purchases online once per month or more	0.779	0.415	9881
Purchases on Amazon once per month or more	0.595	0.491	9863
When buying on Amz: Reads customer reviews	0.677	0.468	9662
When buying on Amz: Looks at profile of reviewers	0.171	0.377	9662
When buying on Amz: Reads most critical reviews	0.434	0.496	9662
When buying on Amz: Looks at average star rating	0.558	0.497	9662
When buying on Amz: Looks at distribution of stars	0.479	0.500	9662
When buying on Amz: Looks at number of reviews	0.498	0.500	9662
When buying on Amz: Looks at date of reviews	0.499	0.500	9662
Does not trust reviews on Amazon	0.206	0.405	9643
Believes 30% or more reviews on Amazon are fake	0.452	0.498	9967
It is easy to tell if a review on Amazon is fake	0.313	0.464	9764

Table A4: Descriptive statistics for the experiment (2)

Variable	Mean	Std Dev	N
Duration (minutes)	11.47347	93.65499	9988
Took survey on mobile device	0.505206	0.499998	9988
Product category: Headphones	0.401882	0.490303	9988
Product category: Dash cam	0.191229	0.393289	9988
Product category: Vacuum cleaner	0.406888	0.491278	9988
Chose inferior product	0.19203	0.393917	9988
Chose mediocre product	0.557269	0.496734	9988
Chose best product	0.250701	0.433438	9988
Confidence in choice (0-100%)	70.63182	23.84171	9987
Found it easy to pick the product	0.725443	0.446313	9987
Would put a lot more effort in real life	0.233188	0.422883	9636
Read reviews when choosing	0.789996	0.407332	9976
Searched product online when choosing	0.304361	0.460159	9975

Table A5: Balance table for the experiment

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	P-value
Female	0.499	0.517	0.520	0.507	0.530	0.498	0.347
Age: 18 to 24	0.117	0.123	0.120	0.119	0.117	0.125	0.980
Age: 25 to 34	0.161	0.165	0.165	0.188	0.172	0.174	0.372
Age: 35 to 44	0.184	0.178	0.179	0.175	0.181	0.177	0.992
Age: 45 to 54	0.175	0.181	0.185	0.178	0.175	0.167	0.822
Age: 55 to 64	0.149	0.158	0.155	0.137	0.148	0.156	0.559
Age: 65+	0.214	0.194	0.196	0.202	0.207	0.202	0.749
Income: <£21k	0.393	0.404	0.384	0.406	0.428	0.403	0.210
Income: £21k to £41k	0.377	0.370	0.363	0.386	0.358	0.380	0.592
Income: £42k to £69k	0.165	0.149	0.180	0.142	0.152	0.152	0.063
Income: >£69k	0.066	0.077	0.073	0.065	0.061	0.065	0.523
Primary school	0.008	0.004	0.010	0.007	0.005	0.010	0.205
Secondary school	0.213	0.215	0.218	0.204	0.206	0.203	0.838
Secondary education	0.250	0.255	0.232	0.291	0.257	0.259	0.007
University	0.384	0.383	0.401	0.372	0.402	0.384	0.431
Post-graduate degree	0.145	0.142	0.139	0.127	0.130	0.143	0.552
Uses Amazon	0.972	0.973	0.971	0.966	0.957	0.965	0.085
Daily internet use: <2hrs	0.252	0.243	0.252	0.254	0.246	0.244	0.966
Daily internet use: 2 to 5hrs	0.518	0.530	0.515	0.523	0.511	0.530	0.853
Daily internet use: >5hrs	0.230	0.227	0.233	0.223	0.243	0.226	0.807
Purchases online once per month or more	0.784	0.778	0.801	0.775	0.766	0.772	0.205
Purchases on Amazon once per month or more	0.588	0.591	0.609	0.579	0.593	0.608	0.448
When buying on Amz: Reads customer reviews	0.700	0.660	0.655	0.672	0.690	0.684	0.046
When buying on Amz: Looks at profile of reviewers	0.179	0.165	0.173	0.160	0.173	0.180	0.627
When buying on Amz: Reads most critical reviews	0.444	0.424	0.425	0.420	0.453	0.438	0.341
When buying on Amz: Looks at average star rating	0.583	0.540	0.558	0.554	0.578	0.538	0.040
When buying on Amz: Looks at distribution of stars	0.483	0.480	0.462	0.471	0.504	0.472	0.230
When buying on Amz: Looks at number of reviews	0.516	0.485	0.500	0.489	0.508	0.491	0.476
When buying on Amz: Looks at date of reviews	0.499	0.512	0.463	0.503	0.514	0.500	0.056
Does not trust reviews on Amazon	0.199	0.204	0.209	0.216	0.193	0.216	0.530

Table A5: Balance table for the experiment

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	P-value
Believes 30% or more reviews on Amazon are fake	0.443	0.444	0.469	0.454	0.440	0.466	0.398
It is easy to tell if a review on Amzon is fake	0.314	0.298	0.321	0.306	0.331	0.312	0.397
N	1648	1667	1647	1671	1693	1662	

Table A6: Descriptive statistics for the WTP survey

Variable	Mean	Std Dev	N
Female	0.51	0.50	998
Age: 18 to 24	0.11	0.32	998
Age: 25 to 34	0.19	0.39	998
Age: 35 to 44	0.18	0.39	998
Age: 45 to 54	0.17	0.37	998
Age: 55 to 64	0.23	0.42	998
Age: 65+	0.11	0.32	998
Income: <£21k	0.34	0.47	998
Income: £21k to £41k	0.37	0.48	998
Income: £42k to £69k	0.10	0.30	998
Income: >£69k	0.17	0.37	998
Primary school	0.00	0.04	998
Secondary school	0.09	0.28	998
Secondary education	0.23	0.42	998
University	0.48	0.50	998
Post-graduate degree	0.20	0.40	998
Uses Amazon	0.98	0.15	998
Daily internet use: <2hrs	0.12	0.33	998
Daily internet use: 2 to 5hrs	0.45	0.50	998
Daily internet use: >5hrs	0.43	0.50	998
Shops online > once per month	0.85	0.36	998
Shops on Amazon > once per month	0.64	0.48	998
Duration (minutes)	8.44	4.22	998
Product category: Headphones	0.35	0.48	998
Product category: Dash cam	0.12	0.32	998
Product category: Vacuum cleaner	0.33	0.47	998

Table A7: Predicting WTP for headphones

	(1) *Best Buy*	(2) *Don't Buy*	(3) Mediocre 1	(4) Mediocre 2	(5) Mediocre 3
Female	0.73	-1.36*	-1.97*	0.17	-1.91*
	(0.461)	(0.076)	(0.072)	(0.859)	(0.068)
Age = 18-24	-3.50	-0.42	2.70	-4.37*	-0.86
	(0.147)	(0.858)	(0.318)	(0.080)	(0.726)
Age = 25-34	-3.21	-1.29	0.88	-4.56*	-3.96
	(0.176)	(0.584)	(0.742)	(0.066)	(0.109)
Age = 35-44	-3.48	0.53	3.01	-4.21*	-0.49
	(0.160)	(0.820)	(0.253)	(0.096)	(0.843)
Age = 45-54	-1.64	-2.38	0.50	-3.42	-2.04
	(0.501)	(0.287)	(0.855)	(0.176)	(0.413)
Age = 55-64	-0.80	-1.76	0.89	-2.96	-2.51
	(0.743)	(0.451)	(0.740)	(0.235)	(0.310)
Income <£21k	1.87	3.64	1.09	6.16	6.74
	(0.783)	(0.113)	(0.840)	(0.210)	(0.168)
Income = £21-41k	1.26	3.07	0.61	4.29	6.42
	(0.852)	(0.184)	(0.910)	(0.380)	(0.186)
Income = £42-69k	2.72	1.90	0.17	4.85	6.55
	(0.693)	(0.442)	(0.976)	(0.334)	(0.194)
Primary school	26.98***	8.61***	19.91***	31.89***	28.42***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Secondary school	14.27***	8.58***	20.76***	17.74***	22.09***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Secondary education	11.96***	5.28**	14.77***	16.04***	17.36***
	(0.000)	(0.010)	(0.000)	(0.000)	(0.000)
University	12.56***	5.38***	13.30***	16.07***	17.09***
	(0.000)	(0.005)	(0.000)	(0.000)	(0.000)
Post-graduate degree	11.66***	6.57***	15.45***	17.69***	17.43***
	(0.000)	(0.003)	(0.000)	(0.000)	(0.000)
Uses Amazon	1.46	-1.95	2.36	-0.90	2.18
	(0.675)	(0.507)	(0.553)	(0.788)	(0.711)
Daily internet use: <2hrs	1.76	9.13***	1.28	5.55	12.37***
	(0.567)	(0.001)	(0.710)	(0.112)	(0.000)
Daily internet use: 2 to 5hrs	1.31	10.61***	3.09	6.38**	12.26***
	(0.620)	(0.000)	(0.320)	(0.040)	(0.000)
Daily internet use: >5hrs	1.54	9.94***	2.73	5.57*	11.99***
	(0.539)	(0.000)	(0.375)	(0.070)	(0.000)
Shops online >once per month	-2.25	-1.98	-1.88	-3.72**	-2.33
	(0.125)	(0.119)	(0.275)	(0.014)	(0.169)
Shops on Amazon >once per month	1.86	1.97**	2.27	1.38	0.79
	(0.125)	(0.025)	(0.101)	(0.167)	(0.533)
Constant	11.89	-8.79	-5.22	-4.79	-16.52*
	(0.177)	(0.119)	(0.529)	(0.540)	(0.063)
Observations	354	354	354	354	354

Table A8: Predicting WTP for dash cams

	(1) *Best Buy*	(2) *Don't Buy*	(3) Mediocre 1	(4) Mediocre 2	(5) Mediocre 3
Female	-0.62	-3.85*	-8.02***	-4.02*	-2.74
	(0.804)	(0.086)	(0.001)	(0.100)	(0.222)
Age = 25-34	-3.71	-2.56	-3.52	-1.63	-7.42
	(0.494)	(0.586)	(0.545)	(0.737)	(0.104)
Age = 35-44	2.43	0.97	-1.17	6.88	1.07
	(0.639)	(0.830)	(0.839)	(0.139)	(0.812)
Age = 45-54	-2.64	0.78	-2.47	-0.44	-1.96
	(0.635)	(0.873)	(0.660)	(0.926)	(0.684)
Age = 55-64	-1.60	-2.57	-9.63*	-1.76	-4.09
	(0.752)	(0.573)	(0.069)	(0.684)	(0.350)
Age = 65+	4.23	-3.59	-3.43	-0.37	-0.35
	(0.434)	(0.484)	(0.566)	(0.942)	(0.946)
Income <£21k	-7.09	-7.08	5.63	6.55	-11.89***
	(0.321)	(0.459)	(0.651)	(0.663)	(0.003)
Income = £21-41k	-0.82	-9.34	4.51	6.25	-10.04***
	(0.902)	(0.317)	(0.715)	(0.674)	(0.004)
Income = £42-69k	-4.85	-6.68	5.11	0.23	-14.02***
	(0.518)	(0.485)	(0.685)	(0.988)	(0.002)
Primary school	10.62	-7.88	8.13	-2.30	15.55*
	(0.165)	(0.556)	(0.388)	(0.782)	(0.074)
Secondary school	23.03***	2.37	27.27***	22.34***	24.85***
	(0.000)	(0.843)	(0.000)	(0.000)	(0.000)
Secondary education	17.90***	10.68	24.14***	21.51***	26.56***
	(0.000)	(0.344)	(0.000)	(0.000)	(0.000)
University	15.68***	4.28	19.84***	18.41***	22.98***
	(0.000)	(0.701)	(0.001)	(0.000)	(0.000)
Post-graduate degree	18.13***	4.33	18.12***	19.11***	23.51***
	(0.000)	(0.702)	(0.004)	(0.000)	(0.000)
Uses Amazon	3.08	-5.38	0.05	1.77	0.42
	(0.501)	(0.456)	(0.993)	(0.772)	(0.939)
Daily internet use: 2 to 5hrs	4.79	3.33	5.97	3.85	-0.05
	(0.172)	(0.295)	(0.110)	(0.321)	(0.987)
Daily internet use: >5hrs	3.79	-0.21	2.37	3.13	-1.41
	(0.302)	(0.947)	(0.528)	(0.428)	(0.673)
Shops online >once per month	-3.19	-2.07	-3.51	1.89	-5.42*
	(0.331)	(0.529)	(0.302)	(0.579)	(0.057)
Shops on Amazon >once per month	1.22	5.31**	4.50	-1.69	2.82
	(0.690)	(0.029)	(0.121)	(0.584)	(0.267)
Constant	36.26***	22.60	10.38	6.63	32.85***
	(0.000)	(0.170)	(0.489)	(0.687)	(0.000)
Observations	288	288	288	288	288

Table A9: Predicting WTP for cordless vacuum cleaners

	(1)	(2)	(3)	(4)	(5)
	Best Buy	*Don't Buy*	Mediocre 1	Mediocre 2	Mediocre 3
Female	10.04	5.83	7.77	12.69*	5.65
	(0.129)	(0.256)	(0.205)	(0.062)	(0.361)
Age = 25-34	14.98	18.03*	26.35**	29.28	4.31
	(0.381)	(0.097)	(0.029)	(0.119)	(0.776)
Age = 35-44	19.29	13.21	39.51***	46.38**	12.64
	(0.245)	(0.173)	(0.001)	(0.010)	(0.390)
Age = 45-54	26.04	25.23**	42.01***	51.10***	21.73
	(0.113)	(0.023)	(0.000)	(0.004)	(0.138)
Age = 55-64	17.49	7.39	26.17**	34.20**	11.75
	(0.266)	(0.442)	(0.021)	(0.050)	(0.408)
Age = 65+	3.82	2.40	22.17*	9.20	-5.96
	(0.820)	(0.809)	(0.067)	(0.610)	(0.681)
Income <£21k	4.16	-13.18	2.84	9.63	7.06
	(0.787)	(0.253)	(0.834)	(0.646)	(0.689)
Income = £21-41k	10.94	-6.77	5.28	23.19	15.78
	(0.476)	(0.559)	(0.695)	(0.262)	(0.367)
Income = £42-69k	-4.47	-9.29	-6.41	11.14	12.62
	(0.799)	(0.470)	(0.680)	(0.620)	(0.512)
Secondary school	-42.38**	19.25	4.15	-11.15	8.77
	(0.047)	(0.266)	(0.840)	(0.626)	(0.665)
Secondary education	-52.31***	4.40	-11.40	-5.83	12.68
	(0.008)	(0.763)	(0.546)	(0.783)	(0.515)
University	-45.01**	6.02	-8.87	-18.85	6.26
	(0.025)	(0.678)	(0.632)	(0.377)	(0.748)
Post-graduate degree	-52.37***	2.59	-22.58	-28.51	-3.41
	(0.004)	(0.849)	(0.188)	(0.138)	(0.847)
Uses Amazon	-35.37**	-24.32	-13.68	-44.35***	-30.17**
	(0.025)	(0.232)	(0.554)	(0.008)	(0.049)
Daily internet use: 2 to 5hrs	80.53***	22.22***	78.03***	106.30***	68.77***
	(0.000)	(0.008)	(0.000)	(0.000)	(0.000)
Daily internet use: >5hrs	61.78***	16.00*	67.71***	95.24***	59.52***
	(0.000)	(0.081)	(0.000)	(0.000)	(0.000)
Shops online >once per month	6.47	-3.09	7.11	0.55	-4.47
	(0.500)	(0.709)	(0.475)	(0.961)	(0.653)
Shops on Amazon >once per month	1.81	10.31*	-3.69	9.74	15.48**
	(0.807)	(0.075)	(0.618)	(0.236)	(0.043)
Constant	106.57***	24.79	-7.64	15.79	13.59
	(0.000)	(0.316)	(0.788)	(0.621)	(0.632)
Observations	356	356	356	356	356

G Supplementary analyses

Table A10: Heterogeneity analysis (1)

	(1) Chose *Don't Buy*	(2) Chose Mediocre	(3) Chose *Best Buy*
Group 2	0.056***	-0.032	-0.024
	(0.017)	(0.025)	(0.022)
Group 3	0.126***	-0.076***	-0.050**
	(0.019)	(0.025)	(0.022)
Group 4	0.097***	-0.051**	-0.046**
	(0.018)	(0.024)	(0.022)
Group 5	0.112***	-0.085***	-0.026
	(0.019)	(0.025)	(0.022)
College degree	-0.016	-0.002	0.018
	(0.015)	(0.024)	(0.022)
Group 2 # College degree	0.003	-0.016	0.014
	(0.024)	(0.034)	(0.031)
Group 3 # College degree	-0.000	0.004	-0.004
	(0.026)	(0.034)	(0.030)
Group 4 # College degree	0.025	-0.030	0.004
	(0.025)	(0.034)	(0.031)
Group 5 # College degree	0.059**	-0.037	-0.022
	(0.026)	(0.034)	(0.031)
Constant	0.113***	0.614***	0.273***
	(0.011)	(0.017)	(0.016)
Observations	8296	8296	8296
R-squared	0.018	0.006	0.002

Table A11: Heterogeneity analysis (2)

	(1) Chose *Don't Buy*	(2) Chose Mediocre	(3) Chose *Best Buy*
Group 2	0.059***	-0.025	-0.034
	(0.020)	(0.028)	(0.025)
Group 3	0.124***	-0.069**	-0.055**
	(0.022)	(0.029)	(0.025)
Group 4	0.119***	-0.062**	-0.056**
	(0.021)	(0.028)	(0.025)
Group 5	0.130***	-0.118***	-0.012
	(0.021)	(0.028)	(0.025)
Income: £21k to £41k	-0.024	-0.010	0.034
	(0.018)	(0.029)	(0.027)
Income: £42k to £69k	-0.013	-0.004	0.017
	(0.023)	(0.037)	(0.034)
Income: >£69k	0.073*	0.009	-0.082*
	(0.041)	(0.053)	(0.044)
Group 2 # Income: £21k to £41k	0.004	0.002	-0.005
	(0.028)	(0.041)	(0.037)
Group 2 # Income: £42k to £69k	0.030	-0.056	0.026
	(0.038)	(0.054)	(0.049)
Group 2 # Income: >£69k	-0.046	-0.085	0.131**
	(0.031)	(0.041)	(0.037)
Group 3 # Income: £21k to £41k	0.028	0.007	-0.035
	(0.031)	(0.041)	(0.037)
Group 3 # Income: £42k to £69k	0.013	-0.032	0.019
	(0.039)	(0.052)	(0.047)
Group 3 # Income: >£69k	-0.115*	-0.040	0.155**
	(0.029)	(0.041)	(0.037)
Group 4 # Income: £21k to £41k	-0.019	0.001	0.018
	(0.029)	(0.041)	(0.037)
Group 4 # Income: £42k to £69k	0.014	-0.036	0.022
	(0.041)	(0.054)	(0.049)
Group 4 # Income: >£69k	-0.109*	0.011	0.097
	(0.030)	(0.041)	(0.036)
Group 5 # Income: £21k to £41k	0.007	0.074*	-0.081**
	(0.030)	(0.041)	(0.036)
Group 5 # Income: £42k to £69k	0.025	-0.021	-0.004
	(0.040)	(0.053)	(0.048)
Group 5 # Income: >£69k	0.041	-0.140*	0.099
	(0.066)	(0.075)	(0.066)
Constant	0.111***	0.613***	0.276***
	(0.013)	(0.020)	(0.018)
Observations	7572	7572	7572
R-squared	0.020	0.008	0.005

Table A12: Heterogeneity analysis (3)

	(1) Chose *Don't Buy*	(2) Chose Mediocre	(3) Chose *Best Buy*
Group 2	0.028	-0.033	0.005
	(0.019)	(0.027)	(0.024)
Group 3	0.071***	-0.048*	-0.023
	(0.020)	(0.027)	(0.024)
Group 4	0.079***	-0.048*	-0.030
	(0.020)	(0.027)	(0.023)
Group 5	0.091***	-0.046*	-0.046**
	(0.020)	(0.027)	(0.023)
Uses Amazon > once/month	-0.033**	-0.015	0.048**
	(0.016)	(0.024)	(0.022)
Group 2 # Uses Amazon > once/month	0.051**	-0.020	-0.031
	(0.024)	(0.035)	(0.031)
Group 3 # Uses Amazon > once/month	0.092***	-0.046	-0.046
	(0.026)	(0.035)	(0.031)
Group 4 # Uses Amazon > once/month	0.053**	-0.038	-0.016
	(0.026)	(0.035)	(0.031)
Group 5 # Uses Amazon > once/month	0.085***	-0.102***	0.017
	(0.026)	(0.035)	(0.031)
Constant	0.125***	0.622***	0.253***
	(0.013)	(0.019)	(0.017)
Observations	8225	8225	8225
R-squared	0.020	0.009	0.004

Table A13: Heterogeneity with controls

	(1) Chose *Don't Buy*	(2) Chose Mediocre	(3) Chose *Best Buy*	(4) WTP (predicted)	(5) WTP (average)
Does not trust reviews	-0.0268*	0.0212	0.00563	3.059	3.333*
	(0.0157)	(0.0227)	(0.0205)	(1.986)	(1.962)
Group 2	0.0284**	-0.0324*	0.00405	-1.402	-1.060
(G1 + fake stars)	(0.0136)	(0.0178)	(0.0158)	(1.572)	(1.502)
Group 2 × no trust	-0.0334	0.0410	-0.00753	0.154	-2.161
	(0.0267)	(0.0390)	(0.0350)	(3.473)	(3.406)
Group 3	0.102***	-0.0719***	-0.0301**	-5.901***	-5.481***
(G2 + fake reviews)	(0.0149)	(0.0179)	(0.0153)	(1.554)	(1.471)
Group 3 × no trust	-0.0586**	0.0550	0.00365	0.936	-0.242
	(0.0296)	(0.0392)	(0.0342)	(3.429)	(3.371)
Group 4	0.0668***	-0.0606***	-0.00618	-3.888**	-3.591**
(G2 + sloppy reviews)	(0.0145)	(0.0180)	(0.0158)	(1.592)	(1.506)
Group 4 × no trust	0.00721	0.0540	-0.0612*	-2.508	-3.519
	(0.0301)	(0.0388)	(0.0330)	(3.317)	(3.255)
Group 5	0.111***	-0.102***	-0.00887	-6.645***	-6.275***
(G3 + endorsement)	(0.0147)	(0.0175)	(0.0153)	(1.514)	(1.434)
Group 5 × no trust	-0.0750**	0.0765*	-0.00150	3.661	2.938
	(0.0293)	(0.0393)	(0.0348)	(3.398)	(3.348)
Constant	0.162***	0.598***	0.240***	33.88***	50.34***
	(0.0229)	(0.0283)	(0.0246)	(2.230)	(2.335)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	8,723	8,723	8,723	8,677	8,723
R^2	0.020	0.021	0.021	0.099	0.059

Table A14: Treatment effects (headphones)

	(1) Chose *Don't Buy*	(2) Chose Mediocre	(3) Chose *Best Buy*	(4) WTP (predicted)	(5) WTP (average)
Group 2	0.039**	0.000	-0.039	-0.676*	-0.683**
	(0.017)	(0.027)	(0.027)	(0.396)	(0.348)
Group 3	0.108***	-0.071***	-0.037	-1.381***	-1.372***
	(0.019)	(0.027)	(0.027)	(0.415)	(0.367)
Group 4	0.087***	-0.036	-0.052*	-1.185***	-1.285***
	(0.019)	(0.027)	(0.027)	(0.412)	(0.359)
Group 5	0.114***	-0.106***	-0.008	-1.269***	-1.137***
	(0.019)	(0.027)	(0.027)	(0.423)	(0.370)
Group 6	0.078***	-0.073***	-0.005	-0.801*	-0.762**
	(0.018)	(0.027)	(0.027)	(0.410)	(0.363)
Constant	0.091***	0.492***	0.416***	20.041***	19.211***
	(0.011)	(0.020)	(0.019)	(0.276)	(0.241)
Controls	No	No	No	No	No
Observations	4014	4014	4014	3557	4014
R-squared	0.012	0.006	0.002	0.004	0.005

Table A15: Treatment effects (dash cams)

	(1) Chose *Don't Buy*	(2) Chose Mediocre	(3) Chose *Best Buy*	(4) WTP (predicted)	(5) WTP (average)
Group 2	-0.005	-0.028	0.033	0.094	0.659
	(0.029)	(0.039)	(0.034)	(1.188)	(0.991)
Group 3	0.044	-0.034	-0.010	-2.201*	-1.052
	(0.031)	(0.039)	(0.033)	(1.211)	(1.015)
Group 4	0.028	-0.046	0.018	-0.381	-0.263
	(0.030)	(0.039)	(0.033)	(1.224)	(0.997)
Group 5	0.080***	-0.073*	-0.008	-1.927	-1.680*
	(0.031)	(0.039)	(0.033)	(1.198)	(1.018)
Group 6	0.020	-0.012	-0.008	-0.290	-0.097
	(0.030)	(0.039)	(0.033)	(1.183)	(0.993)
Constant	0.156***	0.621***	0.223***	37.501***	35.758***
	(0.021)	(0.027)	(0.024)	(0.831)	(0.693)
Controls	No	No	No	No	No
Observations	1910	1910	1910	1702	1910
R-squared	0.006	0.002	0.001	0.004	0.004

Table A16: Treatment effects (cordless vacuum cleaners)

	(1) Chose *Don't Buy*	(2) Chose Mediocre	(3) Chose *Best Buy*	(4) WTP (predicted)	(5) WTP (average)
Group 2	0.105***	-0.087***	-0.018	-6.448***	-6.562***
	(0.019)	(0.025)	(0.020)	(2.209)	(1.731)
Group 3	0.181***	-0.091***	-0.091***	-13.395***	-13.657***
	(0.021)	(0.025)	(0.018)	(2.289)	(1.803)
Group 4	0.174***	-0.106***	-0.068***	-12.370***	-13.295***
	(0.021)	(0.025)	(0.019)	(2.298)	(1.801)
Group 5	0.200***	-0.120***	-0.080***	-14.109***	-14.995***
	(0.021)	(0.025)	(0.019)	(2.276)	(1.810)
Group 6	0.120***	-0.022	-0.099***	-9.327***	-9.554***
	(0.019)	(0.025)	(0.018)	(2.217)	(1.722)
Constant	0.095***	0.725***	0.180***	105.955***	103.983***
	(0.011)	(0.017)	(0.015)	(1.469)	(1.105)
Controls	No	No	No	No	No
Observations	4064	4064	4064	3648	4064
R-squared	0.026	0.009	0.013	0.014	0.021

Table A17: Heterogeneous effects of the education intervention

Interaction with...	Effects
Dummies for age brackets	No significant interactive effects
Dummies for educational attainment	No significant interactive effects
Frequency of Amazon use	No significant interactive effects
Dummies for income ranges	No significant interactive effects
Trusts reviews	No significant interactive effects

Table A18: Mechanisms (1)

	(1)	(2)	(3)	(4)	(5)	(6)
	Based product choice on…					
	Star ratings	# reviews	Review content	Look of prod.	Prod. descr.	Brand
Group 2	-0.0108	0.00451	-0.0128	-0.00714	-0.0313*	-0.00757
(G1 + fake stars)	(0.0171)	(0.0158)	(0.0172)	(0.0171)	(0.0172)	(0.0170)
Group 3	-0.00156	0.0123	-0.0252	-0.00278	-0.00207	0.00207
(G2 + fake reviews)	(0.0172)	(0.0159)	(0.0172)	(0.0172)	(0.0172)	(0.0171)
Group 4	-0.00285	-0.00517	-0.0168	-0.00333	-0.00626	0.00465
(G2 + sloppy reviews)	(0.0172)	(0.0157)	(0.0172)	(0.0171)	(0.0171)	(0.0170)
Group 5	0.0112	-0.00118	-0.00282	-0.00396	-0.0226	-0.0124
(G3 + endorsement)	(0.0171)	(0.0157)	(0.0172)	(0.0170)	(0.0171)	(0.0169)
Constant	0.424***	0.289***	0.439***	0.416***	0.585***	0.401***
	(0.0122)	(0.0112)	(0.0122)	(0.0121)	(0.0121)	(0.0121)
Observations	8,326	8,326	8,326	8,326	8,326	8,326
R^2	0.000	0.000	0.000	0.000	0.001	0.000

Table A19: Mechanisms (2)

	(1) Trust rev. in exp.	(2) Duration (minutes)	(3) Easy to choose	(4) Confident in choice	(5) More effort IRL
Group 2	-0.0157	0.331	-0.000439	-0.0250	-0.0245
(G1 + fake stars)	(0.0161)	(2.387)	(0.0156)	(0.0165)	(0.0165)
Group 3	-0.0190	1.924	0.0223	0.00648	-0.00707
(G2 + fake reviews)	(0.0162)	(2.941)	(0.0154)	(0.0164)	(0.0167)
Group 4	-0.0287*	4.579	-0.0111	-0.00380	-0.0175
(G2 + sloppy reviews)	(0.0162)	(4.471)	(0.0156)	(0.0164)	(0.0166)
Group 5	-0.00659	-1.144	0.00855	-0.00179	-0.0137
(G3 + endorsement)	(0.0160)	(2.058)	(0.0154)	(0.0163)	(0.0165)
Constant	0.692***	10.69***	0.723***	0.667***	0.358***
	(0.0114)	(1.831)	(0.0110)	(0.0116)	(0.0118)
Observations	8,326	8,326	8,326	8,326	8,326
R^2	0.000	0.000	0.001	0.001	0.000

	(6) Buy prod. again	(7) Easy spot fake rev.	(8) Trust rev. on Amz	(9) Read prod. rev.	(10) Searched online
Group 2	-0.00539	-0.0143	-0.00703	-0.0269*	0.000560
(G1 + fake stars)	(0.0152)	(0.0159)	(0.0146)	(0.0140)	(0.0159)
Group 3	-0.0217	0.00504	-0.00985	-0.0224	0.00943
(G2 + fake reviews)	(0.0151)	(0.0161)	(0.0147)	-0.014	(0.0160)
Group 4	-0.00600	-0.00842	-0.0190	-0.0381***	0.00967
(G2 + sloppy reviews)	(0.0152)	(0.0160)	(0.0147)	(0.0141)	(0.0160)
Group 5	-0.0194	0.0137	0.00482	-0.0131	-0.00160
(G3 + endorsement)	(0.0150)	(0.0161)	(0.0144)	(0.0138)	(0.0158)
Constant	0.262***	0.308***	0.774***	0.809***	0.298***
	(0.0108)	(0.0114)	(0.0103)	(0.00969)	(0.0113)
Observations	8,326	8,326	8,326	8,317	8,316
R^2	0.000	0.000	0.000	0.001	0.000

Table A20: Mechanisms (3)

	(1)	(2)	(3)	(4)	(5)	(6)
	Based product choice on…					
	Star ratings	# reviews	Review content	Look of prod.	Prod. descr.	Brand
Group 6	-0.0400**	-0.0145	-0.0219	0.00587	0.0105	0.0403**
(G5 + education)	(0.0170)	(0.0155)	(0.0171)	(0.0170)	(0.0171)	(0.0170)
Constant	0.435***	0.288***	0.437***	0.412***	0.562***	0.389***
	(0.0121)	(0.0110)	(0.0121)	(0.0120)	(0.0121)	(0.0119)
Observations	3,355	3,355	3,355	3,355	3,355	3,355
R^2	0.002	0.000	0.000	0.000	0.000	0.002

Table A21: Mechanisms (3)

	(1)	(2)	(3)	(4)	(5)
	Trust rev. in exp.	Duration (minutes)	Easy to choose	Confident in choice	More effort IRL
Group 6	-0.0155	0.193	-0.0112	-0.0233	0.00161
(G5 + education)	(0.0161)	(2.003)	(0.0154)	(0.0164)	(0.0164)
Constant	0.686***	9.543***	0.731***	0.666***	0.344***
	(0.0113)	(0.940)	(0.0108)	(0.0115)	(0.0116)
Observations	3,355	3,355	3,354	3,354	3,355
R^2	0.000	0.000	0.000	0.001	0.000

	(6)	(7)	(8)	(9)	(10)
	Buy prod. again	Easy spot fake rev.	Trust rev. on Amz	Read prod. rev.	Searched online
Group 6	0.0238	-0.0169	-0.0216	0.00124	0.0211
(G5 + education)	(0.0150)	(0.0160)	(0.0146)	(0.0139)	(0.0159)
Constant	0.243***	0.321***	0.779***	0.796***	0.297***
	(0.0104)	(0.0114)	(0.0101)	(0.00980)	(0.0111)
Observations	3,355	3,355	3,355	3,352	3,352
R^2	0.001	0.000	0.001	0.000	0.001

H Survey questions - Main experiment

H.1 Introduction

Welcome and thanks for participating!

This is a study about purchasing habits on Amazon.co.uk.

In the study, you will be asked to choose a product category that you are interested in. You will then be shown a list of products, and will be asked to select the product that you would most like to receive.

All respondents that complete the survey will be entered into a prize draw to receive the product that they choose. Ten respondents will be randomly selected as winners at the end of the survey, and will be prompted to enter their address in order to claim their prize. Winners will receive their prizes within 6 weeks of completing the survey. The prize draw is completely random, and the distribution of prizes is administered by The Behaviouralist Ltd (contact: info@thebehaviouralist.com).

The survey should take around 5-10 minutes to complete.

The survey must be completed before 13 March 2020 in order to be eligible for the prize draw.

By clicking the button below, you acknowledge that your participation in the study is voluntary, that you are at least 18 years of age, and that you are aware that you can end your participation in the study at any time and for any reason.

H.2 Demographics

To start, please answer the following questions about yourself.

What is your age?

- 18 - 24
- 25 - 34
- 35 - 44
- 45 - 54

- 55 - 64
- 65 - 74
- 75 - 84
- 85 or older

What is your gender?

- Male
- Female
- Other

Where do you live?

- East of England
- East Midlands
- Greater London
- North East England
- North West England
- Northern Ireland
- Scotland
- South East England
- South West England
- Wales
- West Midlands
- Yorkshire and the Humber

What is your yearly income?

- Up to £7,000
- £7,001 - £14,000

- £14,001 - £21,000
- £21,001 - £28,000
- £28,001 - £34,000
- £34,001 - £41,000
- £41,001 - £48,000
- £48,001 - £55,000
- £55,001 - £62,000
- £62,001 - £69,000
- £69,001 - £76,000
- £76,001 - £83,000
- £83,001+
- Prefer not to say

What is your highest level of education?

- Primary school
- Secondary school up to 16 years
- Higher or secondary or further education (A-levels, BTEC etc)
- College or university
- Post-graduate degree
- Prefer not to say

On average, how much time, if any, do you spend online every day? By this, we mean using the internet or an internet-enabled device, such as a computer, smartphone or tablet.

- Less than 1 hour
- 1 – 2 hours
- 2 – 3 hours
- 3 – 5 hours
- More than 5 hours

- Don’t know

How often do you purchase items online?

- Most days
- About once or twice per week
- About once or twice per month
- About once every three months
- About once every six months
- About once every year
- Less often than once a year
- Never
- Don’t know

How often do you purchase items on Amazon.co.uk?

- Most days
- About once or twice per week
- About once or twice per month
- About once every three months
- About once every six months
- About once every year
- Less often than once a year
- Never
- Don’t know

H.3 Product category choice

We will now begin the study.

Please choose the product category you are most interested in.

- Headphones
- Dash Cams
- Cordless Vacuum Cleaners

H.4 Product choice

We would like you to decide what [product category] you would want to buy from Amazon.co.uk. Please do this exercise as you would in real-life.

If you are having trouble reading the text in the images you can zoom in by pressing COMMAND + "+" on a Mac, and CTRL + "+" on a PC.

Click next to begin

Please consider the products shown in the image. You can learn more about the products by selecting them in the question section below the image. You are able to view more than one product before you make your choice.

[An image of product search page shown here]

Select a product below to learn more about it

- Product 1
- Product 2
- Product 3
- Product 4
- Product 5

Explore the product page. Below you can choose to select this product or to go back and continue shopping.

[An image of product page shown here]

Do you want to select this product? [if you want to continue shopping click on the back button]

- Select Product

H.5 Post-experiment questions

You are almost there! Just a few more questions about your experience.

How confident are you that you chose the best product for you?

- Extremely confident
- Very confident
- Moderately confident
- Slightly confident
- Not confident at all

How easy or difficult was it to assess the products' value for money?

- Extremely easy
- Somewhat easy
- Neither easy nor difficult
- Somewhat difficult
- Extremely difficult

Thinking about the effort you put into choosing the [product] you selected, how does this compare to the amount of effort you would put into choosing and purchasing [product] in real life?

- I would put a LOT MORE effort into choosing in real life
- I would put a little more effort into choosing in real life

- I would put the same amount of effort into choosing in real life
- I would put a little less effort into choosing in real life
- I would put a LOT LESS effort into choosing in real life
- Don't know

Imagine that you were given [£24.99/£49.99/£149.99] and asked to buy [a pair of headphones/a dash cam/a cordless vacuum cleaner]. You would have free access to any online retail platform, such as Amazon.co.uk or Argos. Do you think that you would have chosen a different product to the one that you just chose?

- Yes
- No
- Don't know

What did you base your choice on? (choose all that apply)

- The star rating
- The number of reviews
- The content of the reviews
- Information in the product description
- Brand
- The look of the product
- Other

Did you read any customer reviews when choosing a product?

- Yes
- No

[If yes] To what extent did you trust the customer reviews of the product that you selected?

- A great deal
- A fair amount

- Not very much
- Not at all
- Don't know

Did you search online for more information about the products?

- Yes
- No

[If yes] Where did you search for more information? (choose all that apply)

- Search engine
- Expert review site
- Retailer website
- Manufacturer website
- Amazon.co.uk
- Other

To what extent do you generally trust customer reviews on Amazon.co.uk?

- A great deal
- A fair amount
- Not very much
- Not at all
- Don't know
- I don't shop on Amazon.co.uk

How easy or hard is it to tell if a customer review on Amazon.co.uk is fake?

- Extremely easy
- Somewhat easy
- Neither easy nor difficult

- Somewhat difficult
- Extremely difficult
- I don't shop on Amazon.co.uk

What share of reviews on Amazon.co.uk do you believe are fake?

[Choose on a scale of 0 - 100]

Do you typically do any of the following when buying products on Amazon.co.uk? (choose all that apply)

- Read customer reviews
- Look at the profiles of customers who have left reviews on the product
- Find and read the most critical customer reviews (*e.g.*, 1-star reviews)
- Look at the average star rating
- Look at the distribution of star ratings (*e.g.* how many gave 5-stars compared to 1-stars)
- Look at the number of reviews
- Look at the date of reviews
- Use online tools to help determine the quality of reviews
- None of the above
- I don't shop on Amazon.co.uk

In the last 12 months, have you bought or searched online for [product category]?

- Yes
- No
- Don't know

[If respondent wins the prize draw]

Congratulations! You have been selected to receive the prize.

Please provide your name and email address so we can contact you with further details to claim your prize.

Your address information will only be used for the purpose of shipping your reward. We may contact you through your email if we run into issues with the shipment of the prize.

- Full Name
- Email Address
- Address line 1
- Address line 2
- City
- County
- Post Code

End of Survey

I Survey questions - Eliciting willingness to pay

I.1 Introduction

Welcome and thanks for participating!

This short survey is being conducted by researchers at The Behaviouralist to gain a better understanding of the prices consumers are willing to pay for different products. The survey should take around 5-10 minutes to complete.

Please note that by clicking the button below to proceed, you acknowledge that your participation in the study is voluntary, that you are at least 18 years of age, and that you are aware that you can end your participation in the study at any time and for any reason.

We will start by explaining how the exercise in this survey works. Please go through the explanations carefully as your responses will be considered invalid if you do not understand how the exercise works.

We will then give you a short exercise for practice and ask you a few comprehension questions to make sure that you understand the instructions. We will not tell you if you passed the comprehension question, so it is essential that you pay attention.

Once you have finished the comprehension questions, the actual survey will then begin. You will first be asked to choose a product category that you are interested in.

You will then be shown a table containing information about five products within your chosen product category. Please take your time to look at all five products and the information about their features, ratings, and reviews.

We will then ask you how much you are be willing to pay for each of the five products. We will do this by listing a number of prices and asking what the highest price is that you are willing to pay.

All respondents who complete the survey will be entered into a prize draw for £300. Winners will be randomly selected at the end of the survey, and the payment will be made via the Prolific platform.

If you win the prize draw, we will randomly select one of the five products that you were shown. We will then randomly choose a price point and will implement your choice for that price point.

Please see below for an example of this works:

Imagine that we randomly select the Bose Headphones. We then go on to randomly select a price point of £25 for that product. Then:

1. If you said that you are willing to pay £25 or more for Bose Headphones, we will send you Bose Headphones and the remainder of your prize money (which is £275).

2. If you said that you are not willing to pay £25 or more for Bose Headphones, we will only send you the prize money (which is £300).

It is therefore in your best interest to answer as honestly as you can about the price that you are willing to pay for each of the five products.

I.2 Comprehension questions

Before we proceed with the survey, we will ask you to complete a practice exercise to make sure you fully understand the instructions. We may exclude your responses from our analysis if you fail to answer the practice questions correctly.

The table below shows two bluetooth speakers and their product features and other details. Please go through the information and answer the questions below.

Figure A16: Table for comprehension questions

Product image		
Product name	JBL Charge 4 Bluetooth Wireless Speaker	Roberts Beacon 310 Bluetooth Wireless Speaker
Which? recommendation	BEST BUY	Not Applicable
Which? test score	77%	54%
Brand	JBL	Roberts Radio
Colour	Black	Black
Connectivity technology	Bluetooth	Bluetooth
Battery life	19 hours and 15 minutes	12 hours
Product weight	1kg	0.71kg
Special feature	IPX7 waterproof, portable	Portable

Comprehension Question 1: How much is the JBL Charge 4 bluetooth speaker worth to you (in £)?

Comprehension Question 2: What is the highest amount you would pay for the JBL Charge 4 bluetooth speaker?

For example, if the product is worth £100 to you, you should choose that your highest willingness to pay is £85.

Comprehension Question 3: How much is the Roberts Beacon 310 bluetooth speaker worth to you (in £)?

Comprehension Question 4: What is the highest amount you would pay for the Roberts Beacon 310 bluetooth speaker?

Even if you prefer one product to the other, you should still tell us how much you are willing to pay for both, as we will randomly select ONE product and ONE price point for that product.

Comprehension Question 5: Imagine that you won the prize draw (for £300) and we randomly picked the JBL Charge 4 bluetooth speaker. We also randomly picked the price point of £60. Based on the choices that you made above, what will you receive?

Thank you for your responses to the practice questions!

I.3 Product category choice

We will now begin the study.

Please choose the product category you are most interested in.

- Headphones
- Dash Cams
- Cordless Vacuum Cleaners

I.4 Willingness to Pay

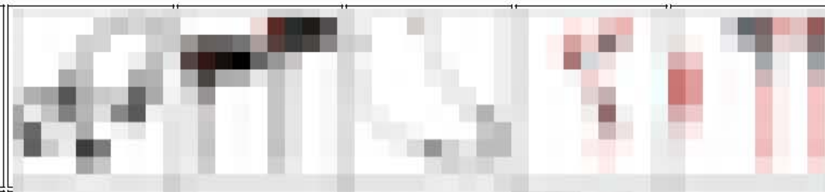
The table on the next page shows five [product category] and their product features and other details. Please go through the information carefully and answer the questions below the table.

To help you evaluate the products, the table includes information about product recommendations and test scores from Which?. Which? is a UK's consumer protection organisation that conducts rigorous testing, rates products, and publishes their recommendations and warnings on their website.

Which? recommendation shows whether a product is labelled 'Best Buy' or 'Don't Buy' by Which?. 'Best Buy' products are the very best products that have satisfied or exceeded specific criteria based on the results of rigorous comparative tests and analysis carried out by Which?. 'Don't Buy' products are the very poor products that are a waste of money at best and a threat to safety at worst.

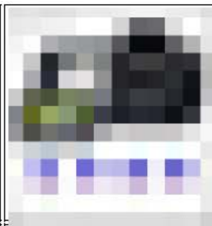
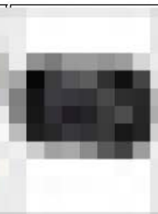
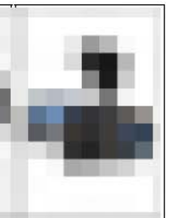
Which? test score is based on rigorous independent lab tests, which measure and assess each of the most important aspects of usability and performance. The results of the expert tests are combined to create a unique overall test score for the product.

Figure A17: Table for headphones

Product Image					
Product Name	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Which? recommendation	Not applicable	Not applicable	Not applicable	DON'T BUY	BEST BUY
Which? test score	Not tested	Not tested	Not tested	24%	71%
Brand	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Colour	Grey	Black	Black	Red	Red
Connectivity technology	Wireless	Wired	Wireless	Wireless	Wired
Noise control	Sound isolation	None	None	None	Active noise cancellation
Cable feature	Detachable	Tangle-free	Tangle-free	-	-
Product weight	36.3g	22.7g	18.1g	20g	19g
Batteries	1 Lithium ion batteries required. (included)	-	1 Lithium Metal batteries required. (included)	1 Lithium ion batteries required	-
Special features	Wireless	Inline Playback Controls, With microphone, Tangle-free Cord	Neckband, integrated remote control	Wireless	With microphone, Integrated remote control

Notes. The product names, brand names, and product photos are available upon request.

Figure A18: Table for dash cams

Product Image					
Product Name	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Which? recommendation	BEST BUY	Not applicable	Not applicable	DON'T BUY	Not applicable
Which? test score	74%	67%	Not tested	44%	Not tested
Brand	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Colour	Black	Black	Black	Black	Black
Resolution	1080p Full HD	1080p Full HD	1080p Full HD	1080p Full HD	4K
Viewing angle	140	150	170	125	Not available
Type of mount	Windscreen suction mount	Windscreen suction mount	Windscreen suction mount	Adhesive pad	Windscreen suction mount
Product dimensions	9.2 x 1.8 x 5.6 cm	3 x 10.5 x 4.7 cm	5.6 x 3 x 5.6 cm	5 x 6.3 x 2.3 cm	Not available
Batteries	-	1 Lithium ion batteries required. (included)	-	1 Lithium ion batteries required. (included)	1 Lithium ion batteries required

Notes. The product names, brand names, and product photos are available upon request.

Figure A19: Table for cordless vacuum cleaners

Product Image					
Product Name	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Which? recommendation	BEST BUY	Not applicable	Not applicable	DON'T BUY	Not applicable
Which? test score	72%	Not tested	Not tested	Not tested	49%
Brand	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Colour	Blue	Red	Blue	Red	Dark blue
Power/Wattage	350	250	300	100	120
Runtime	Up to 25 minutes	Up to 30 minutes	Up to 22 minutes	Not available	Up to 22 minutes
Noise level (dBA)	70	Not available	84	Not available	80
Product weight	2.28kg	5.7kg	4kg	3.28kg	2.4kg
Capacity	0.7 litres	6 litres	0.33 litres	Not available	Not available
Surface recommendations	Hard Floor, Carpet	Upholstery	Dual Action, Hard Floor	Hard Floor, Carpet, Walls, Curtains, Furniture, Ceilings	Hard Floor, Upholstery, Carpet

Notes. The product names, brand names, and product photos are available upon request.

What is the highest price you are willing to pay for these products?

Remember that we will randomly select ONE product and ONE price point, so it is important that you provide us with your willingness to pay for all five products.

[Participants are shown a list of ten prices for each of the five products according to the product category they have chosen.]

I.5 Demographics

Before we end the survey, please answer the following questions about yourself.

What is your age?

- 18 - 24
- 25 - 34
- 35 - 44
- 45 - 54
- 55 - 64
- 65 - 74
- 75 - 84
- 85 or older

What is your gender?

- Male
- Female
- Other

Where do you live?

- East of England
- East Midlands
- Greater London
- North East England
- North West England
- Northern Ireland

- Scotland
- South East England
- South West England
- Wales
- West Midlands
- Yorkshire and the Humber

What is your yearly income?

- Up to £7,000
- £7,001 - £14,000
- £14,001 - £21,000
- £21,001 - £28,000
- £28,001 - £34,000
- £34,001 - £41,000
- £41,001 - £48,000
- £48,001 - £55,000
- £55,001 - £62,000
- £62,001 - £69,000
- £69,001 - £76,000
- £76,001 - £83,000
- £83,001+
- Prefer not to say

What is your highest level of education?

- Primary school
- Secondary school up to 16 years
- Higher or secondary or further education (A-levels, BTEC etc)

- College or university
- Post-graduate degree
- Prefer not to say

On average, how much time, if any, do you spend online every day? By this, we mean using the internet or an internet-enabled device, such as a computer, smartphone or tablet.

- Less than 1 hour
- 1 – 2 hours
- 2 – 3 hours
- 3 – 5 hours
- More than 5 hours
- Don't know

How often do you purchase items online?

- Most days
- About once or twice per week
- About once or twice per month
- About once every three months
- About once every six months
- About once every year
- Less often than once a year
- Never
- Don't know

How often do you purchase items on Amazon.co.uk?

- Most days
- About once or twice per week

- About once or twice per month
- About once every three months
- About once every six months
- About once every year
- Less often than once a year
- Never
- Don't know

End of Survey