

# Deception in Online Auction Marketplaces: Incentives and Personality Shape Seller Honesty

Brian D. Glass, Michelle Baddeley, M. Angela Sasse  
University College London

## Abstract

In online auction marketplaces, item misrepresentation is one of the most common forms of seller deception. The impact of coarse-grained incentive manipulations on deceptive seller behavior in online markets had not been studied. We demonstrate experimental control over seller honesty, quantify this behavior relative to ground truth, and link it to personality. We recruited 62 experienced online auction sellers to handle seven jewelry items (some flawed) and generate auction advertisements. Independent variables included demographics, personality, and an incentive manipulation: maximize returns, avoid negative reputation, or a *Control* condition. Dependent variables involved the inclusion or exclusion of honest information. *Reputation* incentives led to significantly more reported flaws (30%) than other conditions (17%), all ‘hiding’ flaws in the middle of descriptions. More experience in selling items significantly predicted honesty, while extraversion and neuroticism predicted dishonesty. Design features can reduce seller deception and enhance buyer detection of deception: reputation systems should be highly salient to sellers, while buyers should consider images over text, and descriptions could be shuffled to expose honest information.

## 1 Introduction

Item misrepresentation is a major problem in the online auction marketplace sector: misrepresented items represent the 2nd most common buyer complaint on eBay (behind *paid for but not received*) [15]. On eBay alone, 3 million transactions are completed per day from 200 million active item listings [31], each of which represents an opportunity for fraud via item misrepresentation.

Our ultimate aim is to deter seller deception in ways that are low-effort for the user and low-cost for service providers. We draw on knowledge from design, human computer interaction, and behavioral economics. We

achieve this goal by incentivizing seller behavior at the moment of advertisement creation. With the availability of ground truth (rare in deception research [38]) regarding the items for sale, we can accurately quantify deceptive behavior in ways that completed listings data and surveys cannot. In finding ways to quantify seller behavior and the transient incentive frameworks which predict seller behavior, we can begin to develop design principles in order to deter these forms of deception. Given the pervasiveness of auction misrepresentation, reducing deception could have a real and immediate impact on the online economy.

In order to achieve deterrence, we seek to develop a novel procedure to quantify the complex behavior of creating advertisements for online auction marketplaces. The present study achieves this, offering insights into how incentives and product framing can impact deceptive practices in seller behavior. First, we consider the available literature on misrepresentation in online auction systems. Next, we describe our novel methodology for quantifying deceptive seller behavior. We then report statistical analyses which reveal the impact of incentives and characteristics of the individual which predict honest behavior. Finally, we consider the impact of our results for website design and basic research.

## 2 Background

### 2.1 Misrepresentation in the Marketplace

Chiu [7] reports that information quality has an impact on “procedural justice” (e.g., perceived fairness of the transaction policies adopted by the seller). Thus, finding ways to deter misrepresentation fraud could have an important systemic impact on an online auction site’s customers’ sense of trust in the site’s policies. The problem of misrepresentation has been explored in the economic literature on asymmetric information and adverse selection. Akerlof [1] describes “lemons” in car mar-

kets: when the quality of the product is unknown by a buyer, then the seller can exploit their informational advantage to charge the same price for a lemon as others would charge for a good car – a “plum”. Buyers adapt by paying a price that reflects average quality: the sellers of lemons get too much but the sellers of plums get too little. The problem is that the plum-sellers leave the market, quality falls, prices fall and the market may disappear altogether. In the real world, this basic insight explains why people don’t get fair prices for second-hand goods unless the quality is clearly observable. Standard economics would predict that all sellers would be equally likely to conceal flaws in their products and they would not discriminate amongst buyers. In terms of online auction marketplaces, this could mean that the concealment of flaws may not be moderated by buyer feedback or reputation systems. However, evidence suggesting the contrary is discussed below.

## 2.2 Avoiding Increased Buyer Effort

Prior research on the impact of reputation systems on online market places have been done using either survey techniques or customer transaction data sets collected online or provided by eBay Inc. or Yahoo! Inc. to the researchers. Zhang & Cohen [43] have proposed a social-network based reputation system, which clusters buyers and sellers by their reputation similarity and allows them to bid for reputable transaction partners. However, others provide evidence that dynamic reputation systems may promote retaliation, making anonymous reputation systems potentially more reliable [28]. On the other hand, reputation systems that are strongly bound to seller identity (i.e., difficult to change) have been found to be more effective at maintaining honesty [2]. We seek to develop novel methods for deterring deceptive behavior which require minimal change to currently instantiated reputation systems.

Buyer defense has also been proposed, such as insurance [35], or the use of 3rd party trust certificates [2]. Other high-cost mechanisms include performing an internet search for the seller’s ID or company name along with the word “fraud”, establishing lengthy email communication with the seller, and even meeting the seller in person [11]. Others have gone as far as to propose “vigilantism”, in which buyers actively sabotage the auctions of sellers known to be fraudulent [8]. These adversarial solutions are suboptimal for promoting trustworthiness in a socio-technical system because they represent unreliable signals, potential for deceptive mimicry, and systemic instability [30] [3] [32]. Instead, system design supportive of cooperative behavior leads to trustworthiness in the marketplace, beneficial for parties on both sides of the transaction [21] [22].

Moreover, since complex reputation systems, escrows, insurance, and pre-sale communication are all additional mechanisms which online marketplaces would have to actively adopt, the present work focuses on basic seller behavior at the moment of advertisement creation using simple incentive mechanisms. In this way, a fully experimental item description generation paradigm represents a novel contribution.

## 2.3 Reputation: Shaping Seller Behavior

A robust finding in the literature is that sellers with higher reputation also achieve higher sales prices. In studies analyzing completed eBay listings, sellers with better reputations have higher sales and obtain higher prices, at least for various items of electronics such as mobile phones, DVDs [9] and Pentium III processors [17]. Theoretical work has indicated that positive reputation systems may be more effective at promoting market efficiency than negative reputation systems [?]. However, positive-only reputation systems can make it difficult for new entrants into the market to find transaction partners, making it unattractive for online auction providers. Moreover, the vast majority of currently instantiated reputation systems use both positive and negative feedback, and so it is important to investigate the impact of reputation systems which allow both forms of buyer feedback.

One research question is whether people generate item descriptions that are more honest if the seller is seeking to achieve a high final auction price rather than seeking to avoid negative feedback. Seller reputation appears to have an asymmetrical impact on final buying price [34] in that buyers are sensitive to the presence of negative reviews despite being relatively insensitive to the number of positive reviews (although this may only be the case for established seller identities [29]). If one instructional method results in higher information quality (i.e., less misrepresentation) then it could be immediately relevant as a method for deterring online deception. The hypothesis could be that when the negative feedback mechanism of the reputation system is more salient, sellers will engage in fewer deceptive descriptive practices.

## 2.4 Quantifying Auction Behavior

### 2.4.1 Novel Methodology

In order to examine choice behavior at the moment of advertisement creation, we develop a procedure which involves handling actual items in person while using a novel online auction website. Moreover, participants achieve actual monetary rewards, not hypothetical rewards nor secondary rewards (e.g. raffle tickets). Participants construct advertisements from controlled elements

including text and images. A controlled proportion of description elements contain information regarding item flaws. In this way, we are able to quantify and describe task behavior at the point of advertisement creation using a novel and immersive procedure. Selling items on an online auction websites is a real-world task that is in many ways intractable due to the complexities involved. We seek to make this paradigm tractable by developing a controlled approach, which offers not only freedom and flexibility but also the considerable structure required to quantify aggregate behavior and test specific hypotheses.

Importantly, we seek to test the impact of a salient reputation mechanism on seller behavior at the moment of advertisement creation. To do so, we contrast a *Reputation* condition with other incentive-based conditions for which no form of reputation mechanism is visible or made explicitly available. In the non-reputation conditions, *Percentage Commission* (earn 10% of sale price) is tested against a *Flat-Rate Commission* (earn £12 regardless of sale price). Finally, there is a *Control* condition without an explicit commission or reputation mechanism.

#### 2.4.2 Importance of Experimental Design

A popular method for studying online auction systems is to acquire, or scrape, data from popular websites such as eBay. Publicly available data for completed auctions can include item sales price, bid history, and information about the seller's reputation. While this technique can result in large and readily accessible data sets, there are drawbacks. First, ground truth regarding the item for sale is inaccessible, making it difficult or impossible to quantify any item misrepresentation which may have occurred. Second, post-hoc methods are blind to the moment of advertisement creation. Beyond the items for sale, what pre-auction information is driving the seller to create his or her advertisement, and how and when does the seller decide whether to misrepresent an item?

The generalizability of laboratory-based experiments is a common criticism in experimental research, but we would point to strong positive relationships between laboratory and field studies in psychological research. Mitchell [26] found good correlations between the results and effect sizes in 217 laboratory-to-field comparisons. In the meta-analysis, subfields relevant to the present work enjoyed laboratory-to-field correlations of  $R=.89$  (Industrial-Organizational Psychology),  $.83$  (Personality Psychology), &  $.59$  (Consumer Psychology). Unlike observational "scraping" studies, lab-based studies offer true experimental manipulations via random assignment, and embed robust controlled conditions.

The present study design promotes realistic real-world behavior for two main reasons. First, the design ties

monetary rewards to performance. Second, we recruit genuine, active eBay sellers with above-average experience. According to a well-established psychological principle, repeated behavior leads to automaticity [19]. Similarly, in evolutionary game theory, social norms (including those relating to reputation) are internalized via learning and it takes time to "unlearn" these. Thus, experienced sellers will repeat the behaviors they have learned via repeated use of online auction markets. For all these reasons, our design promotes, manipulates, and captures realistic human behavior.

### 2.5 The Person Behind the Auction

Previous research on human deception has demonstrated that several personality and demographic characteristics can impact dishonest behavior. In the present study, personal attributes of the seller may interact with task components in critical ways. For example, prior work suggests the amount of experience of a seller does not correlate with the probability of fraud and deception [23]. This work was based on buyer feedback from completed auctions, and did not consider the content of the advertisements themselves - although it offers a strong prediction that seller experience may not correlate with misrepresentation in the current study. Experience might be linked with age, and younger people are more likely to self-report engaging in online deceptive practices [6]. However, the current study does not include a sufficient age range to assess the association of age with deceptive practices.

Pro-social behavior may be linked to the personal traits of agreeableness and empathy [14]. In online auction marketplaces, the economic perspective suggests pro-social behavior might manifest as the tendency to describe items accurately because reputation is salient for pro-social individuals. For this reason, we hypothesize agreeableness and empathy will correlate negatively with misrepresentation in the online advertisement creation task.

There are other key personality traits that have been associated with various forms of deception. Extraversion has been shown to correlate with deceptive behavior. Although at first glance this may seem counterintuitive, increased deception among extraverts may arise from not only an increased desire to appeal to others, but also due to an increased understanding of the nuances of social interactions [41]. In addition, prior work has linked extraversion to increased self-promotion tactics during job interviews [18]. In online auction marketplaces, extraversion might represent a desire to create advertisements which appeal to others, a goal which may require omitting honest information regarding product flaws.

Neuroticism has also been linked to deceptive behavior. In particular, neuroticism has been linked to disclosure deceptions which involve the act of regulating the disclosure of information to others [27]. In online behavior, neuroticism has been shown to predict presentation of the ideal and false self [25]. Thus, both extraversion and neuroticism may be associated with increased deceptive practices in online auction marketplace behavior. There is no prior literature to suggest conscientiousness may be associated with increased deceptive practices, although face validity suggests there may be a negative association with deception.

Intellectual ability has been shown to be a prerequisite for deceptive abilities. In work with primates, larger neocortex size is associated with increased rates of deception [5]. In humans, cognitive mechanisms such as executive attention which are associated with intellectual ability have been shown to be associated with deceptive actions [13], although not necessarily in deceptive ability [16]. Those with lower intelligence report lying takes more effort [39], and so it is possible that in the present study intellect will correlate positively with misrepresentation.

It is not known how the value of an item for sale may impact the tendency for a seller to misrepresent the item. In the study of deception, there is considerable interest in how behavior may differ between low-stakes and high-stakes scenarios [24]. Recent evidence suggests that the level of stakes may have no impact on verbal or non-verbal deceptive behavior [20], although such behavior may represent actions of dishonesty in the direct presence of others, while creating advertisements on a computer is done without the buyer directly present. Although there is some evidence that percentage-based commissions may lead seller agents to use suboptimal selling techniques in order to make quick sales [42], this concept of optimality may not map onto deceptive behavior in online auction marketplaces. The present study offers a unique insight into whether item value impacts deceptive practices in online auction marketplaces, although prior work suggests that item value may not drive dishonest behavior.

## 2.6 Hypotheses

- H1: Participants are most dishonest for the non-reputation commission conditions (*Flat-Rate Commission* and *Percentage Commission*), and most honest for the *Reputation* condition.
- H2: Item values are positively associated with the level of misrepresentation (i.e., the amount of honest statements included in the generated descriptions).
- H3: Deception via images is more common in the non-

reputation commission conditions, and less common in the *Reputation* condition.

- H4: Misrepresentation will be positively associated with A) extraversion, B) neuroticism, and C) intellect/imagination.
- H5: Misrepresentation will be negatively associated with A) agreeableness, B) empathy, C) conscientiousness, and D) online auction experience

## 3 Methods

### 3.1 Design

#### 3.1.1 Experimental Manipulations

We tested four experimental conditions. In a *Control* condition, participants were told to generate realistic advertisements, and that they would be entered into a random drawing for a cash prize. In a *Flat-Rate Commission* condition, participants were told that they would receive £12 (\$20) for each item that sold. In a *Percentage Commission* condition, participants were told that they would receive 10% of the final sales price for each item that sold. In a *Reputation* condition, participants were told that they would receive 10% of the final sales price for each item that sold, so long as negative buyer feedback was not received for that item.

#### 3.1.2 Ecological Validity

We sought to maximize the ecological validity of our task paradigm in critical ways. First, we sought experienced online auction sellers. Only 20.2% of screener respondents qualified for the study. The median number of online auction sales listings in the previous year for all respondents was 3, with a lifetime median of 10. Among those who qualified, the median number of online auction sales listings in the previous year was 6, with a lifetime median of 50. This demonstrates that our sample included a very experienced group of online auction sellers. Having an experienced sample decreases initial task training, allowing participants to focus on the demands of the task. Moreover, it offers an enhanced insight into how actual online sellers may react to the incentive-based manipulations instantiated in the current study.

The task itself is designed to achieve greater ecological validity and depth of insight than other popular forms of studies. In post-hoc studies which analyze completed auction listings, little or nothing is known about the impact of incentives on the seller at the moment of advertisement creation. Moreover, these scraped studies can do little if anything to measure personality or demographic attributes of the sellers and buyers. In questionnaire based surveys of online auction sales behavior,

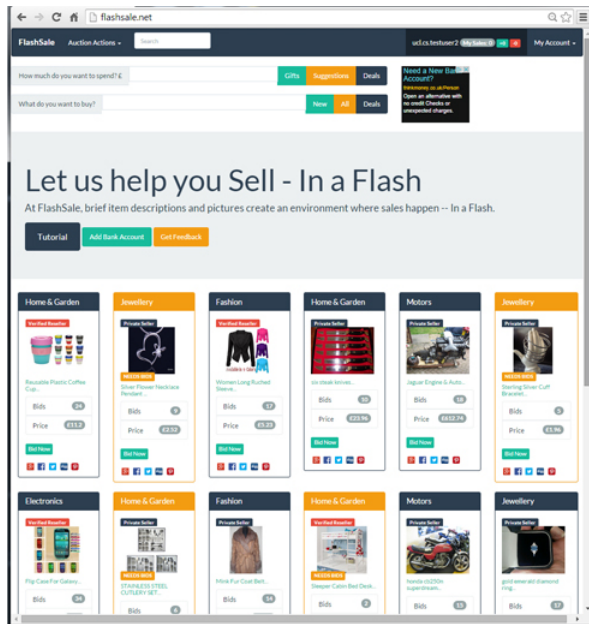


Figure 1a: Screen shot of “Flashsale.net” home page.

no information can be garnered to quantify how sellers actually behave. Additionally, true experimental manipulations cannot be done using scraped or provided data sets, questionnaires, or surveys. In the present study, we develop a laboratory based online auction environment which is lifelike and immersive, yet which allows true experimentation via random assignment. Participants view and handle actual sales items while creating advertisements on a realistic online auction website designed to mimic popular auction websites such as eBay.

### 3.2 Procedure

Participants completed an in-laboratory study which lasted approximately 1 hour (Mean = 57 minutes, SD = 13 minutes). Participants were randomly assigned to the four experimental conditions: *Control*, *Flat-Rate Commission*, *Percentage Commission*, and *Reputation Feedback*. Participants were briefed that they would be creating 7 advertisements on an online auction website called Flashsale.net (Figure 1a). Participants were told that the auction website was similar to eBay in that users post items for sale and other users can bid on those items in order to purchase them. They were instructed to freely click on two items for sale in order to view examples of completed advertisements. In the *Control* condition, participants were told to simply create advertisements, while in the other conditions participants were instructed that they would receive commission based on any sales that occurred within one week. In order to create each advertisement (see Figure 1b), participants selected six

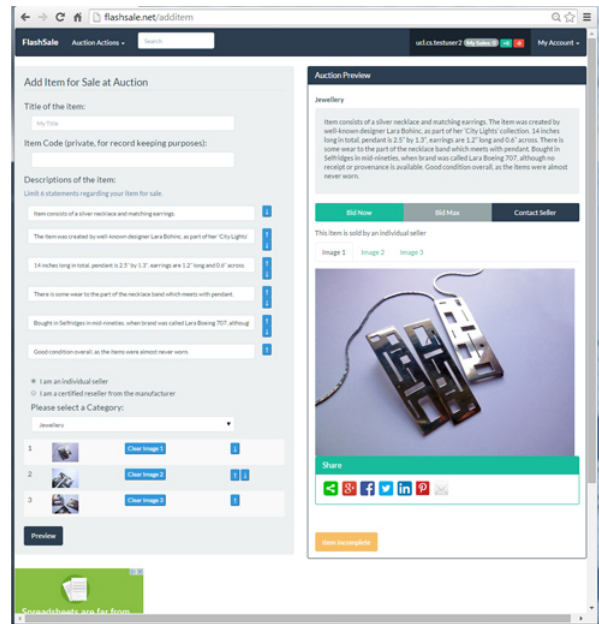


Figure 1b: Screen shot of advertisement creation page.

of 12 available pre-written descriptions and three of six available photographs. Four of the pre-written descriptions contained information regarding flaws of the item (e.g., the material was not real jade, or one of the clasps was missing, or a part of the item had been broken off). Two of the pre-written descriptions contained objective information such as size and weight. The remaining six text-based descriptions were accurate statements regarding the item. The images also varied on various binary feature dimensions such as whether the image contained a flaw, or whether the image was a close-up. In general, the available statements and images were designed to offer a well-rounded portrayal of the available items in the typical style of online auction advertisements.

Participants created the advertisements by copying and pasting descriptions from a text document (the order of the descriptions varied randomly between participants), and by uploading images from a folder which contained digital images of the item for sale (see Figure 2). Participants could click on a "preview" button in order to generate a preview the advertisement on the right hand side of the website. The chosen descriptions were concatenated into paragraph format to form the advertisement text. Below is an example paragraph, with a statement regarding flaws in bold and an objective measurement statement italicized (this formatting was not visible to the participants).

The item was created by well-known designer Lara Bohinc, as part of her 'City Lights' collection. The piece consists of striking ge-



Figure 2: Images available for advertisement generation

ometric patterns, offset by subtle curvatures which give the piece a dynamic appearance. 14 inches long in total, pendant is 2.5" by 1.3", earrings are 1.2" long and 0.6" across. The item is made from silver, and is verified with hallmark stamps. **One of the earrings is missing a clasp, which can be easily replaced.** Good condition overall, as the items were almost never worn.

The items varied in suggested value from £30 (\$49) to £210 (\$343), in steps of £30, and were presented in a pseudo-random order between subjects order using a Latin square. Participants were not briefed about these aspects of the descriptions and images. Participants were allowed to handle the actual jewelry items as they created their advertisements. This was done in order to provide the participants with the ground truth of the item, as they were able to see the quality and potential flaws of the item for themselves. After creating the advertisements, participants completed the Toronto Empathy Questionnaire (TEQ) [33], a short form of the International Personality Item Pool (mini-IPIP) [10], as well as the System Usability Scale (SUS) [4].

A sample of participants ( $n = 8$ ) participated in a follow-up interview in order to determine general self-reported strategies or biases. Questions were varied, but included "Did you have any particular strategy when choosing the descriptions for your advertisements today?", "If you haven't mentioned this yet, how did you handle whether or not to include the item flaws in the descriptions?", "Did you have a particular strategy for selecting the images for your advertisements?", "Did you consider the monetary value that the items typically sell for?", "When creating the advertisements, did you consider how the buyer would react when they receive the item in person?", and "How do you go about deciding to

include or exclude flaws in items that you post for sale in real life?"

Participants returned after one week to collect their bonuses, which were pseudo-randomly generated such that the average bonus was £21 (\$34), with the maximum bonus as £84 (\$137).

### 3.3 Participants

#### 3.3.1 Recruitment

Experienced users of online auction marketplaces were sought in order to minimize task training and maximize the ecological validity of the sample. Participants were recruited from the University of London community, and were screened for their general buying and selling habits. Those who endorsed posting 15 items for sale on online auction websites in their lifetime or more than 7 in items in the past year were eligible for participation. This ensured our sample had average experience or more than average experience as a seller on online auction websites [31]. Our sample consisted of 62 participants (*Control* condition,  $n = 16$ ; *Flat-Rate Commission*,  $n = 15$ , *Percentage Commission*,  $n = 15$ ; *Reputation* condition,  $n = 16$ ). In all, 434 online advertisements were constructed, resulting in 3,038 description choices.

#### 3.3.2 Demographics

The mean age of participants was 28.2 (range from 20.6 to 51.3  $SD = 6.31$ ), and 47 were female. They reported an average of 65.2 ( $SD = 134.1$ ) attempted lifetime sales on online auction marketplaces, and 14.8 ( $SD = 18.6$ ) in the past year. Their responses to the mini-IPIP suggested that our sample did not differ in terms of personality characteristics from a previous study involving a large sample of 2,663 respondents [10].

## 4 Results

### 4.1 Text-Based Misrepresentation

For each item, participants selected six descriptions from twelve pre-written descriptions. The participants arranged these descriptions into a coherent paragraph which became part of the final item advertisement. Four of the descriptions contained honest information regarding a flaw, and two contained object information regarding size or weight measurements. A 4 (Condition: *Control*, *Flat-Rate*, *Percentage*, *Reputation*)  $\times$  6 (Within participants: Selected location of the description within the advertisement) repeated measures ANOVA was conducted on the proportion of chosen descriptions which

Step	Task Activity
0	Complete screener questionnaire [IVs: <b>Number of online auction / in-person online classifieds sales listings created and items bought in lifetime and previous year</b> ]
1	Intake, random condition assignment
2	Verbal instruction
3	Log into Flashsale.net
4	Explore the website, choosing two listed items to view
5	Create sample advertisement for familiarization
6	Create advertisement for Item #1 [DV: <b>Time to completion (seconds)</b> ]
6a	Select 3 from 6 available photographs [DVs: <b>Which chosen; binary features of the images</b> ]
6b	Select 6 from 12 available text description statements (concurrently with Step 6a) [DVs: <b>Order of statements, information of flaws, or objective information regarding size and weight</b> ] (Repeat Step 6 for all seven items)
7	Mini-IPIP personality questionnaire [IV: <b>Multiple choice</b> ]
8	Toronto Empathy Questionnaire (TEQ) [IV: <b>Multiple choice</b> ]
9	Simple Usability Survey (SUS) [IV: <b>Likert Scale</b> ]
10	Interview questionnaire (sample of participants) [Free text responses]
11	One week later, return to collect bonuses

Table 1: Overview of the task procedure.

contained honest information (i.e., information regarding a product flaw). There was a statistically significant main effect of the Condition factor ( $F_{3,38} = 8.18$ ,  $p < 0.001$ ), and a significant main effect of description location ( $F_{5,290} = 9.26$ ,  $p < 0.001$ ). A similar ANOVA was conducted for whether an objective statement was included as a description. There was a significant main effect of description location ( $F_{5,290} = 14.50$ ,  $p < 0.001$ ). All other main effects and interactions were non-significant. Figures 3a and 3b illustrate these findings. In short, all participants were more likely to include both flaws and objective statements in later versus earlier parts of the constructed text description paragraphs. Importantly, those in the *Reputation* condition were more honest in their inclusion of flaws in the text descriptions, although no different in their inclusion of objective state-

ments.

The order of presentation of the item during the task was also considered as a within subjects factor. A 4 (Condition) x 7 (Within participants: Order of presentation) repeated measures ANOVA was conducted on honest information as well as objective information. As expected, there was an equivalent main effect of Condition as above, although no significant main effect or interaction involving presentation order ( $F_s < 0.5$ ). Thus, deceptive behavior did not vary over time during the course of the experiment.

## 4.2 Image-Based Misrepresentation

Images varied along four binary feature dimensions: Whether an item flaw was in view (for an example, see Figure 2a), whether the image was a close-up (Figure 2b), whether the image contained the entire item in view (Figure 2c), and/or whether the image contained a hallmark stamp (Figure 2d). The analysis from the previous section was conducted using the binary feature characteristics of items chosen for the advertisements as the dependent variable. There were no statistically significant main effects, interactions, or correlations.

The probability of randomly selecting images with flaws was calculated using a hypergeometric cumulative distribution function [37]. The actual number of image flaws selected by each participant was used to calculate the probability of randomly selecting the number of chosen image flaws or fewer. Table 2) also includes the probability of selecting the number of text-based honest descriptions. It is apparent that participants likely suppressed text-based honesty, while image-based honesty was no different than a random selector.

## 4.3 Assessment of Usability

The System Usability Scale (SUS) ranges in score from 0 (not at all usable) to 100 (most usable). A previous meta-analysis ( $n = 2,324$ ) of usability studies revealed a negatively skewed distribution of SUS scores with a mean of 70.14 (Median = 75, SD = 21.71). In the present

Condition	$P(\leq \text{Honest Images})$	$P(\leq \text{Honest Text})$
Control	0.53 (0.08)	0.06 (0.04)
Flat Comm.	0.51 (0.09)	0.01 (0.01)
% Comm.	0.53 (0.09)	0.03 (0.01)
Reputation	0.63 (0.09)	0.20 (0.01)

Table 2: Probability of a random selector including the measured number of honest information, calculated using a hypergeometric cumulative distribution function. Standard error is in parentheses.

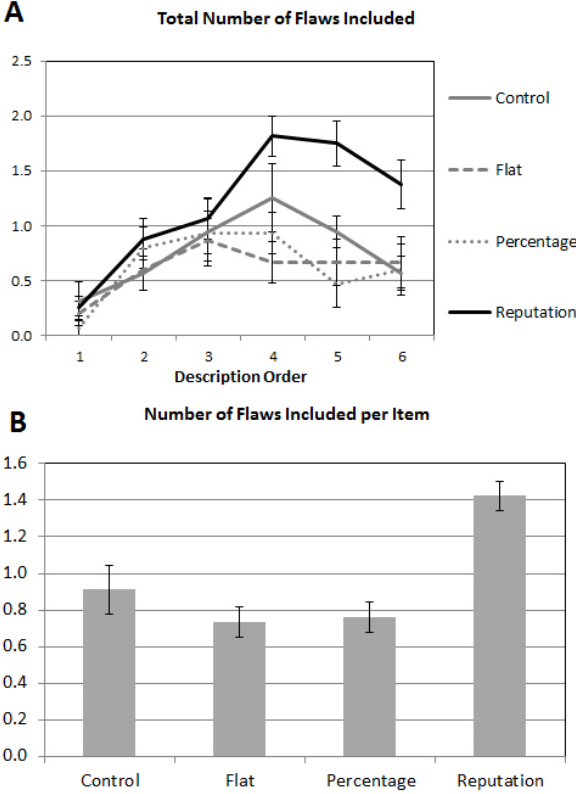


Figure 3: Number of flaws included (A) at the different location orders, and (B) overall, out of four available flaws.

study, the mean SUS score was 87.18 (Median = 92.5, SD = 11.73). This suggests that subjective usability in the present study was higher than the average study. A 4 (Condition)  $\times$  2 (Gender) ANOVA revealed that SUS did not vary by condition ( $F_{3,54} = 0.73, p = 0.54$ ), nor gender ( $F_{1,54} = 2.24, p = 0.14$ ). Although, the gender imbalance towards females (binomial test,  $p < 0.001$ ) limits the interpretation of a lack of gender effect. SUS also did not correlate with the number of flaws included in text descriptions ( $R = -0.02, p = 0.88$ ), nor flaws included in images ( $R = -0.05, p = 0.73$ ). In summary, subjective usability did not appear to play a significant role in the performance aspects of the present study.

#### 4.4 Demographics and Personality Characteristics

The experimental data were analyzed using logistic estimation techniques, estimated using the econometrics package STATA 13. The experimental data capture binary responses from the experimental subjects about whether or not to use a statement revealing a flaw in their advertisements. This behavior can be modelled as a dis-

crete choice problem and estimated using the binary dependent variable estimation technique logit in which a logistic function is used to capture the probability that a flaw will be reported ( $P_{flaw}$ ):

$$P_{flaw} = G(x\beta) = \frac{\exp(x\beta)}{1 + \exp(x\beta)} \quad (1)$$

where  $x\beta$  is a matrix of explanatory variables and accompanying parameters. Panel fixed effect (FE) estimation was used to capture the fact that the preferences of individual experimental subjects may vary and to overcome problems of endogeneity created by heterogeneity bias in a panel estimation context. FE estimation was used in preference to random effects because FE generates concrete parameter estimates for subject-specific differences and the heterogeneity in the econometric models is not purely the outcome of randomness [40].

We use  $z$ -tests to test the individual significance of each explanatory variable. For interpretative purposes, conventional parameter estimates are not illuminating for logistic estimation because marginal impacts of explanatory variables will be non-constant reflecting the non-linear function form of the logistic function. For this reason we have reported STATA’s estimates of the odds ratio. For a given explanatory variable, if the odds ratio is less than 1, flaws are less likely to be reported; for an odds ratio greater than one, flaws are more likely to be reported.

Table 3 reports the probability of including honest information (i.e., not engaging in deceptive behavior) in relation to personality characteristics and demographics. The logistic regression featured 1860 total observations, with a model  $\chi^2 = 607.9$  ( $p < 0.001$ ). Importantly, assignment to the *Reputation* condition was a significant predictor of including honest information regarding flaws ( $\beta = 1.79, z = 3.01, p < 0.01$ ). This finding was consistent with the ANOVA conducted above, which found a main effect of Condition in relation to the dependent variable of number of flaws included in the text based descriptions.

#### 4.5 Follow-up Interviews

A qualitative follow-up interview was conducted on a sample of participants ( $n = 8$ ). A brief summary of the typical responses is provided here. The critical insight comes from the response to the question, “*Did you have any particular strategy when choosing the descriptions for your advertisements today?*” All eight participants mentioned a desire to create an honest representation of the item for sale. For example, one participant responded, “*I chose the statements I felt gave the most honest description of the items whilst still seeming positive*



*and creating appeal. When I am shopping online I hate when there is a lack of information about the item I am considering.”* Yet, despite this sentiment, this participant included only 2 of 20 available item flaws. As illustrated by the above results, there was an overall tendency to exclude information about flaws. The interview responses suggest that participants may not be aware of this bias, and believe they are being honest and open regarding items, when in fact they are hiding critical information which may lead to buyers feeling they have been misled. In other topics, respondents accurately self-reported. For example, respondents reported not attending to the value of the objects, and this was confirmed in the statistical analysis. Despite the interview involving only a subset of the participants, future work should consider whether this mismatch between self-reported behavior and measured behavior results from self-deception, active deception in self-reports, or other metacognitive biases [36].

## 5 Discussion

### 5.1 Conclusions

The present study represents the first attempt to capture behavior at the moment of advertisement creation. We tested the role of incentives on seller behavior, as well as personality traits of the seller. In doing so, we provide insights into basic research questions regarding human deception. Furthermore, these insights can inform the design of online auction websites.

In the present study, we found that sellers tended to hide critical item flaws in the middle or end of the advertisement. We also found that sellers were most honest when the saliency of the website’s reputation system was promoted (supporting H1). Importantly, while the saliency of the reputation system had an impact on the text-based information channel, it had no impact on the image-based information channel (supporting H3). However, probability analysis (see Table 2) suggests active suppression of honest text-based information relative to image-based information. The value of the items did not correlate ordinally with deception (supporting H2), although future work could consider higher stakes.

We predicted a specific signature of personality traits to be associated negatively with honesty: Extraversion, neuroticism, and intellect/imagination. We found significant support for neuroticism (H4B), moderate support for extraversion (H4A), but no support for intellect/imagination (H4C). We predicted a separate signature of personality traits to correspond positively with honesty: Agreeableness, empathy, and conscientiousness. However, we found no support for these predictions (no support for H5A, H5B, and H5C). The current study demonstrates that experience as a seller (but importantly not a buyer)

in online market places is associated with increased honesty (partial support for H5D).

## 5.2 Implications

### 5.2.1 Design of Online Auction Marketplaces

The implications of these findings lead to design suggestions for online marketplaces. First, the reputation or feedback system should be at the forefront of the advertisement creation tool. We have shown in the present study that this can lead to the suppression of deception via decreased item misrepresentation.

Second, an overall finding in the present study is that sellers place honest information in the middle or end of the advertisement. Thus, one idea to increase the visibility of critical information is to continuously shuffle the order of information presented in the advertisement. To prevent issues of comprehension or flow, sellers could be informed of this upfront.

Third, buyers should be notified or reminded to closely assess the images included on advertisements, and to not rely on text-based descriptions to provide critical information regarding potential item flaws. This could be done using a text or icon-based warning that appears when the user attempts to purchase an item before viewing all available images. By orienting the potential buyers to all available images, it might be possible to increase buyers’ awareness of potential flaws. This awareness will shape buyer expectation, hopefully precluding the deduction of having been deceived. Similarly, sellers could be notified of the overall system-wide tendency for misrepresentation.

Fourth, the free-form interviewing procedure revealed that sellers may believe that they are being honest, when their behavior reveals that they are engaging in deceptive practices. This implies that simple reminders and warnings may not be sufficient to shape seller behavior. Online auction websites should not overly rely on targeting sellers without also alerting buyers to the likelihood of deceptive practices.

### 5.2.2 Insights about Behavior

We assessed the demographic and personality characteristics of our participants in terms of their deceptive behavior. Importantly, our sample represented a typical cross section of personality types from the general population. In the current study, we found the personality traits of extraversion (moderately significant) and neuroticism to be associated with honest behavior. This builds on prior work by demonstrating that certain personality traits are associated with deceptive behavior, although the present study shows these findings are robust to real-time behavior in an online social environment.

Measure	DOI on Honesty	Hypothesized DOI	$\beta$	z-score	p-value
Reputation feedback	+	+ (H1)	1.79	3.01	0.001***
Extraversion†	-	- (H4A)	0.67	-1.84	0.07*
Neuroticism†	-	- (H4B)	0.54	-2.29	0.02**
Intellect/Imagination†	n/s	- (H4C)	1.24	0.76	0.45
Agreeableness†	n/s	+ (H5A)	1.01	0.03	0.98
Empathy	n/s	+ (H5B)	0.84	-0.4	0.69
Conscientiousness†	n/s	+ (H5C)	0.95	-0.27	0.79
Experience as seller	+	+ (H5D)	1.19	2.2	0.03**
Experience as buyer	n/s	+ (H5D)	0.92	-1.32	0.19

Table 3: Logistic regression of personality and demographic characteristics, in relation to the probability to include honest information in the constructed advertisements; DOI = direction of impact; n/s = non-significant; \*statistically significant at the  $\alpha=0.10$  level; \*\* statistically significant at the  $\alpha=0.05$  level; \*\*\* statistically significant at the  $\alpha=0.01$  level; † Mini-IPIP subscore

Detecting personality traits linked with deceptive behavior signifies that certain cognitive mechanisms might underlie deception generation. By honing in on these traits, it may be possible to infer how the brain generates deceptive action, leading to discoveries which can speak to deterrence techniques.

It is interesting to note that no personality traits were found to correlate positively with honesty. At first sight, this may seem to be related to task characteristics. However, the personality traits measured by the mini-IPIP could easily have been reverse coded and thus labelled conversely. For example, the extraversion sub-score could be considered to be the converse of an introversion sub-score. Thus, it is important not to reach unwarranted conclusions about the lack of detecting personality traits which correlated positively with honesty.

We did not find that older participants were less likely to engage in deceptive behavior, although we did find that online experience in auction marketplaces does predict more honesty (i.e., fewer item misrepresentations). This is important in that it poses important questions regarding the causal direction of seller experience and honesty. Are those who are more honest simply more likely to survive in the market, or does the process of repeated sales lead to a learning process? This is considered below in the future work section.

Robust to all conditions in the present study, honest information regarding flaws was almost never included early in the text based advertisements. This demonstrates a strong tendency to begin the advertisement with a positive statement. Likewise, there was a tendency to end the advertisements with a positive statement. While at first blush this seems to simply be “good practice”, sellers may be taking natural advantage of primacy and recency effects by crafting descriptions such that negative information is situated where the seller believes the buyer is less likely to attend or remember.

### 5.3 Future Work

The current study demonstrates that those with more seller experience are also less likely to engage in deceptive practices. We consider the effects of experience to represent the result of repeated instances of engaging in the act of selling on an online market place. Future work should incorporate a fuller range of older adults to assess the role of normal aging in deceptive online behavior.

It could also be possible that those with honest traits simply survive in the market place, although in the present study there was no correlation with seller experience and the personality traits which were associated with honesty. Thus, there appears to be a real impact of learning on seller behavior in terms of deceptive practices. Still, the age range of the current study does not include those who would be typically considered to be older adults [12], and so future work should consider the role of normal aging in online deceptive behavior.

Due to the potential importance of learning to curb deceptive practices from experience, future work should also include repeated sessions of advertisement creation in order to assess the characteristics, saliency, and modality of buyer feedback communication which can lead to changes in seller behavior over time. By defining the various feedback channels involved in a seller’s learning experience, and quantifying the impact of these various channels, future work can hone in on the social, personal, and economic mechanisms which can lead to the deterrence of deception in online interactions. Finally, the current study included pre-written descriptions in order to facilitate the collection of quantifiable behavioral metrics. Future studies should consider free-form responses in order to build upon the generalizability of the current results.

## 6 Acknowledgments

The authors gratefully acknowledge those who provided helpful discussion. The work was supported by the EP-SRC under grant #EP/K033476/1.

## References

- [1] AKERLOF, G. A. The market for "lemons": Quality uncertainty and the market mechanism. *The quarterly journal of economics* (1970), 488–500.
- [2] BA, S., WHINSTON, A. B., AND ZHANG, H. Building trust in online auction markets through an economic incentive mechanism. *Decision Support Systems* 35, 3 (2003), 273–286.
- [3] BOHNET, I., FREY, B. S., AND HUCK, S. More order with less law: On contract enforcement, trust, and crowding. In *American Political Science Association* (2001), vol. 95, Cambridge Univ Press, pp. 131–144.
- [4] BROOKE, J. Sus-a quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
- [5] BYRNE, R. W., AND CORP, N. Neocortex size predicts deception rate in primates. *Proceedings of the Royal Society B: Biological Sciences* 271, 1549 (2004), 1693.
- [6] CASPI, A., AND GORSKY, P. Online deception: Prevalence, motivation, and emotion. *CyberPsychology & Behavior* 9, 1 (2006), 54–59.
- [7] CHIU, S.-P., CHOU, H.-W., AND CHIU, C.-M. The antecedents of buyers' perceived justice in online markets. *Cyberpsychology, Behavior, and Social Networking* 16, 7 (2013), 536–542.
- [8] CHUA, C. E. H., WAREHAM, J., AND ROBEY, D. The role of online trading communities in managing internet auction fraud. *MIS Quarterly* (2007), 759–781.
- [9] DIEKMANN, A., JANN, B., PRZEPIORKA, W., AND WEHRLI, S. Reputation formation and the evolution of cooperation in anonymous online markets. *American sociological review* 79, 1 (2014), 65–85.
- [10] DONNELLAN, M. B., OSWALD, F. L., BAIRD, B. M., AND LUCAS, R. E. The mini-ipip scales: tiny-yet-effective measures of the big five factors of personality. *Psychological assessment* 18, 2 (2006), 192.
- [11] GAVISH, B., AND TUCCI, C. L. Reducing internet auction fraud. *Communications of the ACM* 51, 5 (2008), 89–97.
- [12] GLASS, B. D., CHOTIBUT, T., PACHECO, J., SCHNYER, D. M., AND MADDOX, W. T. Normal aging and the dissociable prototype learning systems. *Psychology and aging* 27, 1 (2012), 120.
- [13] GOMBOS, V. A. The cognition of deception: the role of executive processes in producing lies. *Genetic, social, and general psychology monographs* 132, 3 (2006), 197–214.
- [14] GRAZIANO, W. G., HABASHI, M. M., SHEESE, B. E., AND TOBIN, R. M. Agreeableness, empathy, and helping: A person  $\times$  situation perspective. *Journal of personality and social psychology* 93, 4 (2007), 583.
- [15] GREGG, D. G., AND SCOTT, J. E. A typology of complaints about ebay sellers. *Communications of the ACM* 51, 4 (2008), 69–74.
- [16] GRIEVE, R., AND MAHAR, D. The role of fluid and emotional intelligence in malingering. *Australian Journal of Psychology* 62, 2 (2010), 103–111.
- [17] HOUSER, D., AND WOODERS, J. Reputation in auctions: Theory, and evidence from ebay. *Journal of Economics & Management Strategy* 15, 2 (2006), 353–369.
- [18] KRISTOF-BROWN, A., BARRICK, M. R., AND FRANKE, M. Applicant impression management: Dispositional influences and consequences for recruiter perceptions of fit and similarity. *Journal of Management* 28, 1 (2002), 27–46.
- [19] LALLY, P., VAN JAARSVELD, C. H., POTTS, H. W., AND WARDLE, J. How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology* 40, 6 (2010), 998–1009.
- [20] LEVINE, T. R., BLAIR, J. P., AND CLARE, D. D. Diagnostic utility: Experimental demonstrations and replications of powerful question effects in high-stakes deception detection. *Human Communication Research* 40, 2 (2014), 262–289.
- [21] LOPES, L. L. Psychology and economics: Perspectives on risk, cooperation, and the marketplace. *Annual Review of Psychology* 45, 1 (1994), 197–227.
- [22] LORENZ, E. Trust, contract and economic cooperation. *Cambridge journal of economics* 23, 3 (1999), 301–315.
- [23] MACINNES, I. Causes of disputes in online auctions. *Electronic Markets* 15, 2 (2005), 146–157.
- [24] MANN, S., VRIJ, A., AND BULL, R. Suspects, lies, and videotape: An analysis of authentic high-stake liars. *Law and human behavior* 26, 3 (2002), 365–376.
- [25] MICHIKYAN, M., SUBRAHMANYAM, K., AND DENNIS, J. Can you tell who i am? neuroticism, extraversion, and online self-presentation among young adults. *Computers in Human Behavior* 33 (2014), 179–183.
- [26] MITCHELL, G. Revisiting truth or triviality the external validity of research in the psychological laboratory. *Perspectives on Psychological Science* 7, 2 (2012), 109–117.
- [27] PHILLIPS, M. C., MEEK, S. W., AND VENDEMIA, J. M. Understanding the underlying structure of deceptive behaviors. *Personality and Individual Differences* 50, 6 (2011), 783–789.
- [28] REICHLING, F. Effects of reputation mechanisms on fraud prevention in ebay auctions. Tech. rep., Working Paper, Stanford University, 2004.
- [29] RESNICK, P., ZECKHAUSER, R., SWANSON, J., AND LOCKWOOD, K. The value of reputation on ebay: A controlled experiment. *Experimental Economics* 9, 2 (2006), 79–101.
- [30] RIEGELSBERGER, J., SASSE, M. A., AND MCCARTHY, J. D. The mechanics of trust: A framework for research and design. *International Journal of Human-Computer Studies* 62, 3 (2005), 381–422.
- [31] SCOTT, S. Ebay marketplaces. JMP Securities Research Conference.
- [32] SHIMER, R., AND SMITH, L. Assortative matching and search. *Econometrica* 68, 2 (2000), 343–369.
- [33] SPRENG, R. N., MCKINNON, M. C., MAR, R. A., AND LEVINE, B. The toronto empathy questionnaire: scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of personality assessment* 91, 1 (2009), 62–71.
- [34] STANDIFIRD, S. S. Reputation and e-commerce: ebay auctions and the asymmetrical impact of positive and negative ratings. *Journal of Management* 27, 3 (2001), 279–295.
- [35] TREVATHAN, J., AND READ, W. Undesirable and fraudulent behaviour in online auctions. In *SECRYPT* (2006), Citeseer, pp. 450–458.
- [36] TRIVERS, R. Deceit and self-deception. In *Mind the gap*. Springer, 2010, pp. 373–393.
- [37] VIDAKOVIC, B. *Statistics for bioengineering sciences: with MATLAB and WinBUGS Support*. Springer Science & Business Media, 2011.

- [38] VRIJ, A. *Detecting lies and deceit: Pitfalls and opportunities*. John Wiley & Sons, 2008.
- [39] VRIJ, A., GRANHAG, P. A., AND MANN, S. Good liars. *J. Psychiatry & L.* 38 (2010), 77.
- [40] WOOLDRIDGE, J. M. Fixed-effects and related estimators for correlated random-coefficient and treatment-effect panel data models. *Review of Economics and Statistics* 87, 2 (2005), 385–390.
- [41] YAMAMOTO, H., ISHIDA, K., AND OHTA, T. Modeling reputation management system on online c2c market. *Computational & Mathematical Organization Theory* 10, 2 (2004), 165–178.
- [42] YAVAŞ, A. Economics of brokerage: an overview. *Journal of Real Estate Literature* 2, 2 (1994), 169–195.
- [43] ZHANG, J., AND COHEN, R. Design of a mechanism for promoting honesty in e-marketplaces. In *PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE* (2007), vol. 22, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, p. 1495.