

# On the Influence of User Reviews in E-Commerce: The Expertise Tag

**Nick Colvin**

University of Washington  
nccolvin@uw.edu

**Michael Gelon**

University of Washington  
mgelon@uw.edu

**Gabby Mehlman**

University of Washington  
gnmehlman@uw.edu

**Matt Soave**

University of Washington  
soave@uw.edu

## Abstract

Customers buying products online often consult user reviews to inform their decisions. A typical user review on an e-commerce site might include a rating from 1 - 5 with accompanying text. Currently, there lacks research investigating the effect of reviewer credential tags on a user's perception and intent to purchase. In this paper, we conduct a study utilizing Amazon's Mechanical Turk crowdsourcing platform. 388 participants encountered one of nine review treatments that varied in star rating and reviewer expertise credential, and reported their likelihood to recommend the product. Our findings reinforce that star ratings play a large role in customer perception, but find no significant effect of supporting the impact of reviewer credentials in the format we presented. Further research is recommended to determine if credentials might influence reviewers in other contexts.

*Keywords: user reviews, e-commerce, credentials, star rating, online shopping*

## 1. Introduction

In the realm of e-commerce, consumers must make decisions about purchasing a product without being able to directly examine, test, or evaluate it in a tangible way. They must rely on the medium of electronic information to decide if the product will satisfy their desires and requirements. This information comes in the form of images, descriptions, videos, and reviews. Since prevalent e-commerce services like Ebay implemented a customer review system, shoppers continue to rely more and more on the appraisal of their peers; as of 2013, customer reviews are now the second most trusted online source of information, with 70 percent of online shoppers stating that they trust reviews (Nielsen, 2013).

There has been myriad research in both academic and commercial landscapes studying how customer-created reviews affect the behavior and buying intention of shoppers. One particular area of interest is how shoppers come to assess the quality and believability of the reviews left by seemingly anonymous members of the internet community. While social networking sites have become a source trusted feedback, it is not always the case that a customer's reliable peers have used and reviewed a product he or she might be interested in. This leaves online shoppers at the mercy of reviews written by strangers.

To increase a reviewer's credibility, websites like Newegg.com and Amazon.com will indicate within a review that the reviewer has legitimately purchased the product with a "verified purchase" tag. Further, the website reddit.com utilizes credential tags that allow users to showcase that they have a particular vocation or degree to bolster the credibility of their statements and reviews. Quora.com and stackoverflow.com, websites for answering questions posted online, also provide a rating system which enables users to evaluate how helpful a poster has been in the past. Quora.com also allows answerers to label themselves with an expertise tag. These are all examples of design solutions that these companies have implemented to provide credibility to their online users. However, we are unaware of any research that explores the effectiveness or impact of these designs.

Correlational research has been conducted to study how users determine if a review is credible or not, but there is a distinct lack of experimental inquiry on this topic. We aim to conduct an experiment to investigate if online shoppers are influenced by reviewers that are indicated as having domain specific knowledge related to the product. We also seek to examine the moderating factor of star rating. Will a review made by a user indicated as having domain specific knowledge of the product influence a consumer's perception of the product more than a review without credentials? Through experimental analysis, we hope to determine if our 'credential tagging' design decision could help customers in assessing which online product to choose.

## 2. Literature Review

### 2.1 Online Trust

Online trust plays a key role in our experimental study. Since users cannot interact with or examine a product in an online market, trusting the provider and the information presented is an important aspect of their perception of the product. By examining past studies on trust, we sought to identify influential pieces that provide insight into how users develop and lose trust for online services.

E-commerce as an enterprise has continued to command a growing market share of the exchange of commercial goods and services. As the internet becomes a more readily available service, e-commerce more widely provides a fast and convenient platform for customers to shop from the comforts of their home. Thus, knowledge of how businesses could tap into this online market has become a heavily researched topic. It became apparent to businesses and researchers alike that trust was a major factor in convincing people that shopping online is safe and reliable. Early sociological studies of trust revealed that it largely stems from familiarity (Luhmann, 1979), thus online systems have presented a challenge for business and design.

(Ruyter et. al 2001) investigated factors playing into customers' willingness to adopt an e-service. Their experimental study looked at variables including how well a service is known, the perceived advantages of a service, and security risk to see how they affected a customer's trust in a web service. They found that well known companies were perceived as being less risky, and that this was the dominant effect regarding customer trust. Perceived advantages proved to be the least important influence because subjects didn't know if they could trust electronic services in the first place. This paper set some of the early groundwork for why online trust is a vital and potentially profitable factor to study when creating or managing an online business.

## **2.2 Effects of User Reviews**

Research on e-commerce user reviews has shown that user input has an impact on merchant and product perception, ultimately affecting purchase intent and sales.

(Flanagin et. al 2011) explored the factors that contribute to trust in online information and found that user-generated content (e.g. user reviews) played an important role in judging the credibility of online information, suggesting that such content is invaluable to e-commerce sites in their efforts to sell products and retain customers. They also conducted a quasi-experiment to determine the effects of review quantity and average rating on the perceived quality of different product types. When the content of user reviews was hidden, they found that the average rating (in number of stars) had a significant positive effect on perceived quality, while the quantity of reviews alone had no significant effect. Overall, their results suggest consumers use peer-generated ratings in a significant way to assess both the information being communicated by the e-commerce company and the quality of the product itself. The study's results, however, aggregated all reviews on a product; our current research intends to explore the effects of individual reviews and the factors therein.

In opposition to some findings in (Flanagin et. al, 2011), an earlier study performed by (Park et. al, 2006) found that review quantity had a significant effect on a user's intention to purchase a product, but only while they were not highly 'involved'. This experimental study looked to determine if the role of involvement (i.e. if they motivated subjects with a statement like, "You are purchasing this product for a friend.") altered the way they looked at online reviews while shopping for a portable music device. If a subject was not motivated (i.e. not given any role-playing scenario), the number of reviews and star rating had the most significant effect on their intent to purchase the product. However, during the 'high involvement' condition, subjects focused more on the quality of the written reviews rather than the satisficing heuristic of checking the star rating and number of reviewers. This article provides grounds to consider different types of customers and how they might approach utilizing reviews in a disparate way.

## **2.3 Credibility**

### ***2.3.1 Effect of user reviews on merchant credibility***

User reviews can also increase the perceived trustworthiness of the merchant (as opposed to the trustworthiness of the review itself).

Recognizing the increased level of risk in purchasing from a remote online website as opposed to a tangible brick & mortar location, (Flanagin et. al 2011) investigated the factors that consumers use to judge the credibility of online information and mitigate risk. They found that in addition to primary source information (i.e. from the site owner), second-hand input (i.e. information from peers, experts, or other individuals not associated with the site owner) was considered very important in assessing the credibility of online information. User reviews were highlighted as a vital component of e-commerce sites, whose success generally relies on customer trust.

(Utz et. al 2012) explored the quantitative effects of consumer reviews of stores on perceived trustworthiness of online stores. They performed two experiments in order to test eight different hypotheses, treating "dispositional trust" (i.e. how trusting a person tends to be) as a moderator. The first experiment explored how reviews (positive or negative) and store reputation (also positive or

negative) independently affected participants' trust in the online store. They found that positive reviews were positively correlated with perceived trustworthiness, but did not establish a significant correlation between store reputation and perceived trustworthiness, suggesting that reviews were more influential than the store reputation. Furthermore, individuals who were generally more trusting were more heavily influenced by reviews. Their second experiment compared the addition or absence of assurance seals against positive, negative or neutral consumer reviews. With the addition of the neutral condition, it was proven that a participant was not trusting by default since the positive review improved trustworthiness (and the negative reduced it).

These studies show how second-hand information, particularly user reviews, can greatly affect a the perceived trustworthiness of a store, regardless of its general reputation.

### **2.3.2 Credibility of user reviews**

There are many factors that can affect the perceived credibility of the review itself, such as the source of the recommendation or the type of website it appears on. Perceived credibility is often separated into perceived trustworthiness and perceived expertise.

In "*The Ironic Effect of Source Identification on the Perceived Credibility of Online Product Reviewers*" (Willemsen et al 2012), postulated and found that perceived trustworthiness and expertise are not one and the same. A consumer may believe that the review writer is an expert, but still not trust his or her review to be of genuine intent. In fact, a self-proclaimed expert was perceived as more expert but less trustworthy than a typical consumer. Our credential tag will, in a way, seek to confirm or deny this finding by Willemsen. If we discover that consumers are more influenced by the review of a user tagged as having expert domain knowledge, we can make a claim that they are in fact perceived as more trustworthy.

(Senecal et al, 2004) explored how consumers' usage of online recommendation sources influenced their online product choices and perceived trust in a review, and how the type of website, source of recommendation, and type of product might moderate this effect. In this study, they found that recommender systems (i.e. computer generated) were more influential than typical consumers or human experts. Typical consumers were perceived as less expert than human experts or recommender systems, and human experts were perceived as more expert than recommender systems. Consumers were perceived as more trustworthy than either human experts or recommender systems, between which there was no significant difference in trustworthiness. The type of website (non-commercially linked 3<sup>rd</sup> party, commercially linked 3<sup>rd</sup> party, and retailer) did not influence the perceived trustworthiness. It is interesting that the recommender system was most influential even though it was seen as less expert than human experts and less trustworthy than consumers.

In our research we want to further explore what factors might moderate the perceived trustworthiness of a self-proclaimed expert's review since it has been found experts are often perceived as having high expertise but not necessarily as being trustworthy. These findings could have important implications on the future design of online review systems.

### 3. Hypotheses

**H1:** The likelihood to recommend the product to a friend will be positively correlated with the number of stars given by the single review.

(Flanigan et. al 2011) unsurprisingly found that perception of product quality increased as the product's average rating increased. In this experiment, we are exploring a different dependent variable (likelihood to recommend the product to a friend) and are manipulating the star rating for a single review rather than an aggregate. Despite these differences, we expect to see a similar effect of positive correlation between the rating and product perception.

**H2:** There will be no significant difference in likelihood to recommend between the exclusion of a profession tag and the inclusion of an irrelevant profession tag.

In order to isolate the effects of the content of the tag and the presence of the tag itself, our experiment includes an irrelevant tag ("account manager") whose contents we do not expect to affect the perception of the review. Across each star rating, we expect the likelihood to recommend to not be statistically different between the two tag conditions.

**H3:** For a given star rating, the expert profession tag will amplify the effect on likelihood to recommend compared to the irrelevant profession tag.

We expect that participants will perceive an expert's review to be more valuable than a non-expert's review and will therefore mentally "weight" the review more heavily as they compare it to the aggregate rating. That is, assuming that likelihood to recommend is correlated with star rating, we expect an expert's 5-star rating to increase the likelihood to recommend (compared to a 4-star rating) than a non-expert's 5-star review. Likewise, we expect an expert's 3-star review to decrease the likelihood to recommend more significantly than a non-expert's 3-star review. Finally, we expect a 4-star expert review to match a 4-star non-expert review because the single review rating is exactly the same as the aggregate rating.



\*A visualization of hypothesized results

## 4. Methods

### 4.1 Participants

We gathered data from 471 subjects using Amazon's Mechanical Turk Platform. Since our study was relatively short in length and could be administered online, we found that Mechanical Turk gave us the option of collecting data from a large number of subjects in a relatively short amount of time. Our study consisted of 9 separate conditions. In the interest of achieving a worthy degree of statistical significance we recruited and randomly assigned, on average, 52 subjects per condition. Each participant could only participate in one trial and afterwards was restricted from taking part in our study again. Subjects were informed that the study should take approximately 3-5 minutes of their time, and were compensated 5 U.S. cents for their participation.

Mechanical Turk has been shown to not only be capable of gathering high quality data, but also to reflect a more diverse demographic of subjects than recruiting from a college campus (Buhrmester et al, 2011).

### 4.2 Design

Using a custom template that was made to resemble common e-commerce websites like Amazon and Ebay, we presented the participants with a page that depicts a review for a Sony digital camera. We chose a stock image by searching for 'digital camera' on Amazon.com, and altered the name to not reflect a particular series of Sony cameras. Subjects were also shown the average star rating for the product, at 4 stars. Following, we display basic information about the camera's features and end with a user review that contains its own textual product description.

All of the page attributes described above remained static as a control. For our treatments, we manipulated the user reviewer credential tag and the reviewer's star rating. Both variables are depicted in context by **Figure 1**. In **Figure 1**, the example credential tag is the expert "professional photographer tag" and the individual star rating is five stars.

**Sony 5X Zoom Digital Camera with 20.1 Megapixels**  
Average rating: ★★★★★

**Camera Features:**

- 5x optical zoom plus an astounding 68x Dynamic Fine Zoom
- Comfortable design, elegantly simple controls and intuitive menu system
- 3.0-inch 921,000-dot LCD display
- Full HD 1080p videos with stereo sound for dazzling movies

**User Reviews:**

**Alex Johnson - Profession: Portrait Photographer** ★★★★★

I bought this camera to keep in my car for occasions that require something better than my phone's camera. For that purpose, it meets my expectations and sometimes exceeds them. However, this is still a snapshot device with a relatively low price tag and therefore a DSLR level quality (like high shutter speed, good low light performance, etc.) should not be expected.

**Pros:**

- Lightweight & slim: It can fit in small spaces and feels like a high-quality item.
- Easy menu mode: Practicality should be the main focus in a snapshot device. This is one of the most user-friendly interfaces I have seen.
- Image quality: Not quite DSLR quality, but this one is not bad at all, especially in daylight.

**Cons:**

- Low light performance: I was not expecting professional quality from this device, but I still found the low light performance to be just okay.
- Slow shutter speed: It has a continuous shooting mode which is supposed to take multiple frames as long as the shutter button is depressed, but sometimes it is too slow.
- Zoom buttons: The zoom button is not very convenient IMO.

Overall, whether you will like it or not depends on your expectations and how to plan to use it. For me it works well enough and has several useful extra features.

**Figure 1.** Sample of what participants saw for the review with independent variables highlighted

Our credential tag has three levels: absent, product-domain irrelevant, and product-domain relevant. In the absent condition the subjects were simply presented with the reviewer's name, 'Alex Johnson.' Domain-irrelevant conditions showed 'Alex Johnson' as an 'Account Manager,' and finally domain-specific trials presented 'Alex Johnson' as a 'Portrait Photographer.' We included the domain-irrelevant condition to ensure that the mere presence of a credential tag did not have an effect on how users perceived the product. Our single review's star ratings were manipulated to levels of either three, four, or five stars. These contrasted or confirmed the average star review control variable that is presented at the top of the product page, effectively making the single user's review lower than, the same as, or higher than the aggregate rating.

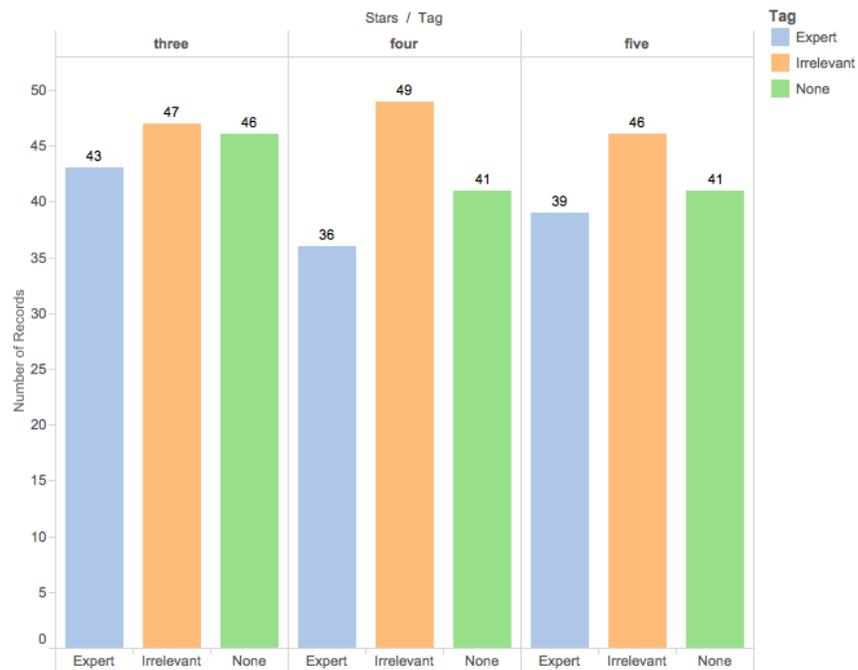
Our dependent variable consisted of a 7 choice Likert scale with the question, "How likely would you be to recommend this product to a friend based on the information that you see?". **Figure 2** shows how this question was presented in the study.

**How likely would you be to recommend this product to a friend based on the information you see?\***  
(on a scale of 1 to 7)

- 1. Very unlikely to recommend
- 2. Unlikely to recommend
- 3. Somewhat unlikely to recommend
- 4. Neutral
- 5. Somewhat likely to recommend
- 6. Likely to recommend
- 7. Very likely to recommend

**Figure 2** shows how we measured our dependent variable ‘willingness to recommend’.

Finally, at the conclusion of the study we included a screener question to judge if subjects had read the content on the page in order to identify participants who may have rushed through the study just to get their compensation. The question inquired about content present in the textual review by asking which of four statements was mentioned in the review. We found that 83 of 471 subjects were unable to answer this question correctly, and their data was removed from the study. **Figure 3** below shows the number of subjects that successfully completed each condition and answered the screener question correctly.



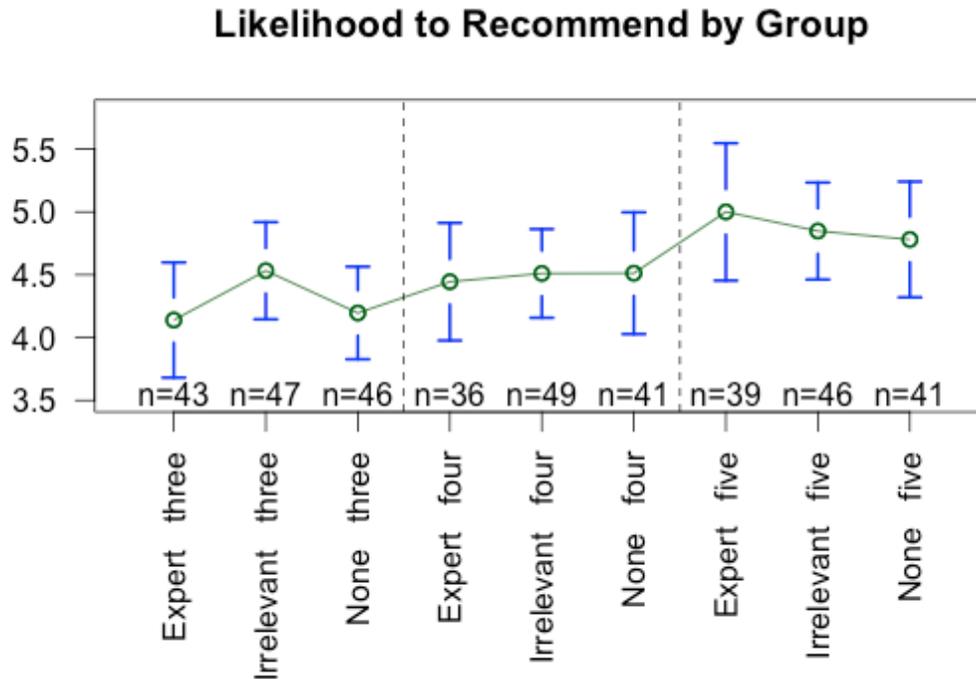
**Figure 3.** The number of subjects in each condition, post-parsing the data.

After collection, we performed a two-way ANOVA analysis to determine any significant effects as well as interactions. Further, a Post-Hoc Tukey HSD test was used to determine which interactions proved significant.

## 5. Findings

All statistical tests were performed on the data post-parsing (N = 388).

The Levene's test of equality of error variances proved to be insignificant ( $p=0.827$ ), and thus equal variances could be assumed. **Figure 4** below shows the means and confidence intervals for each test group.



**Figure 4.** Means and confidence intervals by test group

A two-way between groups ANOVA was conducted to explore the impact of the credential tag and number of stars on participants' likelihood to recommend the product to a friend. The interaction effects between the tag and the number of stars were found to be insignificant ( $F(4,379)= 0.494$ ,  $p=0.74$ ). The number of stars on the individual recommendation did have a significant effect for  $p<0.5$  ( $F(2,379)=5.685$ ,  $p=0.0037$ ), though it was a small to moderate effect (partial  $\eta^2 = 0.029$ ). The credential tag was not found to be significant ( $F(2,379)=0.756$ ,  $p=0.681$ ).

A post-hoc Tukey HSD test showed that there was a significant difference between a reviewer rating of three stars ( $M=4.29$ ,  $SD=1.35$ ) and five stars ( $M=4.87$ ,  $SD=1.47$ ) for  $p<0.05$  ( $p=0.003$ ). This supported **H1** that a higher star rating would lead to a significant increase in the average recommendation value. For a rating of four stars ( $M=4.49$ ,  $SD=1.37$ ) versus five stars, there was not a significant effect for  $p<0.05$ , but was for  $p<0.10$  ( $p=0.077$ ). Otherwise there were no significant differences between test groups. **H2**, our hypothesis that an irrelevant tag ( $M=4.63$ ,  $SD=1.28$ ) would be no different than no tag at all ( $M=4.48$ ,  $SD=1.42$ ), was thus supported since there was no significant difference found between these two groups. As there were no significant differences between any of the credential tag groups, **H3**, that the expert tag would amplify the effect in comparison to an irrelevant tag, was not supported.

## 6. Discussion

### 6.1 Hypotheses and Implications

Our first hypothesis, **H1**, regarding the number of stars was supported since there was a significant increase in recommendation value between a rating with three stars and a rating with five stars. While three versus four and four versus five stars were not significantly different, there was still a positive trend from three to four to five stars. It is possible that the four star rating may have been unnecessary to witness the effect, since the difference between three to four and four to five stars was insignificant while the difference between three and five was. This finding was similar to (Flanagin et. al, 2011)'s finding that the higher the aggregate star rating, the higher the product's "perceived quality."

**H1** was partially intended as a benchmark to compare to the effect of the credential tag. In the design of the experiment, the number of stars on the rating was located right next to where the credential tag might be. The fact that the number of stars did make a difference proved that this manipulation did work.

For **H2**, we did not find a significant difference between adding an irrelevant credential tag and no tag at all, which is what we predicted. While this does not prove that the two groups are equivalent and that there is no effect at all, it does show that in the context of this experiment, adding the irrelevant tag did not have a statistically significant effect.

The most surprising part of this study is that our third hypothesis, **H3**, the crux of the experiment, was not at all supported. We believed that adding an expert tag would amplify the strength of the star rating versus an irrelevant tag. It was very surprising that there was no significant difference between any of the credential tag groups. It is possible that with more participants, a significant effect would be seen, but given that the p-value is 0.681 for the credential tag, it seems unlikely.

Our bigger question was whether indications of expertise on user reviews would affect e-commerce decisions, and how they would interact with the star rating of the review. Based on this study alone, we did not find a significant effect. This is particularly surprising since (Willemsen et al 2012) found that the identification of the reviewer did have significantly different outcomes. In our *Limitations* section below, we discuss possible reasons for why we did not observe a significant effect.

Despite the fact that we did not find significance in our third hypothesis, we believe that our research and study still may have implications for design. Our findings did not show that an expert review was more influential than a non-expert's review, but our literature research suggests that expertise may still be valuable. Further investigation may be able to find an effect, but at this point we at least know that including the tags would not have a negative effect and may indeed have more nuanced or hidden benefits, such as the user coming to feel more confident about customer reviews on that site.

However, even if a significant effect were found in our experiment, additional thought would need to be put into whether this is even an effect that should be implemented. For example, do experts tend to be more critical in their reviews? If so, and if consumers did place more weight on expert reviews, this could result in an average lower likelihood to recommend. Such a result could be detrimental to businesses looking to maximize sales, or beneficial to businesses looking to maximize customer satisfaction. Therefore, any significant effects observed in a study like this cannot be said to

necessarily mean that a business should implement expert tags; it could be evidence to support their exclusion for certain business models.

## **6.2 Experimental Limitations**

Despite careful efforts and considerations in our experimental design, there were inevitably limitations that resulted from the way the test was structured and administered. In measuring our dependent variable, we asked users to report the likelihood that they would recommend a product to a friend. This measure was used as a proxy for additional buying decisions (such as perception of product quality, purchase intent, etc.), but cannot be said to directly represent those other measures. Therefore, our dependent variable is an approximation of these other buying decisions, and we cannot say whether an effect in this study would necessarily impact sales or other consumer behavior.

One potential threat to our experiment's internal validity is that our pruned data may still have contained invalid data. While we eliminated responses from participants who failed the post-test screener, it is still possible some participants guessed the correct answer on the screener but did not actually read the review carefully and/or critically assess their response to the question. If this is the case, it adds noise to our results that we are unable to eliminate. Since our screening question had four answers, there was a 25% chance that a participant could guess the correct answer. This means that some of the data we used may have been from participants who did not truly read or engage with the study and instead just made random choices.

In some regards, the experimental setting was artificial and therefore may decrease the external validity of our results. Each participant was asked to make a decision based on a brief viewing of a product page mock-up, which is not necessarily a natural situation for making purchase decisions. Additionally, in this experimental situation, the participant had no real incentive to research the product as diligently or for as long as they might in real life.

When it comes to the participants involved in our study, MTurk limited us to an unknown demographic of people. We can say that our results hold for MTurkers, but not necessarily for the general population or any population in particular. It is possible that other groups would have responded differently to our scenario. This could diminish our external validity.

The display of only a single review also represents an artificial environment and potentially reduces external validity because a real consumer would be expected to browse multiple reviews (or at least have them available). Additionally, in our experimental conditions, we showed only one type of review (e.g. irrelevant) without any other types of reviews (e.g. expert) present. With more reviews, people may be able to disregard non-expert reviews and place more emphasis on expert reviews. Instead, in seeing only a single review, participants may have regarded the review as important (despite the reviewer's lack of expertise) since there was no other user-generated information to compare it to. A future study could look at the relative importance placed on reviews when more than one is displayed at a time.

Finally, we can't definitively say that subjects even noticed the profession tag. If we were to repeat this experiment, we could include a post-test question asking subjects if they noticed the tag, or quiz them on the contents of the profession tag.

### 6.3 Future work

In addition to efforts to reduce the above limitations, there are a number of possibilities for future research in this area.

Our literary research revealed a distinction between “search products” and “experience products”, defined as:

Search qualities are those that “the consumer can determine by inspection prior to purchase,” and experience qualities are those that “are not determined prior to purchase” (Nelson, 1974, p. 730)

For our experiment, we chose a search product (a digital camera) and explored the effects of the domain-relevant profession of “portrait photographer.” A future study could look for an effect among other search products, perhaps some less familiar to consumers. A specialized power tool, for example, may see consumers paying more attention to the reviewer’s credential because it is highly domain specific. A future study could also explore experience products whose qualities are more visceral and less apparent from the product description alone. For example, a product category like fine wine, whose nuances might be better described by experts than by average reviewers, may increase consumer reliance upon expert reviews. (Senecal et. al 2004) found that consumers were more strongly influenced by recommendations for experience products than for search products. Would a relevant credential tag, such as a “wine connoisseur” or “sommelier” have a significant effect on a fine wine review?

Additionally, our experiment used two specific tags (“account manager” and “portrait photographer”) that we arbitrarily chose to represent irrelevant and relevant expertise, respectively. We believe that the tags we chose were appropriate for our testing purposes, but a future study could look at additional tags that may more accurately reflect what users deem to be relevant or with more expertise for a given product.

In this experiment, we consciously decided to minimize the visual prominence of the expertise tag to more closely match how the tag might be introduced on a live website. However, this minimal treatment could have reduced our experimental effect if users overlooked it. A future study could look for an effect of a more visually prominent tag. Although a dramatic visual treatment could reduce the external validity of the results, it would help us determine whether the tags have any effect at all. From there, an optimal design might be reached.

As mentioned above, our experiment was limited by displaying a single review on the page. A more complex study could be run to explore the effects of an expertise tag when multiple reviews are competing for attention. For example, showing a 3-star expert review and 5-star non-expert review could result in a lower likelihood to recommend than showing a 5-star expert review and a 3-star non-expert review.

## 7. Conclusion

In conclusion, our study did not find significance regarding credential tagging. We did, however, find significance related to the star rating we displayed, which gives us confidence that our study was legitimate to some degree. It is likely that our specific credential tag design had a negligible effect on

our participants. Further permutations of this concept might yield more significant results, whether this means using a more prevalent tag design or working with a different type of product. This study only accounted for a 'search product', and by looking at 'experience products', it may be found that credential tagging indeed has a place in helping online consumers assess the quality of a potential purchase.

E-commerce continues to be a growing marketplace, and it is important that resources are devoted to creating an informative and satisfying experience for customers. We encourage other researchers to investigate how designers can intelligently mold review content to better accommodate the needs of online shoppers. As online shopping becomes ubiquitous, it demands an efficient and transparent design- usable and relevant user reviews are likely to remain a vital component.

## 8. References

- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data?. *Perspectives on Psychological Science*, 6(1), 3-5.
- Chevalier, Judith and Dina Mayzlin. "The Effect of Word of Mouth on Sales: Online Book Reviews." *Journal of Marketing Research* (August 2006).
- Flanagin, Andrew J., Miriam J. Metzger, Rebekah Pure, and Alex Markov. "User-Generated Ratings and the Evaluation of Credibility and Product Quality in Ecommerce Transactions." *Proceedings of the 44th Hawaii International Conference on System Sciences* (2011): n. pag. *IEEE Xplore*. Web. 22 Oct. 2014.
- Hu, Nan, Ling Liu, and Jie Jennifer Zhang. "Do Online Reviews Affect Product Sales? The Role of Reviewer Characteristics and Temporal Effects." *Information Technology and Management* 9.3 (2008): 201-14. *Researchgate.net*. Web. 20 Oct. 2014.
- Lee, Sheng-Hsien. "How Do Online Reviews Affect Purchasing Intention?" *African Journal of Business Management* 3 (2009): 576-81. *Academicjournals.org*. Web. 20 Oct. 2014.
- Liebermann, Y., & Stashevsky, S. (2002). Perceived risks as barriers to Internet and e-commerce usage. *Qualitative Market Research: An International Journal*,5(4), 291-300.
- Luhmann, N. (1979). *Trust; and, Power: two works by Niklas Luhmann*. Chichester: Wiley.
- Mackiewicz, Jo. "Reviewer Bias and Credibility in Online Reviews." *Proceedings of the 2007 Association for Business Communication Annual Convention*. Illinois Institute of Technology, Chicago. Association for Business Communication, 2007. 2007. Web. 20 Oct.2014.

<<http://www.businesscommunication.org/CMS/Resources/proceedings/2007annual/2007annual/06abc07.pdf>>.

- Mudambi, Susan M., and David Schuff. "What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com." *MIS Quarterly* 34.1 (2010): 185-200. Web. 20 Oct. 2014.
- Nelson, Phillip. "Advertising as Information." *Journal of Political Economy* 82.4 (1974): 729-54. *JSTOR*. Web. 21 Oct. 2014.
- Park, Do-Hyung, Jumin Lee, and Ingoo Han. "The Effect of On-Line Consumer Reviews on Consumer Purchasing Intention: The Moderating Role of Involvement." *International Journal of Electronic Commerce* 11.4 (2007): 125-48. *JSTOR*. Web. 20 Oct. 2014.
- Ruyter, Ko De, Martin Wetzels, and Mirella Kleijnen. "Customer Adoption of E-service: An Experimental Study." *International Journal of Service Industry Management* 12.2 (2001): 184-207. *Emerald Library*. Web. 19 Oct. 2014.
- Senecal, Sylvain, and Jacques Nantel. "The Influence of Online Product Recommendations on Consumers' Online Choices." *Journal of Retailing* 80.2 (2004): 159-69. *ScienceDirect*. Web. 20 Oct. 2014.
- Wang, Alex. "The Effects of Expert and Consumer Endorsements on Audience Response." *Journal of Advertising Research* (2005): 402-12. *Walsh.edu*. Web. 20 Oct. 2014.
- Willemsen, Lotte M., Peter C. Neijens, and Fred Bronner. "The Ironic Effect of Source Identification on the Perceived Credibility of Online Product Reviewers." *Journal of Computer-Mediated Communication* 18.1 (2012): 16-31. *Wiley Online Library*. Web. 21 Oct. 2014.
- Zhu, Feng, and Xiaoquan Zhang. "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics." *Journal of Marketing* 74 (2010): 133-48. *AMA Journals*. Web. 22 Oct. 2014.