

**ONLINE DECEIT: THE USE OF IDIOSYNCRATIC CUES IN IDENTIFYING
DUPLICITOUS USER-GENERATED CONTENT**

by

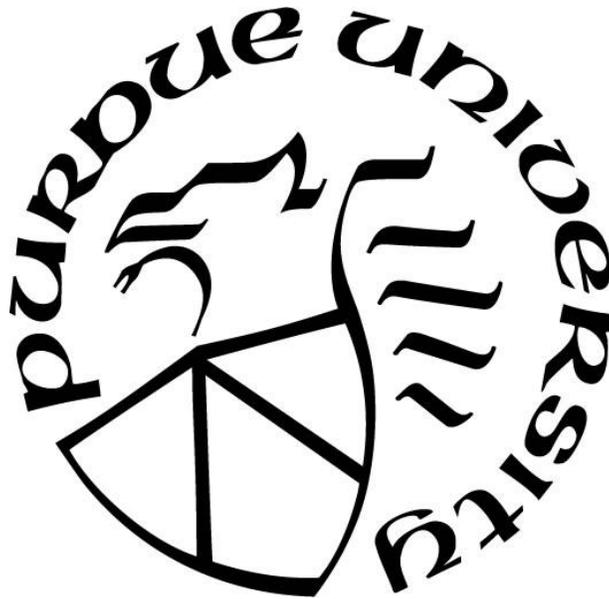
Christopher R. Roland

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy



Brian Lamb School of Communication

West Lafayette, Indiana

August 2019

**THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL**

Dr. Torsten Reimer, Chair

Department of Communication

Dr. Sabine Brunswicker

Purdue Polytechnic Institute

Dr. William Collins

Department of Communication

Dr. Stacey Connaughton

Department of Communication

Approved by:

Dr. Marifran Mattson

Head of the Graduate Program

For my family.

ACKNOWLEDGMENTS

The completion of this dissertation would not have been possible without the contributions of numerous individuals. I am grateful to my advisor and mentor, Torsten Reimer, who has consistently provided sound advice concerning the strategic development of the project throughout the research process. Further, the other committee members: Sabine Brunswicker, William Collins, Stacey Connaughton certainly merit recognition. Each committee member provided diverse and thoughtful insights that were crucial in honing the arguments advanced in the dissertation. In conjunction with Purdue acknowledgements, I would like to thank my friend and mentor, Tillman Russell for his advice and support throughout the development of the dissertation. I would like to issue a special thank you to my parents for their unconditional support for my education. Though they are too numerous to list, I would like to thank my friends and family for the many ways in which they have contributed to the development of my doctoral education.

TABLE OF CONTENTS

LIST OF TABLES	7
LIST OF FIGURES	8
ABSTRACT	9
CHAPTER 1. INTRODUCTION	10
Information Independence	14
Independence Assessment	18
Identifying Idiosyncratic Indicators	22
Current Research Overview	33
CHAPTER 2. STUDY 1	34
Participants and Design	34
Attribute co-occurrence.....	35
Text duplication	35
Timestamp similarity	36
Username similarity	36
Review set.....	36
Measures	37
Information independence.....	37
Procedure	37
Results.....	38
Discussion.....	39
CHAPTER 3. STUDY 2	42
Participants and Design	42
Review number	43
Measures	43
Message credibility	43
Purchase intention.....	44
Choice behavior	44
Procedure	44
Screening procedures	46

Results.....	47
Independence	47
Message credibility	49
Purchase intention.....	50
Discussion.....	51
CHAPTER 4. GENERAL DISCUSSION	55
Idiosyncratic Indicators as Online Detection Strategies	56
Independence and Online Choice Processes.....	60
CHAPTER 5. FUTURE DIRECTIONS	64
REFERENCES	69
APPENDIX A. SAMPLE REVIEW PAIR MESSAGE CONDITION.....	76
APPENDIX B. STUDY 1 SCENARIO INFORMATION.....	78
APPENDIX C. STUDY 2 SCENARIO INFORMATION.....	79
APPENDIX D. HYPOTHETICAL PRODUCT PAGE	80
APPENDIX E. SAMPLE STUDY 2 FEATURED REVIEWS	81
APPENDIX F. STUDY 2 INDEPENDENCE DESCRIPTIVES	82

LIST OF TABLES

Table 1: Intercorrelations, means, and standard deviations for dependent measures	50
---	----

LIST OF FIGURES

Figure 1: Mean independence judgments by indicator number 48

ABSTRACT

Author: Roland, Christopher, R. PhD

Institution: Purdue University

Degree Received: August 2019

Title: Online Deceit: The Use of Idiosyncratic Indicators in Identifying Duplicitous User-Generated Content

Major Professor: Torsten Reimer

The emergence of online information-seekers harnessing the aggregated experiences of others to evaluate online information has coincided with deceptive entities exploiting this tool to bias judgments. One method through which deceit about user-generated content can occur is through single entities impersonating multiple, independent content providers to saturate content samples. Two studies are introduced to explore how *idiosyncratic indicators*, features co-occurring between content messages that implicate a higher probability of deceit, can be used as a criterion to identify content that is not independently authored. In Study 1, analyses of a pairwise comparison of hypothetical reviews revealed that ratings of content *independence* were significantly lower when review pairs co-occurred in the attributes, text, and usernames compared to being heterogenous. In a high-fidelity experiment, Study 2 assessed if the effect of idiosyncratic indicators on independence is increased in the presence of multiple indicators, if it is attenuated with a high number of reviews, and if it impacts factors relevant to the choice selection process. As expected, the findings of Study 1 were replicated in addition to further revealing that the presence of multiple idiosyncratic cues yielded lower independence ratings. An interaction effect with idiosyncratic indicators and high review number was observed such that the effect of the former on independence was attenuated when there were a high number of reviews to obscure the presence of these indicators.

CHAPTER 1. INTRODUCTION

User-generated reviews are a readily available and frequently used source of information that has been shown to influence consumer choices online. When knowledge about an object or its source is low, people often pursue information from similar users that have experience with an object (Walther & Jang, 2012). The pooling of consumer information collectively or user-generated content (UGC) has been shown to be an important factor in influencing the attitudes and decisions of individuals online (Brown, Broderick, & Lee, 2007; Lee & Shin, 2014). In the domain of online choice, people may not always have enough information to interrogate the advantageousness of a product or the importance of an object's attributes. Though this information is potentially used frequently to help facilitate online decisions, it is often anonymously generated, high in volume, and unregulated which may render its veracity uncertain (Cheung, Luo, Sia, Chen, 2009). The accuracy of information online may be inconsistent because of the potential that exists for deceit or the misrepresentation of information (Donath, 2007; Ellison, Heino, & Gibbs, 2006; Flanagin & Metzger, 2007; Walther & Parks, 2002).

There are numerous examples of the presence of misleading information in the context of consumer choice. A recent investigation from an online store found that 5% of their reviews were falsified by entities that did not purchase their products (Anderson & Simester, 2014). Forrest and Cao (2010) documented multiple cases of online information misrepresentation including a textbook company providing Amazon giftcards for favorable reviews, an online marketing company providing a 'pay per post' service, and a PR firm providing false ratings of products provided by their clientele. While it is difficult to precisely identify the pervasiveness of the false presentation of online information, it is evident that it does occur with a measure of

regularity and can cause problems for consumers by skewing consumer perceptions of objects that review content references (Malbon, 2013).

Several explanations can intimate how and why entities are motivated to falsify online content. Review information on the internet is not always accurate but is at times strategically altered in the pursuit of self-benefit which can skew how the content is interpreted by online information-seekers (Brandt, Vonk, & Knippenberg, 2011). If there is a personal interest in presenting information, entities frequently alter their own impressions more compared to the impressions of others (Brandt, Vonk, & Knippenberg, 2009). One way that information can be distorted is through entities fabricating reviews by writing numerous reviews while purporting that the information is authored by different individuals.

This dissertation aims to describe how people judge the independence of user-generated information that is presented online. The ability to identify information as unique for outside influences should be impacted by a person's ability to identify information as authored by truly independent individuals providing unique observations about an object. In contrast to traditional media, the source of information online is often obfuscated (Sundar, 2008). Information can be transferred between sources or edited for content which can promote misinformation and skew the choices of deciders. Consequently, it is a valid pursuit to investigate what cues people utilize to evaluate if content has been produced by unique contributing individuals or *information independence*. When user content is falsely purported to be authored by numerous, discreet sources while being produced by singular, fraudulent entities, this represents a specific form of online *duplicity*. While prior research has assumed that the independence of UGC defines its utility, we sought to articulate the nature of this fundamental yet ill-defined property of online content. While there are conceivably numerous criteria by which people judge if information is

independent, it is unclear precisely what indicators impact independence perceptions, under what conditions this is likely to occur, and how these perceptions influence evaluations of UGC and the objects for which it advocates. This invites the question as to how people investigate if user-generated content is produced by independent sources and how this affects judgments about the objects that the content references.

. This investigation advances a novel model that describes the information that people attend to and illuminates how they utilize it to formulate judgments about the probability of deception in the context of UGC. It is reasoned that a cue-based decision-making approach can be used to predict what types of information are likely to prompt judgments of content independence when people formulate online choices. Moreover, we sought to provide novel evidence concerning key indicators which drive the judgments that user-generated content is authored independently. Through attending to the context in which user-generated content is provided, people can have clues to generate probabilistic inferences as to whether information is independent or conversely, judge that duplicity is likely occurring. Specifically, this study explores how idiosyncratic indicators can provide evidence to prompt suspicion that apparently independent content is authored by the same source and how this evaluation impacts the tendency for individuals to utilize this when formulating choices.

We argue that information-seekers often attend to cues which do not explicitly implicate that different content sources are colluding, but which suggest a high probability of duplicity based on certain cues co-occurring between content entries. These cues, here referred to as *idiosyncratic indicators*, can prompt judgments of information independence under certain conditions. We reason that there are two commonly used classes of cues which can prompt judgments of deceit about independence. Specifically, co-occurrence in content of messages and

source characteristics which are external to message content but describe characteristics of the contributor who produced it. Detection strategies include situations where content providers reference the same attributes of an object or duplicate text within a message (content cues) and publish content at the same time or with similar source usernames (source cues). This may suggest that duplicity is more likely and thus induce the judgment that information is not independent. However, it is further advanced that people generally assume by default that content is produced by independent sources. Consequently, idiosyncratic indicators should only prompt judgments that information is not independent when the indicators are accessible to observers. When this information is inaccessible, the default assumption of independence should be retained. It is reasoned that these idiosyncratic indicators likely exhibit multiple effects on the processes of online information evaluation and choice.

First, information independence and its importance to formulating judgments among user-generated content entries is explicated. Second, the role of online features in obfuscating information independence is discussed. Third, the utility of idiosyncratic indicators in detecting information duplicity by prompting the judgment that it is more probable is introduced. Fourth, it is argued that a judge's propensity to evaluate that duplicity is occurring depends on the accessibility of these cues in the judgment context where they interrogate content. Next, Study 1 proposed to test if individuals are sensitive to idiosyncratic indicators and if they shape impressions of independence. Finally, Study 2 tested if the proposed effect of idiosyncratic indicators on independence is attenuated when a greater volume of content obscures accessibility and if multiple indicators compound to impact judgments of duplicity.

1.1. Information Independence

User-generated content online has been widely viewed as an important and influential tool in evaluating information and helping to facilitate decisions online (Brown, Broderick, & Lee, 2007). The ability for non-affiliated web users to provide messages associated with an individual or object and affix them to content for an associated evaluation is a defining feature of user-generated content (Walther & Jang, 2012). Content produced by independent users has been shown to influence impressions on social media (Carr, Hall, Mason, & Varney, 2017; DeAndrea, 2012; Fox & Moreland, 2015), consumer attitudes about products (Doh & Hwang, 2009; Lee & Youn, 2009), and purchase intentions (Huang & Chen, 2006; Park & Lee, 2009).

The weight of UGC in evaluating claims online can be explained by the tacit notion that interactions tend to be truthful which is consistent across multiple approaches to human communication. Grice (1989) advanced that people prototypically engage in and assume others engage in communication which is cooperative. That is, people evaluate communication with the assumption that the purpose of a speaker's utterance is to be understood. More specifically, a feature of this principle is that people assume speakers try to avoid expressing things which are known to be false. This principle is mirrored in the principle of truth-default. This tenet suggests that unless suspicion of falsehood is activated, people assume automatically that communication is honest (Blair, Reimer, & Levine, 2018). Thus, the influence conveyed by UGC may be due to people assuming that it represents information provided by truthful others.

The principle of default truthfulness is instantiated in numerous assumptions information-seekers hold about UGC providers. One example is that individuals often consider this information as being produced by impartial peers with specific knowledge about the object that they reference (Walther & Jang, 2012). The participatory nature of UGC enables information-

seekers to incorporate the experiential qualities of people's knowledge and information about an object, which affords people to consult the individual experiences of others (Flanagin & Metzger, 2013). This knowledge or experiential expertise enables would-be consumers to evaluate the first-hand knowledge of large numbers of independent users (Flanagin & Metzger, 2008). Through assuming content providers have purchased or otherwise engaged with the object to which they refer, information-seekers can leverage this as evidence about the object to help guide their choices.

Consumers may regard user-produced information as more trustworthy than information that comes directly from the proprietor of a product or object that they reference, because other consumers and users are perceived to be independent (Walther, Van Der Heide, Hamel, & Shulman, 2009). People tend to strategically represent claims about themselves compared to how they are described by others (DeAndrea, 2014). Consequently, content which is viewed as being objective is more likely to be influential compared to information that is susceptible to outside influence. For example, when object proprietors have the ability to control which third-party content is observed, the third-party content tends to impact their impressions less.

One tacit assumption about UGC providers which has received notably less attention is that contributions represent discrete observations of separate individuals. The information produced by other users may exert influence because they are perceived by information-seekers to be unique observations of separate individuals which reference an object. When aggregated with other producers, these unique contributions may inductively impact impressions of objects and serve as an indicator of information credibility. This collective information provides authority through providing knowledge about judgments in aggregation (Flanagin & Metzger,

2013). Attending to aggregate information supposes that components represent autonomous contributions of separate individuals to garner influence.

While the expectation of individual observations serves as a seminal feature of user-generated content, the unique impact it has on impressions of user-generated content is ill-defined. We here refer to the perceived extent to which two or more user-generated messages of a single object represent unique observations of separate individuals as *independent information*. This may serve as an important and widely utilized means of assessing information which references objects in online settings. If people rely on the independence of information when assessing content to capitalize effectively on their individual, aggregated observations, individuals must have some confidence that user-generated content was produced by discreet providers.

In some situations, the extent to which UGC is independent from a producer of an object claim may not be immediately apparent, particularly in the context of digital media. Information may be intentionally or unintentionally obfuscated due to content being transferred across numerous users (Neubaum & Krämer, 2017; Sundar, 2008). A news article (for example) may be transferred from its original writer, to a news aggregator, before being presented to a user which can lead to numerous levels of a source with different degrees of information accuracy (Sundar & Nass, 2000).

The assumption of information independence is routinely violated in online settings in numerous ways. One means is proprietors of the objects can sometimes exert control over the user-generated content (DeAndrea, 2014). As some digital media platforms allow users to edit information, this has prompted the study of the perception that a source has altered third-party information. DeAndrea, Van Der Heide, and Easley (2014) found that perceptions of a source

editing third-party content affects judgments about the source and the utility of the third-party information. Alternatively, website creators may selectively delete or restrict user-generated content from sites they control, affecting what appears online (Walther & Jang, 2012).

DeAndrea, Tong, and Lim (2018) created a mock LinkedIn page and found that when website proprietors had the ability to edit user-generated content, the website's usefulness and the perceived employability of the target source was reduced. Thus, it is possible for the source of a claim to manipulate third party information which may conceivably support or refute the claim.

In addition to the potential for third party information to be edited, it is also possible for sources to mask or misrepresent their identity, thus making information which is produced by a source appear to be authored by independent users. Common characteristics of digital platforms can be used to obscure the identity of a source to serve a range of functions (Hancock & Guillory, 2015). For example, content may be purportedly produced by independent sources while being authored or manipulated by the provider of the original claim which stands the most to gain (Dellarocas, 2006). Researchers have differentially referred to the practice of masking source identity online as review and opinion spamming, masked marketing, or astroturfing, (Jindal & Liu, 2008; Lankes, 2008; Malbon, 2013; Sprague & Wells, 2010).

Despite variation in terms used for the concealment of a content producer's identity, this practice has presented challenges for information-seekers attempting to evaluate user-generated content. Recently, a study of an online clothing store suggested that 5% of product reviews were not authored by independent users who had experience with the product (Anderson & Simester, 2014). Further, using Yelp's detection algorithm as a proxy for fake product reviews, Luca & Zervas (2016) found that reviews that were detected as being duplicitous for about 16% of their reviews. This practice may be commonplace because independent users are viewed as more

influential in shaping impressions about an object than claims provided by the object's proprietor (Walther & Parks, 2002; Walther et al., 2009). Thus, the ability to obfuscate may enable individuals without first-hand knowledge of an object to falsely present themselves as representatives of online peers to exploit the assumption that viewers trust the observations of other users. This represents a specific form of deception which we refer to as *duplicity* whereby single entities purport to be multiple, independent users referring to an object through the provision of numerous content messages to fraudulently influence the impressions of information-seekers.

The extent to which duplicitous content providers are successful in impacting impressions is contingent on whether information-seekers judge the content they provide to be authored by independent users separately describing an object. For example, an overly favorable online review of a restaurant may be anonymous, prompting speculation that it was fabricated by the owner to attract customers. This notion was supported in a recent experiment that found when participants were unsure about the identity of online reviewers, the less the reviews impacted ratings of the object they described (DeAndrea, Van Der Heide, Vendemia, & Vang, 2015). As information-seekers become more uncertain that content is produced by independent users, this can influence how much the information impacts impressions about the object (DeAndrea, 2014). If the extent to which fraudulent content influences impressions is contingent on a person's judgment that it is independently authored, this invites an investigation into the conditions under which duplicity should be detected.

1.2. Independence Assessment

One approach which may intimate the strategies which individuals adopt in identifying duplicity online can be sourced from research on interpersonal deception detection. Informed by

research on deception detection, we propose a criterion by which the independence of information can be assessed. Specifically, observers may attend to the similarity between user content as a basis for judging what information is duplicitous and that which is truly independent. It is reasoned that by evaluating contextual cues surrounding the object claim and reference information, individuals make inferences about posting behavior that is atypical of independent sources and use this as a criterion for judging collusion.

How duplicity specifically and deception generally can be detected has a strong empirical tradition which can inform the diagnosis of duplicity online. To wit, a primary objective in this line of research has been to identify the cues that people utilize to diagnose the accuracy of claims to make inferences about the truthfulness of a speaker in an offline setting. A commonly argued belief in this research suggests that people are generally poor at detecting deception. A meta-analysis of 206 studies found that the average detection accuracy was 54% (Bond & DePaulo, 2006). A different meta-analysis on cues to deception detection supported this finding and argued that low detection accuracy is likely a result of deceptive and accurate claims exhibiting minimal behavioral differences (DePaulo et al., 2003). Due to these minimal differences, it is reasoned that perceivers have little material through which they can diagnose deception, thus accounting for low accuracy rates (Hartwig & Bond, 2011).

Recently, an alternative approach has been proposed which denotes a class of cues which can predict greater accuracy in detecting deception. Specifically, through examining contextual information surrounding communicators and comparing them against what is said, individuals have a basis for determining if deception is occurring. Blair and colleagues (2010) referred to this information as content in context and regarded this as a powerful tool for detecting deception which has numerous forms. It is argued that this serves as a basis not only for determining

situations in which people are prompted to suspect that deception is occurring but provides a criterion for distinguishing which messages are deceptive. This approach assumes that a judge's knowledge about a situation may be limited but it can nonetheless provide useful information to detect deception.

The content from which individuals potentially distinguish deception can be derived from three different types of information. One form is normative information or what is possible or likely in a situation such as a source's usual behavior, understanding of typical situations, or physical/natural constraints. Alternatively, individuals may compare statements to what is already known to identify explicit contradictions between content and prior knowledge.

We aim to explore the role of the final category of detection information: idiosyncratic knowledge. Specifically, information cues of idiosyncratic knowledge may indicate a higher probability that deception occurs by the co-occurrence of information that is peculiar to a source. For example, consider when a bank employee is interrogated about a theft which occurs regularly when they are present but stops when they are absent. Idiosyncratic knowledge does not explicitly indicate deception or offer contradictions but points to a higher chance of deception occurring (Blair et al., 2010). The means through which idiosyncratic knowledge prompts suspicion of deceit consists in individuals attending to co-occurring cues, triggering an if-cue, then-judgment inferential process. This is consistent with persuasion models of communication which treat the relationship of cues and judgments as probabilistic (Reimer, Hertwig, & Sipek, 2012). In other words, idiosyncratic cues can be utilized as uncertain but potentially valid indicators by which to infer deception.

This contextual information as a basis for detecting deception has yielded empirical support to bolster detection accuracy. Blair et al. (2010) tested the effect of numerous types of

contextual information on the accuracy of deception detection and found that the presence of this information yielded greater accuracy compared to its absence. Across several studies, accuracy was significantly higher when contextual information was provided (75%) compared to the control condition where it was not (57%). Consistently, Reinhard, Sporer, Scharmach, and Marksteiner's (2011) found people that were more familiar with contextual information were more accurate at formulating judgments about deception. Moreover, Park, Levine, McCornack, Morrison, and Ferrara (2002) asked people to recall when they detected a lie and found that contextual information was the most useful tool for detecting deception. Collectively, these findings provide empirical support for the suggestion that contextual cues serve as highly valid indicators to detect deception.

Much as contextual information can be used to make inferences about deception generally, it may be used as a highly valid indicator of detecting when content online is independent or duplicitous. In an online context, information-seekers consider user-generated content to be produced by objective, third-party sources (Walther & Jang, 2012). However, entities are sensitive to this assumption and duplicitous reviews are commonplace; funded by companies which even result in financial restitution (Malbon, 2013). If individuals assume that UGC represents independent observations of separate individuals, and duplicitous content is widely produced for exploitation, this constitutes a specific form of deception regarding the independence of information. What then is the process by which individuals can detect duplicity in the context of online information search?

Recently, Blair et al. (2018) proposed a model which provides a framework describing the process through which contextual information can prompt deception detection. The model suggests that individuals assume that by default, people think that statements are truthful until

there is otherwise reason to suspect deception. Accordingly, the model advances that people only engage actively in detecting deception when behavioral or contextual indicators trigger the suspicion of deception. Trigger situations are prompted when conditions suggest that there is a projective motivation to lie or misrepresent. The process of detection involves information search wherein a person searches and attends to critical information as a criterion for identifying deception.

1.3. Identifying Idiosyncratic Indicators

One straightforward way that these suspicion triggers may be expressed in online settings is by using the context in which user-generated content is provided to infer its independence. By comparing individual sources of content and searching for cues, individuals may be prompted to suspect duplicity. Specifically, we focus on the presence of idiosyncratic indicators in the form of two classes of cues. The first is by consulting text content through either observing the co-occurrence of attributes between content entries and duplication of text between messages. The second is through observing source cues: similarity of provider characteristics among UGC referencing the same object such as timestamp and username similarity. When there is a co-occurrence of content between entries and provider characteristics are similar, this may prompt an individual's assessment of unusual behavior and thus that there is a higher probability of duplicity occurring between the entries.

Consider a situation in which an individual observes an object claim produced by a restaurant with which they are unfamiliar. To verify the claim, the person may attend to user-generated content information in the form of review entries from former patrons of the restaurant. The context in which this reference information is sourced may impact judgments of independence. If entries reference different characteristics of the restaurant in the reviews, there

is little reason to suppose duplicity and this may serve as a basis for independence assessment. Conversely, if entries reference the same characteristics of the restaurant consecutively, this may prompt suspicion of duplicity as it would be highly improbable that independent reviewers would reference the same restaurant attributes in the same sequence. In this case, the attribute cues embedded within content entries prompts an inferential process whereby co-occurrence stimulates suspicion of deceit.

Research in online information search has provided some evidence to support that individuals are sensitive to the context in which UGC is presented and may use it as a basis for inferring information independence. For example, an approach has begun to theorize one means by which the potential duplicity of information sources can be assessed consists in the ability of the source of a claim to modify reference information. Adopting the perspective of Warranting Theory (Walther & Park, 2002), research has begun to explore how the capacity for sources to influence reference information affects attributions of its independence from the source of a self-claim (DeAndrea, 2014; DeAndrea & Carpenter, 2016). DeAndrea and colleagues tested restaurant reviews and found that when reviews came from the restaurants website, participants rated the likelihood of information manipulation higher and the participant likelihood of recommending the restaurant lower compared to reviews from an online platform which typically provides reference information (Yelp). Further, it was found that the extent to which the website had the capacity to control information impacted evaluations of the reviews being produced by objective, third-party sources.

Similarly, individuals rate sources who produce content as more trustworthy when large numbers of reference information is provided compared to merely the source's reputation or guarantees as it is easier for entities to influence the latter (Utz, Kerkhof, & van den Bos, 2012).

Further, people are aware of the motivations that sources of recommendations have to provide reference information rather than just the content (Sen & Lerman, 2007). This suggests that individuals attend not only to the content provided within reference information but the context in which it is produced. When the context in which reference information is provided departs from that which is anticipated of independent, third-party sources, this may trigger suspicions as to whether information is independent and thus questions about a claim's trustworthiness.

This body of research offers some insight in to how contextual information may prompt judgments of information independence. However, it speaks to judgments about the capacity of object proprietors to influence user-generated content which references it rather than the evaluation that multiple entries represent unique, independent observations of the object. It is conceivable that proprietors don't appear to exert influence over content entries but the entries themselves are duplicates and thus do not represent unique observations, while still influencing consumer judgments. It follows that additional explanation is needed to demonstrate how people form judgments about the independence of content entries and the cues which may prompt duplicity judgments.

Assessing the extent to which content is produced by discrete, independent sources online has emerged as a seminal pursuit across several disciplines. Data mining approaches have attempted to compile information about content and source features to distinguish independent from duplicitous content. For example, researchers have explored linguistic features like words and POS n-grams to identify differences between legitimate and false content. Alternatively, researchers have studied source cues, information embedded within the review such as reviewer username, date of posting, geolocation, and approval rating. Further, attributes of the target

product such as product description, price, or sales volume have been tested distinguish user types of content (see Tavakoli, Heydari, Ismail & Salim, 2015 for review).

These data-mining approaches have suggested numerous strategies for identifying duplicity in UGC. Liu (2007) advanced that duplicitous reviews may be detected through producing evaluations of different products at roughly the same time. It is reasoned that this behavior is atypical for most content producers and thus may serve as an indicator for duplicity online. Research has begun to explore how duplicate and near duplicate user-generated content can be identified (Ma & Li, 2012). Using language models, some scholars have attempted to duplicates of reviews using natural language processing. Patil and Bagade (2012) proposed an n-gram language model and review spamming of brands using the features describes between reviews of the same product. This is derived from the assumption that most fake content providers attempt to be economical when producing content in volume.

How idiosyncratic indicators serve as a basis to prompt duplicity suspicion online consists not only in evaluations of object proprietors but in the co-occurrence of information between content entries. This follows from the observation that the production of numerous duplicitous reviews with heterogenous content is extremely costly and prohibitively time-consuming to enact. Consequently, duplicitous content providers rarely generate different false entries but instead frequently duplicate content from existing fake entries for the same object. Judges may interrogate duplication in the textual content and similarity of the concepts emphasizes as a basis for distinguishing duplicitous from truly independent content providers (Heydari, Tavakoli, Salim, & Heydari, 2015). If individuals observe the co-occurrence of content, it stands to reason that this may function as cues that prompt suspicion of a higher duplicity probability.

While this body of research involves a promising explanation for how duplicitous content can be detected, it does not describe the strategies which individuals adopt in judging the independence of information. Tavakoli and colleagues (2015) reason that this is due to the perceived difficulty in identifying deception among UGC through manually reading content between reviewers. Rather, the function of automated approaches is to provide a more comprehensive basis for evaluating fake content despite the practice still being commonplace (Heydari et al., 2015). Thus, while the co-occurrence of content advances a plausible explanation for detecting duplicity in UGC, the extent to which individuals are sensitive to these differences warrants empirical inquiry.

Information-seekers online assume by default that user-generated content entries represent discrete observations from different content providers with experiential expertise. This assumption should be violated when contextual information provides reason to suspect deception (Blair et al., 2018). When information online points to the high probability of deception occurring, this should trigger suspicion that deception is occurring and that content entries are duplicitous. As duplicitous content providers attempt to be economical when falsifying UGC, they may duplicate both a. the textual content and b. the attributes of the referencing object between content entries. When different content entries duplicate the text or object attributes which they reference, this may trigger the evaluation that there is a higher probability that content is duplicitous and thus that entries are not independent. In contrast, if content entries provide different text or reference different attributes of an advocated object, there should be no suspicion of duplicity should not be prompted and thus the default judgment that entries are independent should be retained.

In addition to duplicating attributes in different content entries, false providers may be distinguished from unique providers based on scrutinizing the characteristics surrounding content providers. Liu (2007) advanced that duplicitous reviews may be detected through producing evaluations of different products at roughly the same time. It is reasoned that this behavior is atypical for most content producers and thus may serve as an indicator for duplicity online. Similarly, Liu (2012) advanced that a high posting frequency within a certain time interval may be used by spammers. Studies which looked at high content production frequency in a short amount of time were highly accurate at detecting duplicitous content (Heydari et al., 2015).

It follows that the similarity of source cues such as production time and source name similarity may serve as a cue for detecting duplicitous content. As a high content production frequency in a short time interval is a common behavior of false content producers, this may signal unusual behavior to judges evaluating UGC. Consistently, if reviews have similar source names, this may suggest atypical behavior of independent users as the probability of near duplicate usernames is low. If judges are sensitive to this tendency, a high production frequency or similarity in producers' usernames should indicate that duplicity is occurring and thus that independence cannot be assumed. In contrast, different content production times or dissimilar source names should offer no indication that duplicity is occurring, and judges should retain the default judgment of information independence.

Hypothesis 1: Judgments that content is independently authored will be lower when there is the presence of a. attribute co-occurrence, b. text duplication, c. timestamp similarity, d. source similarity between different content entries compared to its absence.

While there are conceivably numerous means through which contextual information can be utilized to detect duplicity, these cues may also compound to strengthen duplicity judgments. This is derived from the assumption that contextual information can increase deception detection accuracy in an additive fashion. Specifically, as there is an increase in the volume of available information which may prompt suspicion, deception detection accuracy increases. This is supported by Blair et al. (2018) who tested this in a mock crime scene scenario and found that as suspects provided details which were inconsistent with known facts, they were perceived to be less truthful. This suggests that as indicators are added, this accentuates the probability that deception will be judged. It follows that in the presence of multiple idiosyncratic indicators, people should become more attuned to a higher probability of duplicity compared to one or no indicators, thus reducing evaluations of content being independently authored.

Hypothesis 2: When there are multiple idiosyncratic indicators present, independence judgments will be lower compared to when there are single or no idiosyncratic indicators.

Whereas idiosyncratic indicators may prompt suspicion that duplicity is occurring while observing UGC, this may not hold in all situations. Among many choice situations, deciders may face high levels of uncertainty about information, time constraints, and limited cognitive capacity in evaluating objects (Gigerenzer & Kurz, 2001). These situational constraints may render it more difficult for people to attend to idiosyncratic indicators and thus their ability to utilize them in judgments. This lack of attention to these cues may affect the inferences which people glean about UGC and the extent to which this information is utilized to evaluate the objects that they describe generally.

As choice situations may entail high uncertainty, too much information to process, and limited time to integrate it, deciders may adopt simple decision rules to operate effectively in their environment. These mental shortcuts, or *heuristics* can serve as useful tools to formulate approximately accurate judgments with minimal effort by exploiting features of the environment in which individuals are embedded. A large body of research has found that applying these simple decision rules can serve as adaptive responses which can be more accurate at predicting a criterion compared to more cognitively expensive strategies (Gigerenzer & Goldstein, 1996; Reimer & Katsikopoulos, 2004).

One judgment rule which can explain a person's propensity to assume truthful behavior in the communication of others is the default. The default choice or judgment is one that is selected by doing nothing given the features of the choice environment. There are several explanations as to why defaults can impact judgments and decisions. One is that by having a default within the choice context, this serves as an implicit recommendation which is activated by individuals through observing the behavior of others (Johnson & Goldstein, 2003). Alternatively, individuals may select a default as it requires less effort compared to switching between alternatives (Huh, Vosgerau, & Morewedge, 2014).

Building on insights about defaults in other areas of judgment and decision making, Blair et al. (2018) propose a truth default. According to their model, people assume by default that communication is honest unless otherwise prompted to suspect deception. This represents a cognitively inexpensive, automatic process which people incorporate while investigating human communication. The default is retained when a person either does not actively examine the potential for deception or doesn't find inculcating evidence to prompt suspicion of deception

occurring (Blair et al., 2018). Thus, a more deliberative process of suspicion is only triggered when judges become aware of the possibility of deception occurring.

Some empirical support for the truth-default assumption comes from the reliable observation that participants in deception studies have a strong tendency to judge most claims to be truthful (e.g., see Levine, Park, & McCornack, 1999). Moreover, studies often prime individuals to pursue deception judgments thus attuning them to the potential for deception and inflating the default tendency. One explanation for the propensity to assume truthful communication is that doing so is an adaptive response due to most people engaging in truthful communication in many situations (Blair et al., 2018).

While there is the potential for this default to be challenged when suspicion is triggered by inculcating information, this should only hold when people can reasonably attend to these cues. Individuals may not incorporate these indicators into their judgments when their choice environment is saturated with information as this environmental noise renders it difficult to integrate cues into their judgments. Novel communication technologies which afford the capacity to evaluate information quickly and accessibly has contributed to this saturation of information (Bawden & Robinson, 2009). In an online choice context, the high information volume available can render excess noise and lead to more difficult judgments. Lee and Lee (2004) demonstrate that the attribute number and attribute level distribution serve as robust predictors of information being saturated in a choice environment. Thus, the more available information within a choice context, the lower the probability of information being integrated in to judgments.

The amount of total information available within a choice environment can impact the extent to which people evaluate idiosyncratic information. In the context of online deception detection, a higher proportion of alternative information compared to idiosyncratic indicators

may reduce the likelihood of judges attending to the latter. It stands to reason that this higher noise-to-signal ratio may obfuscate the indicators which would otherwise prompt suspicion of duplicity occurring. When individuals do not become aware of or attend to idiosyncratic indicators, suspicion of duplicity should not be prompted, and the assumption of truthfulness should be retained. In contrast, if there is a comparatively smaller proportion of alternative information to idiosyncratic indicators, the probability of them triggering attention should be higher and suspicion of duplicity more likely.

***Hypothesis 3:** The effect of idiosyncratic indicators on information independence judgments will be attenuated when there are many reviews available compared to when there are few.*

In conjunction with impacting independence ratings, it is plausible that the presence of idiosyncratic indicators could influence other dimensions associated with online choice selection processes. The extent to which information is judged to be authored by independent sources who provide discreet observations about an object is likely to have implications on other judgments associated with the product being reviewed. When information prompts suspicion that independence of content has been violated, this may reduce credibility evaluations of UGC messages. Specifically, in the absence of first-hand knowledge, people are likely to attend and adopt the attitudes of others and rate their claims as valid. Consistently, user-generated content can exert influence on the attitudes of observers toward the object that they reference.

Recommendation systems such as UGC have been shown to exert strong influence on the attitudes on information-seekers online. For example, user-generated ratings have been found to influence observer ratings about movies (Cosley, Lam, Albert, Konstan, & Riedl, 2003). As this content is assumed to be authored by independent users, information-seekers are likely to attend

to and be influenced by information about the objects that the content describes because they rate it as having a high degree of veracity or accuracy. When idiosyncratic indicators are present and suggest a violation of this assumption, it is tenable that credibility ratings of message content are diminished.

In addition to influencing impressions about the object claims, independence violations of user-generated content could influence purchase intentions and choice behavior. Particularly among research on online consumer purchase decisions, the content and ratings of user reviews has been found to influence the likelihood that people will select a product or purchase intent (Flanagin, Metzger, Pure, Markov, & Hartsell, 2014). Consistently, online consumer research has been shown that the suggestions advanced by consumer ratings have been associated with an increase in product sales (Chevalier & Mayzlin, 2006). The weight of user-generated content in influencing these outcomes is due to the ability to aggregate experiences of individual observers (Flanagin & Metzger, 2013). If aggregated content is shown to be saturated with single entities claiming to be multiple sources, their effect on purchase intention and choice behavior may be attenuated. It stands to reason that the presence of idiosyncratic indicators may prompt suspicion of duplicity and thus reduce purchase intentions or the likelihood of formulating a choice.

The extent to which duplicity is detected and thus independence of content determined should impact the message credibility, purchase intention, and choice behavior about the object that the user-generated content references. If attitudes about message credibility, purchase intention, and choice behavior are influenced by independent users, then the utility of the information produced by content providers should hold insofar as they are perceived as being independent observers. If idiosyncratic indicators implicate that this assumption does not hold, it follows that the experiential expertise of the providers will not be perceived and the ability for

the content information to influence attitudes about claim veracity, purchase intention, and choice behavior should be attenuated.

***Hypothesis 4:** 1. Message credibility, 2. purchase intention, and 3. choice behavior will be lower when there is the presence of a. attribute co-occurrence, b. text duplication, c. timestamp similarity, d. source similarity between different content entries compared to its absence.*

1.4. Current Research Overview

To test the nature of idiosyncratic indicators and their impact on independence judgments, we advanced two studies which sought to test the proposed effects in different ways. Study 1 examined whether people attend to each cue discreetly when prompted to examine review pairs and if independence judgments are lower when they are present (H1). To test this, participants were asked to review multiple pairs which were systematically varied to exhibit consistency along each indicator for some pairs and for other review pairs to be heterogenous. To further investigate this concept, Study 2 examined whether the co-occurrence of multiple indicators simultaneously impels lower evaluations of independence than the presence of single or no indicators (H2). Further, this study tested if the presence of more reviews moderates the proposed effects of indicators on judgments by obscuring the attention to their presence (H3). Finally, Study 2 explored if idiosyncratic indicators reduce evaluations of message credibility, purchase intention, and the propensity to formulate a choice (H4). We proposed an experiment to test the effects of different indicator combinations and high/low review numbers in a task which closely approximated consumer judgment situations.

CHAPTER 2. STUDY 1

The present study explored if individuals are sensitive to idiosyncratic indicators which are embedded within user-generated content that describes an e-commerce product. Varying the degree to which different hypothetical product reviews exhibit attribute and text co-occurrence as well as timestamp and source name similarity, this study aimed to sequentially test the degree to which participants attend to each of these indicators. This study tested Hypothesis 1 through assessing if these indicators impact participant judgments as to whether they are authored by independent sources or represented duplicitous content.

2.1. Participants and Design

Two hundred and two participants (95 male, 106 female, 1 unspecified; mean age, $M = 20.23$, $SD = 1.43$) were recruited from a large, Midwestern university for which they received course credit as compensation. The design entailed a repeated measure, pairwise comparison judgment task between pairs of reviews. The study consisted of the within-subjects factor *Indicator* (attribute co-occurrence, text duplication, timestamp similarity, username similarity, mismatched), with the mismatched level inhering an aggregation of 8 randomly assigned, heterogenous review pairs. The study further included the between-subjects factor, *Review set* (Set 1, Set 2, Set 3, Set 4) comprised of differing, randomly assigned combinations of cues within each review. The experiment involved participants looking at twelve pairs of hypothetical product reviews and making a judgement about information independence following each comparison. Four pairs of reviews (eight reviews in total) were systematically varied to exhibit co-occurrence along each of the idiosyncratic indicators (attribute co-occurrence; text duplication; timestamp similarity; source similarity). Each review pair co-occurred on a single indicator and were controlled for others such that there is no parity along other indicators. The

remaining eight pair comparisons varied systematically by randomly providing pairs of reviews which did not exhibit co-occurrence along any indicator and aggregated to a single mismatched pair judgment. The order in which participants received review pairs was randomized.

2.1.1. Attribute co-occurrence

Hypothesis 1 predicted that if the attributes referenced in a review are the same as those in another review, the reviews should be judged to be less independent compared to reviews that reference different attributes. Attribute co-occurrence consisted of whether the product features described between review pairs were the same or different. If these attributes were the consonant, the sampled reviews referenced the same attributes in tandem. Pairs with different attributes provided dissimilar characteristics of a focal object between the review producers. Attributes were described in a single sentence within a four-sentence review for all review pairs (see Appendix A for sample review pair messages).

2.1.2. Text duplication

The extent to which the text is the same across multiple reviews may impact the judgment that information is not independent. Text duplication consisted of whether the content is the same between multiple reviews. Text duplication occurred when the text within a review other than the object attributes was the same across review pairs. The absence of text duplication occurred when text within the review other than the object attributes are different among the review pairs. Text duplication varied along three of the four sentences within each review message. The introduction sentence was completely replicated between review pairs whereas the remaining sentences exhibited the same content and sentence structure but varied slightly in word choice.

2.1.3. Timestamp similarity

Timestamps consisted of the relative disparity in production time associated with different review messages. Timestamp similarity was exhibited when the production timestamp between two reviews was identical. Conversely, the absence of timestamp similarity occurred when the timestamp between review pairs was staggered thus indicating variation in post time.

2.1.4. Username similarity

Username similarity was manipulated in the following manner. Source name similarity entailed the usernames affixed to each review being the same across review pairs with minor variations in the numbers at the end of each username (e.g. “John 02; John 04). The lack of source name similarity occurred when the usernames will be completely different with no similarity in their name content.

2.1.5. Review set

There was the potential for the combinations of review content to interact with co-occurring indicators and thus impact independence judgments. To counteract this possibility, participants were randomly assigned to one of four review set combinations. Each set consisted of eight cue variations per indicator. Consonance was exhibited among two of the eight variations to denote the presence of attribute co-occurrence, text duplication, timestamp similarity, and username similarity respectively whereas the remaining six variations were heterogenous. The four different sets of review pairs exhibited different combinations of mismatched content such that each of the 8 reviews contained a differing amalgam of attributes imbedded in the text, review descriptions, timestamps, and usernames. These sets differed by randomly assigning combinations of these varied cues such that no review contained the same

amalgam across sets. For example, Review 1 would contain different text, timestamps, and usernames in Set 1 compared to Set 2.

2.2. Measures

2.2.1. Information independence

This consisted of the extent to which participants judged that two or more content entries were authored unique sources. Information independence was assessed using three items measured on 7-point scales with endpoints ranging from strongly disagree (1) to strongly agree (7). Items included, “Separate individuals created the reviews”; “People created the reviews independently”; “Different people authored the reviews that I read”. This measure was assessed following each review pair and aggregated to form a complete measure of information independence (all $\alpha > .89$).

2.3. Procedure

Participants were informed that they would read different product reviews and subsequently asked to evaluate them (see Appendix B for scenario information). A paired comparison task systematically varied eight distinct reviews and exhibited different review pairs for a total of twelve comparisons, each time prompting participants to judge the perceived independence of the pair. Four of the twelve total comparisons co-occurred along one of the idiosyncratic indicators while being different along all other indicators. The mismatched pairs were systematically varied such that single reviews were presented multiple times but were paired with a different review for each judgment. After the presentation of each pair, participants evaluated whether they perceive the reviews to be authored by discrete sources (i.e. independence). The order of presentation for these pairs was randomized across participants.

2.4. Results

We first tested the hypothesis that independence judgments should be lower when there is consonance in idiosyncratic attributes among review entries compared to dissonance. This was conducted by examining differences in independence judgments aggregated from mismatched review pairs (comprised of 8 heterogenous review pairs) and pairs which exhibit similarity in a. attribute co-occurrence, b. text duplication, c. timestamp similarity, and d. username similarity.

A Two-way, Mixed ANOVA was conducted to test differences in information independence judgments for each indicator with *indicator* as a within-subjects factor and *review set* serving as a between-subjects factor. The results showed that there was a significant main effect of indicator on independence, $F(4, 49) = 64.15, p < .001$, partial $\eta^2 = .291$. Post hoc tests using the Bonferroni correction revealed that independence judgments for the mismatched pairs ($M = 5.46, SD = .88$) were significantly higher compared to attribute co-occurrence ($M = 4.60, SD = 1.70$), text duplication ($M = 3.08, SD = 2.03$), and username similarity, ($M = 4.39, SD = 1.89$). In contrast, no significant difference between mismatched pairs and timestamp similarity was identified ($M = 5.38, SD = 1.43$). Thus, it was found that independence judgments were lower than heterogenous pairs for attribute co-occurrence, text duplication, and username similarity, but not for timestamp similarity. In contrast to the indicator factor, no significant main effect of review set was observed, $F(3, 49) = 2.16, p = .1$, partial $\eta^2 = .03$. Further, no interaction of indicator and review set was found, $F(12, 49) = 1.11, p = .35$, partial $\eta^2 = .02$.

To test if there was a significant difference in independence judgments between idiosyncratic indicators collectively and mismatched pairs, a one-sample t-test was conducted. An effect of aggregated idiosyncratic indicators was observed, $t(198) = 57.66, p < .001$. Independence judgments were significantly lower when idiosyncratic indicators co-occurred (M

= 4.36, $SD = 1.07$) compared to mismatched, heterogenous pairs ($M = 5.46$, $SD = .88$). Taken together, these results suggest that partial support for Hypothesis 1 was observed.

2.5. Discussion

If idiosyncratic attributes can serve as a basis through which deception of duplicitous content can be detected, it follows that there must be evidence that individuals are sensitive to these cues and are thus likely to utilize them in formulating independence judgments. The principle goal of Study 1 was to explore if individuals are cognizant of these indicators when they co-occur between reviews. This represents an important first step in describing the cues that individuals utilize to form independence judgments of content. A pair-wise comparison judgment task was implemented wherein review pairs were systematically varied to exhibit consonance along each indicator in tandem and compared to eight heterogenous pairs to explore differences in independence judgments.

Consistent with Hypothesis 1, Study 1 revealed that judgments of information independence were significantly lower when review pairs exhibited consistency along attribute co-occurrence, text duplication, and username similarity compared to when review pairs were different. Though independence ratings were marginally lower when reviews exhibited timestamp similarity, this was not significantly different to ratings of heterogenous review pairs. Despite this negligible difference overall it was found that evaluations of independence were significantly lower when idiosyncratic indicators were present compared to heterogenous pairs. Further, it was found that there was a large effect size of indicator presence overall. This suggests that idiosyncratic indicators clearly albeit varyingly prompt lower judgments of information compared to heterogenous content.

As paired and unpaired content attributes may be confounded, four different sets of review combinations were created. We next explored if these different combinations varied systematically by testing if they impacted independence judgments or interacted with the effect of indicators on independence judgments. As expected, it was identified that the varying combinations of review content did not drive evaluations of review independence. These results suggest that the impact of idiosyncratic indicators on independence evaluations is not contingent on the content which is peculiar to a certain review. Rather, it implicates the information content is substitutable such that the co-occurrence of content is the primary trigger of reduce independence judgments compared to different pairs. This provides some evidence for the robust effect that co-occurrence of cues can have on the likelihood that duplicity is present between review pairs. Collectively, this provides strong support that individuals attend to idiosyncratic indicators, that many of them prompt lower independence judgments compared to dissonant pairs, and that this effect holds even when these cues are interchanged between reviews.

The findings affirm that each of these idiosyncratic indicators can be detected in isolation and that their co-occurrence can serve as cues to prompt suspicion that UGC may not be the product of independent observations. This reinforces prior research on user-generated content which posited that e-commerce algorithms could detect duplicity when purportedly different information contributions exhibit consistency along attributes of review content. Importantly, this provides new evidence that people are sensitive to individual co-occurring attributes even as other features of review content are varied.

While establishing that individuals are sensitive to idiosyncratic indicators when attending to review content is an important first step, further exploration of how they impact independence judgments is warranted. With the attention of idiosyncratic indications in

evaluating UGC, it begs exploration in to how these cues are utilized to formulate inferences about deception judgments and its implications for attributions of content credibility.

Furthermore, if user-generated content is useful to users because it entails an aggregation of information from numerous content providers, it stands to reason that a nuanced understanding of the processes underlying independence judgments requires exploring how they function when a multitude of reviews are present.

CHAPTER 3. STUDY 2

In this study, we investigated if the presence of multiple idiosyncratic indicators compounds the capacity for people to detect duplicity (H2). Further, this study tested if the high-information context where judges are most likely to encounter UGC impacts independence judgments by concealing idiosyncratic indicators among a higher volume of reviews (H3). Finally, this study explored whether the presence of idiosyncratic indicators can impact other dimensions associated with consumer choice (H4). If people assume by default that user-generated content is independent, then their duplicity judgments are contingent on attending to and utilizing cues to formulate these observations. When an observer's capacity to identify idiosyncratic indicators is constrained by increasing the number of reviews available, their ability to detect duplicity may be inhibited. This study advances a highly valid choice scenario by closely replicating a typical e-commerce situation in which people are likely to encounter idiosyncratic indicators.

3.1. Participants and Design

Study 2 sampled 886 participants (533 male, 340 female, 13 undisclosed; mean age, $M = 33.53$, $SD = 17.92$) from Amazon's Mechanical Turk. Participants received \$1.00 toward their account for a task taking approximately 10-15 minutes to complete. Consistent with prior social scientific research, the selection criteria for drawing Mechanical Turk participants was 96% approval rating to determine that individuals have a record of providing quality responses (Buhrmester, Kwang, & Gosling, 2011; Paolacci, Chandler, & Ipeirotis, 2010). The study design consisted of: 2 (attribute co-occurrence: with vs. without) X 2 (text duplication: with vs. without) X 2 (timestamp similarity: with vs. without) X 2 (source similarity: with vs. without) X 2 (review number: high vs. low) between-subjects design. Participants evaluated a series of alternatives,

attribute information, and review information before choosing between one of two alternatives per choice task for a total of two choices. Participants completed 2 hypothetical consumer judgment tasks for which the order of presentation was randomized.

3.1.1. Review number

Study 2 replicated the manipulation of attribute co-occurrence, text duplication, timestamp similarity, and source username similarity used in Study 1 whereby the similarity or difference between review features determines the presence of an indicator. Differently, indicators in Study 2 were manipulated between-subjects such that multiple indicators could co-occur within review pairs and an additional factor was added.

Review number. As independence judgments may be impacted by the ability to detect indicators which may be obfuscated with more information, the number of reviews was varied. In the low-review number condition, a single review pair was provided for each alternative. In contrast, the high-review number condition consisted of 10 reviews per alternative while only two had the potential to co-occur. In the 10 review, with-idiosyncratic indicator conditions, only two reviews exhibited attribute co-occurrence, text duplication, timestamp similarity, or source name similarity. The remaining 8 reviews were always heterogenous and the order of presentation for all reviews was randomized.

3.2. Measures

The measures in this study replicated those used in Study 1. The independence measure was assessed once following the completion of the judgment task and preceding all other measures.

3.2.1. Message credibility

If cues can trigger judgments that different pieces of information have independent sources, independence may serve as a reliable predictor that information is accurate. We

measured message credibility based on Appelman and Sundar (2016) measure of the extent to which people consider the content of communication is accurate. Participants were asked to rate the extent to which reviews are judged to be “accurate”, “authentic”, and “believable”. This was assessed on a 7-point, Likert-type scale with endpoints ranging from “strongly disagree” to “strongly agree”. Items were aggregated to form a complete scale of message credibility, ($\alpha = .89$). These items were aggregated to form a complete index of perceived message credibility.

3.2.2. Purchase intention.

The likelihood that an individual will select an object or purchase intention was adapted from Putrevu and Lord (1994) on a 7-point, Likert-type scale ranging from “strongly disagree” to “strongly agree”. Five items included “I would consider buying this product”; “I have no intention to buy this product (R)”; “It is possible that I would buy this product”; “I will purchase (brand) the next time I need a (product)”; “If I am in need, I would buy this (product)”. These items were aggregated to form a completed measure of purchase intention, ($\alpha = .85$).

3.2.3. Choice behavior

As independence may serve as an indicator that information is accurate, this may affect a person’s likelihood of being influenced to select the advocated object. Participants were prompted to indicate if they would select one of the two object alternatives being advocated in the reviews. This task entailed the ability to select one of the two advocated products or refrain from selecting any product.

3.3. Procedure

The general procedure for this study entailed 1. presenting participants with hypothetical electronic product descriptions, accompanying product reviews and selecting an alternative then 2. evaluating the product and review information. First, participants read background information

indicating that they would be selecting two sets of products (see Appendix C for scenario information). Participants were prompted with the information that “You should consult and evaluate product information and reviews to ensure that you make an informed opinion about what is best for you”. This entailed asking participants to review product information and select one based on their preferred choice alternative or to refrain from making a choice. This procedure was replicated for a total of two different consumer products (digital camera, color laser printer) and the order in which tasks were completed was balanced across participants.

The specific operation for each task entailed participants being directed to an artificial consumer purchase website. Participants were directed to a hypothetical e-commerce website and informed that the search had been narrowed to two 2 different but similar choice alternatives which meet their criteria. In advance of forming a selection, participants had the option to select and evaluate each alternative by being directed to a new web page which contained an image of the alternative, the cost of the product, a set of key attributes, and reviews corresponding to each alternative see (Appendix D for hypothetical product page). To access the reviews, participants were directed to a separate page containing featured reviews (see Appendix E). This review page varied systematically along the factors with participants receiving a single review condition. The attributes of each alternative were varied systematically such that each alternative will have 50% unique and 50% shared attributes with the other alternative and all attributes randomly distributed to alternatives.

After reviewing the available alternatives, participants were asked to select one of the alternatives or not make a choice if no alternative was acceptable. Participants were then asked to respond to several decision measures including first whether participants believed that the

reviews were independent from each other, the accuracy of the online reviews, and their intent to purchase one of the products.

3.4. Screening Procedures

As Study 2 required that participants carefully evaluated review information to assess information independence, several checks were implemented to ensure the stimulus was attended and satisficing minimized. First, participants were required to have observed the pages containing review information for enough time to process the stimulus material. Excluded data were computed by aggregating the response latency of pages containing review information across both judgment tasks and removing participants who fell two standard deviations below the average response time.

Checks on participant attention were implemented to attenuate satisficing in the sample. Past behavioral science research has concluded that some participants may not sufficiently attend to study materials which can inflate noise and substantially decrease statistical power (Oppenheimer, Meyvis, & Davidenko, 2009). Participants were asked to describe the products (digital camera, laser printer) that were being evaluated by the reviews. Based on this prompt, 94% of participants were able to correctly identify the advocated products. Finally, an attention filter question was introduced at the end of the instrument to confirm participant attention and avoid automated responses based on recommendations of prior communication research utilizing Mechanical Turk data (Sheehan, 2018). Specifically, participants were given a paragraph of text detailing the importance of their attention and then instructed not to select any items on a later question. If participants selected any of the items, this was used as a criterion for exclusion.

These three indicators functioned as criteria through which data were excluded from the analyses. The checks identified 85 participants who failed to meet one or more inclusion criteria

with most of these participants failing to meet multiple checks (74%). The remaining sample yielded 801 participants meeting these criteria which were retained for the analyses.

3.5. Results

3.5.1. Independence

The hypothesis of the additive effects of idiosyncratic indicators was conducted with a simple regression analysis. A variable denoting the number of idiosyncratic indicators ranging from 0 to all 4 predicts independence judgments. The number of indicators significantly predicted independence judgments, $b = -.39$, $t(799) = -6.80$, $p < .001$. The number of indicators explained a significant proportion of variance in independence scores, $R^2 = .055$, $F(1, 799) = 46.29$, $p < .001$. On average, participants in single indicator conditions rated independence judgments higher ($M = 5.35$, $SD = 1.41$) compared to those in multiple indicator conditions ($M = 4.87$, $SD = 1.76$), $t(727) = 17.91$, $p < .001$. As Figure 1 indicates, the serial addition of indicators yielded a decrease in independence judgments. An independent samples t-test revealed that there was a significant difference in independence judgments between one indicator ($M, 5.46$, $SD, 1.38$), and two indicators ($M, 5.02$, $SD, 1.67$), $t(494) p < .001$. No significant difference was observed between adding a third indicator ($M, 4.75$, $SD, 1.86$), $t(505) p = .09$, or a fourth ($M, 4.43$, $SD, 1.81$), $t(253) p = .28$. These results support the hypothesis that multiple idiosyncratic indicators were associated with lower evaluations of review independence.

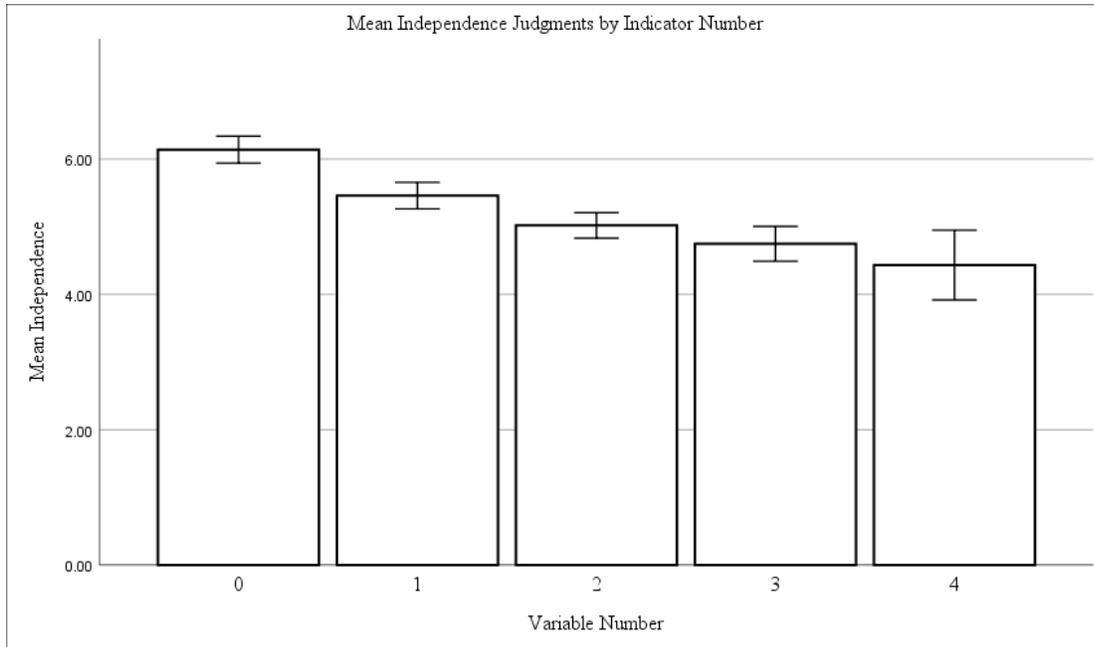


Figure 1. Mean independence judgments by indicator number.

The hypothesis that a high number of reviews should attenuate the effect of idiosyncratic indicators on independence judgments was tested using a Factorial ANOVA with each idiosyncratic indicator and review number as factors. Main effects on independence judgments were observed for attribute co-occurrence, $F(1, 791) = 4.57, p < .05$, partial $\eta^2 = .006$, text duplication, $F(1, 791) = 73.75, p < .001$, partial $\eta^2 = .085$, username similarity, $F(1, 791) = 13.12, p < .001$, partial $\eta^2 = .016$, and review number, $F(1, 791) = 133.82, p < .001$, partial $\eta^2 = .15$. In contrast, no main effect was identified for timestamp similarity, $F(1, 791) = 1.73, p = .189$, partial $\eta^2 = .002$. As Appendix F demonstrates, independence judgments were lower when attribute co-occurrence, text duplication, and username similarity was present compared to when it was absent.

As expected, the results further yielded interaction effects between attribute co-occurrence and review number, $F(2, 791) = 5.55, p < .05$, partial $\eta^2 = .007$, text duplication and

review number, $F(2, 791) = 40.43, p < .001$, partial $\eta^2 = .05$, as well as username similarity and review number, $F(2, 791) = 5.70, p < .05$, partial $\eta^2 = .007$. No interaction effect was observed between timestamp similarity and review number $F(2, 791) = .08, p = .78$, partial $\eta^2 = .000$. Independence judgments scores were lower for with-attribute co-occurrence, text duplication, and username similarity conditions when the number of reviews was low. However, these effects did not persist when the review number was high. This suggests that the number of reviews present attenuated the effects of a. attribute co-occurrence, b. text duplication, and c. username similarity, but not for timestamp similarity which suggests partial support for the hypothesis of indicator and review number interaction.

3.5.2. Message credibility

As idiosyncratic indicators were predicted to have multiple effects on online information evaluation, the effects on other measures were explored. Independence exhibited medium to large correlations to message credibility and purchase intention as indicated in Table 1. We further conducted a Factorial ANOVA with each idiosyncratic indicator and review number as factors and message credibility as a dependent variable to explore potential main and interaction effects. The results yielded main effects on message credibility for text duplication, $F(1, 798) = 24.64, p < .01$, partial $\eta^2 = .030$, and review number, $F(1, 798) = 45.34, p < .001$, partial $\eta^2 = .054$. Message credibility was rated lower for with-text duplication ($M = 6.20, SD = 1.84$) compared to without text duplication ($M = 6.74, SD = 1.38$) whereas it was rated higher for high review number ($M = 6.84, SD = 1.30$) compared to low review number ($M = 6.09, SD = 1.87$). No main effects were identified for attribute co-occurrence, $F(1, 798) = .36, p = .55$, partial $\eta^2 = .000$, timestamp similarity, $F(1, 798) = .00, p = .99$, partial $\eta^2 = .000$, or username similarity, $F(1, 798) = .96, p = .33$, partial $\eta^2 = .001$. An interaction effect between text and review number was

identified, $F(2, 798) = 20.07, p < .001$, partial $\eta^2 = .025$. In contrast, no interaction effects were observed between attribute co-occurrence and review number, $F(2, 798) = 2.38, p = .12$, partial $\eta^2 = .003$, timestamp similarity and review number, $F(2, 798) = .06, p = .80$, partial $\eta^2 = .000$, or username similarity and review number, $F(2, 798) = 3.62, p = .06$, partial $\eta^2 = .005$.

Table 1. Intercorrelations, means, and standard deviations for dependent measures

Measure	1	2	3
1. Independence	-		
2. Message credibility	.69*	-	
3. Purchase intention	.40*	.56*	-
<i>M</i>	5.09	6.46	5.60
<i>SD</i>	1.67	1.65	1.43

Note. * $p < .01$.

3.5.3. Purchase intention

Similarly, a Factorial ANOVA with consistent factors was conducted with purchase intention as the dependent variable. Consistent with message credibility, this yielded main effects of text duplication, $F(1, 788) = 4.63, p < .05$, partial $\eta^2 = .006$, and review number, $F(1, 788) = 6.37, p < .05$, partial $\eta^2 = .008$ on purchase intention. Purchase intention was rated lower for with-text duplication ($M = 5.49, SD = 1.56$) than without-text duplication ($M = 5.71, SD = 1.38$). Conversely, when a high number of reviews were present, purchase intention was higher ($M = 5.73, SD = 1.41$) than when there was a low number of reviews ($M = 5.47, SD = 1.43$). No main effects were identified for attribute co-occurrence, $F(1, 788) = .44, p = .51$, partial $\eta^2 = .001$, timestamp similarity, $F(1, 788) = .16, p = .69$, partial $\eta^2 = .000$, or username similarity, $F(1, 788) = .96, p = .33$, partial $\eta^2 = .001$. Further, no interaction effects were observed between

review number and attribute co-occurrence, $F(2, 788) = 1.15, p = .28$, partial $\eta^2 = .001$, text duplication, $F(2, 788) = .99, p = .32$, partial $\eta^2 = .001$, timestamp similarity, $F(2, 788) = .22, p = .64$, partial $\eta^2 = .000$, or username similarity, $F(2, 788) = 3.68, p = .06$, partial $\eta^2 = .005$.

Whereas idiosyncratic indicators yielded mixed effects in predicting message credibility and purchase intention, a comparable association with choice behavior could not be computed.

Participants were given the option to select one of the two advocated alternatives for each choice task or refrain from formulating a choice altogether. It was found that 95.3% of participants opted to select an alternative in the digital camera task and 94.6% in the laser printer task. As a result, there was insufficient variation in the choice behavior measure to assess its association with independence judgments. These results suggest that partial support for Hypothesis 4 was identified.

3.6. Discussion

Study 2 provides new empirical support for how and under what conditions idiosyncratic indicators impact evaluations that user-generated content is authored by independent entities. Respondents rated independence lower when multiple indicators were present compared to single or no indicators. With the addition of an indicator, independence judgments were lower with a significant difference between serial additions of one and two indicators. While the serial addition of three and four indicators did not yield a significant reduction in independence, the average evaluations still tended in the predicted direction. This suggests that while idiosyncratic indicators may accumulate in prompting suspicion of duplicity, a small number of indicators may be enough to trigger lower independence judgments.

In conjunction with the additive effects of idiosyncratic indicators on independence judgments, it was observed that these effects disappeared when more reviews were present to

obscure the co-occurrence of information. When a high number of reviews were present, the difference between the presence and absence of idiosyncratic indicators along independence evaluations was mitigated. This provides evidence that individuals tend to assume by default that content was authored by independent entities and this assumption is only violated when attention to indicators is facilitated by a smaller corpus of information.

Interaction effects of indicators and review numbers yielded variable strength with a medium-large interaction with text duplication, and small effects on text duplication and username similarity. This suggests that the attenuation on independence judgments may be contingent on the accessibility of independence judgments. Moreover, the strength of this interaction effect may have been accentuated by increasing the volume of reviews available in the review number factor. While 2 and 10 reviews respectively constituted a considerable difference, individuals evaluating UGC may engage with a far higher number thus increasing the potential for idiosyncratic indicators being obscured. Despite clear differences in the capacity for content to constrain the effect, there was a clear attenuation of the effect of numerous idiosyncratic indicators on independence judgments.

A similar main effect and interaction effect was not observed for timestamp similarity. This replicated the finding in Study 1 that the presence of timestamp similarity yielded a negligibly lower difference in independence judgments compared to heterogeneous review samples. Though non-significant, the co-occurrence of timestamp similarity represented a modest reduction in independence evaluations when parity of content post time was observed. One potential explanation for why this effect was not observed across studies has been a consequence of the indicator's utility in evaluating UGC. Specifically, individuals may be less likely to attend to this cue when evaluating online reviews. If the function of UGC is to evaluate

the viability of a product referenced by the reviews, timestamp similarity may be less likely to directly provide information to distinguish product quality. Consequently, people may be less likely to identify when timestamps co-occur across reviews and be less likely to prompt suspicion of duplicity, thus failing to reduce independence evaluations.

In addition to testing the dynamics of idiosyncratic indicators, we further explored the implications that indicators may have on other dimensions associated with selecting alternatives and evaluating information in the context of online consumer choice. It was found that idiosyncratic indicators had varying effects on message credibility. Main effects of text duplication and review number as well as an interaction effect between the two factors were observed with a modest effect size. Neither main effects of the other indicators nor interaction effects on message credibility were observed. However, a marginally non-significant main effect of attribute co-occurrence as well as an interaction effect between username similarity and review number were observed. Independence ratings were found to exhibit a strong correlation with message credibility as well as purchase intention. While there is some potential for idiosyncratic indicators to reduce evaluations of message credibility, this effect was modest and not consistent between all indicators. These results suggest that further exploration is required to understand the precise means by which idiosyncratic indicators impact the evaluations of message credibility.

Consistent with the results of message credibility, main effects of text duplication and review number as well as a marginally non-significant effect of attribute co-occurrence on purchase intention were observed. Further, a marginally non-significant interaction effect of username similarity and review number on purchase intention was identified. No main or interaction effects were observed for the other idiosyncratic indicators. One explanation for the

mixed results of idiosyncratic indicators on message credibility and purchase intention concerns the nature of the cues utilized to inform credibility and purchase judgments. The content of reviews is tenably the primary point of evaluation through which the message credibility and the quality of the referenced is assessed. Thus, co-occurrence along this quality may be more readily accessible to information-seekers and more likely to be integrated in to their judgments. This may explain why the effects of text duplication were observed whereas the co-occurrence of other indicators was marginal or non-significant.

Study 2 was unable to assess the impact of idiosyncratic indicators on choice behavior directly. It was hypothesized that choice formulation should be lower when independence judgments were low. This could not be determined as Study 2 identified that there was insufficient variation in choice behavior with approximately 95% of respondents opting to formulate, rather than defer, making a choice. One explanation for this finding is that choice task in Study 2 did not exhibit conditions which would have made choice deferral likely. To minimize the potential variation in evaluations of the content of reviews, two alternatives were selected per task. This may have rendered the selection of an alternative more likely as there was no large sample of possible alternatives available to promote choice apprehension and thus increase the chance of choice deferral. As the task was designed to facilitate choice with reviews and few alternatives, participants may not have felt compelled to defer making a choice as they may in a context with more alternatives available.

CHAPTER 4. GENERAL DISCUSSION

As user-generated content becomes an increasingly utilized tool for assessing the quality of online choice alternatives, there is also expanding potential for entities online to misrepresent or distort their credentials. One pervasive but ill-defined method through which this can occur is for entities to produce multiple pieces of content while purporting that this content entails separate observations from independent entities. As emerging evidence suggests that the assumption of independence is routinely violated, this invites inquiry in to the information that consumers utilize to distinguish the quality of UGC. We advance that idiosyncratic indicators, cues embedded within message content which suggest a high probability of deception by co-occurring between content entries, can be utilized to detect when content is not authored by independent sources. This provides a descriptive model about a class of cues of which people are cognizant and how they are integrated to formulate judgments about content independence. We provide novel evidence that individuals are sensitive to the independence of information when evaluating online review content and that independence evaluations are seminal for information search and choice selection processes.

Through two experiments, it is demonstrated that idiosyncratic indicators can serve as cues through which people identify that content is duplicitous and the conditions under which they are most likely to assume content is independently authored. The findings suggest that individuals may not immediately be aware that deception is occurring and assume that content is independent by default. This invites a novel line of inquiry related to the cognitive processes which underly source and message credibility judgments in the context of online information search. Based on these findings, we review the central claims advanced by the experimental evidence in the next sections. First, we examine the evidence that idiosyncratic indicators can

function as veracity assessment criteria as well as the underlying mechanisms through which indicators can impact user-generated content. Next, how modelling idiosyncratic indicators can serve as a basis to build theory concerning online credibility judgments is discussed. Finally, the limitations and future directions of this line of inquiry for computer-mediated communication research are explored.

4.1. Idiosyncratic Indicators as Online Detection Strategies

An important consequence of digital technology's emergence as an information tool is the proliferation of enormous quantities of content from as many sources. Amid this landscape of high information volume, a seminal pursuit of digital credibility theorists has been to explore how false or misrepresented content can be identified. Recently, scholars have begun conceptualizing credibility as a product of inferences that are affected by social cues (Metzger et al., 2010). Many scholars have endeavored to catalogue a body of cues available in online platforms which can be utilized to formulate credibility judgments (Sundar, 2008). Though this represents an important contribution, there is a paucity of research which has explored which cues should serve as valid credibility indicators or the conditions under which they should be efficacious at distinguishing information of varied credibility.

To expand understanding of what information search strategies can be utilized to detect duplicity, we sought to introduce a descriptive model illuminating what cues are utilized to infer whether content was independently authored. This entailed establishing a novel criterion for digital credibility that articulates underlying information search processes that prompt independence judgments. It was argued that the co-occurrence of cues between content entries can implicate a higher probability that deception about source independence has occurred. This was based on the premise advanced in prior communication research which suggests that there is

a probabilistic relationship between cues and judgments (Reimer et al, 2012). That is, cues vary in the degree to which they predict a criterion and the extent to which people incorporate them when formulating judgments. If the co-occurrence of review content is a widely utilized cue to determine independence judgments, independence ratings should decrease when content co-occurs.

The goal of Study 1 was to test if people are sensitive to idiosyncratic indicators embedded in isolated review pairs. It was predicted that independence ratings would be lower when review pairs exhibited co-occurrence in their referenced attributes, text content, posting timestamps, and usernames compared to heterogenous review pairs. As expected, the results showed that independence ratings were significantly lower when pairs exhibited similarity along each indicator compared to when pairs were different. Independence ratings were significantly lower when co-occurrence was exhibited between review pairs along attributes, text, and source names compared to mismatched, heterogenous pairs and this yielded a large effect size. Though this provides strong evidence of an effect of idiosyncratic indicators, the difference in independence judgments between timestamp co-occurrence and mismatched pairs was non-significant. While individuals may be aware of idiosyncratic indicators generally, it is possible that people have different estimates for the extent to which cues predict duplicity. These results suggest that if content exhibits co-occurrence in product attributes, text, or usernames, individuals are cognizant of this cue congruence and are consequently more likely to infer duplicity.

While pairwise comparisons of reviews represent an important first step, one limitation is that comparisons occur in isolation and thus the stimuli do not generalize to the environment in which it is typically observed. Study 2 builds on these findings of Study 1 by assessing if people

are sensitive to idiosyncratic indicators in a high-fidelity choice context where they can integrate more product and user-generated content information in to their judgments which they would likely encounter in real-world judgment situations. Co-occurrence was exhibited for very different indicators and the task was designed to refrain from directing attention to reviews by providing product review information. We found that individuals were significantly more likely to rate independence lower when indicators co-occurred compared to mismatched pairs. The results bolster support for the claim that individuals rely on user-generated content when considering online choices and that they are likely to utilize co-occurrence as an independence criterion when evaluating online consumer information.

If idiosyncratic indicators are likely to prompt lower independence judgments due to co-occurring features implicating a higher probability of duplicity, it follows that this effect on independence should compound when multiple indicators are present. Whereas Study 1 established that independence judgments were reduced when co-occurrence of attributes, text, and usernames was exhibited, Study 2 tests if this effect is increases when numerous indicators are present. It was revealed that the presence of multiple idiosyncratic indicators yielded significantly lower independence ratings than single indicators. With the serial addition of each indicator, lower independence judgments were observed with significantly lower judgments between 0, 1, and 2 indicator conditions. This supported our hypothesis that independence judgments should decrease as the number of indicators increases. It stands to reason that individuals integrate multiple co-occurring cues in to their judgments to infer the likelihood of UGC duplicity.

A tenable explanation of this finding that is consistent with our predictions is that indicators prompt the evaluation that there is a higher probability of duplicity through co-

occurrence that is unique to sources. Through demonstrating that only content pairs exhibiting co-occurrence are judged to have lower independence ratings and that adding multiple indicators sequentially lowers ratings, we reason that individuals may utilize multiple indicators to infer a higher probability of duplicity than any single indicator. This supports the central principle that co-occurrence can function as a valid and frequently utilized criterion through increasing the probability of duplicity which is bolstered with the addition of multiple indicators. The findings help to articulate the mechanisms through which the credibility of UGC is evaluated by describing what cues are retrieved and how they are integrated in to credibility judgments.

While supporting the addition hypothesis is a key step in describing the processes through which indicators can function as duplicity detection criteria, we also examined the conditions under which these indicators are less likely to be utilized. If idiosyncratic indicators prompt the judgment that duplicity is more probable, this inference is contingent on the awareness that co-occurring features are present in a sample of content. To reduce this awareness, Study 2 increased the noise-to-co-occurrence ratio by adding a higher number of heterogenous review pairs. It was revealed that there was a significant interaction effect between indicators and higher numbers of reviews on independence apart from timestamp similarity as this main effect was not significant. The effect of lower independence judgments observed from the presence of attribute, text, and username co-occurrence was attenuated. Consequently, this affirms the principle that the efficacy of idiosyncratic indicators as a duplicity detection criterion requires co-occurrence to be accessible for online information-seekers.

Based on the principle of truth default, we explain this finding according to the notion that individuals will not suspect duplicity unless warranted by the availability of idiosyncratic knowledge. Individuals often assume communication is honest as this affords more efficient and

cooperative interaction (Grice, 1989). This truth default is assumed in absence of active deception detection or if evidence for duplicity is not found (Blair et al., 2018). It stands to reason that even with the presence of idiosyncratic indicators, people judged independence higher when more reviews were present due to the review sample being saturated with differing content, thus reducing their capacity to identify co-occurrence in content features. With the reduced attention to co-occurrence due to being augmented with more reviews, participants were more likely to retain the default assumption that content was independently authored. This represents an important contribution to assessing online credibility by providing novel evidence for an important boundary condition of co-occurrence as a duplicity detection criterion and further illuminates the process through which observers of UGC infer its quality as an information assessment tool.

4.2. Independence and Online Choice Processes

The advent of digital media has coincided with innumerable opinions and purchase alternatives generated through a host of sources. Amid this burgeon of available information online, user-generated content has emerged as an important tool for assessing information credibility in exception and addition to traditional sources. The capacity for digital systems to pool the experiences of credentialed individuals in aggregation has emerged as an important source of evidence for evaluating information online (Metzger & Flanagin, 2008). Despite this potential utility, there is ample evidence suggesting that samples of aggregating information are commonly saturated with single entities impersonating multiple individuals (Malbon, 2013; Jindal & Liu, 2008).

To understand the descriptive strategies through which people identify deceptive content, it stands to reason that the important role of independence to evaluating UGC must first be

explicated. This project provides novel support for the principle that the extent to which user-generated content is created by independent sources is a salient dimension for the process of utilizing online consumer information. This has strong implications for the conceptualization of credibility online by illuminating a seminal notion that aggregated content should be valid only insofar as it comprises independent contributions. The insight concerning the salience of independence to online credibility judgments provides novel evidence for the dynamic nature of online credibility which has previously only served as a tacit feature of prior credibility approaches. These conceptions share the assumption that credibility consists in the extent to which content is deemed to be unbiased, objective, and authentic (Sundar, 1999). Consistently, scholars have posited that one important reason for the utility of UGC in information assessment is that it can remove the bias of individuals by aggregating them among a large group of opinions (Flanagin & Metzger, 2013).

Across two studies, we demonstrated that when co-occurrence is exhibited between content providers, people are aware that the assumed unbiased nature of UGC can be violated. This yields key evidence to suggest that an important prerequisite for user-generated content being viewed as credible requires that it is comprised of a sample inhering unique contributions. Whereas this has been treated as a tacit feature in prior research, we yield new support for the notion that the variability in content independence has strong implications for theorizing online credibility evaluations. Through identifying cues which prompt suspicion that content has not been independently authored, this provides a novel criterion through which credible UGC can be distinguished.

In addition to defining the nature of independence and the cues which prompt suspicion that it has been violated, an important objective of Study 2 was to explore possible connections

to existing research on digital credibility assessment. Many theorists have advanced that credibility is a latent feature that is indicated by message effects (Borah, 2013). Recently, Appelman & Sundar (2016) proposed that the credibility can be construed as an individual judgment about the accuracy or veracity of communication content and created a seminal index of online message credibility. Study 2 found that the presence of text duplication reduced message credibility evaluations. Though this effect was modest and did not persist across other indicators, this provides some evidence suggesting that co-occurring indicators may impact evaluations of message credibility. Moreover, independence were strongly correlated with message credibility. This affirms the notion that violations of independence constitute an important factor in assessing if content is unbiased and is thus a worthy source for consideration. To the extent that recommendations inhered within content are taken as evidence in favor of the objects or products they reference, content independence serves as a precondition for the credentials and persuasiveness of UGC.

In conjunction with connecting independence to established credibility measures, our findings further suggest that credibility judgments may be influenced not only by cues embedded in message content but also the online medium through which it is conveyed. More recent approaches to credibility have explored how features of digital technology which support messages can provide cues about the identity of a source or the quality of information contained (Chung, Nam, & Stefanone, 2012; Sundar, 2008). Both studies confirmed that individuals are likely to rate independence lower when co-occurrence of external source information is exhibited whereas timestamp co-occurrence did not significantly impact independence. This suggests a more nuanced characterization of information assessment by describing that low-quality content can be identified through cues retrieved from facets of UGC.

We further sought to identify the implications of idiosyncratic indicators on the online selection process and choice behavior. Specifically, we found that the presence of text duplication yielded lower ratings of purchase intention compared to its absence whereas this effect was not observed for other indicators. Though this evidence does not generalize to all tested indicators, this implicates that suspicion of duplicity may impact the degree to which people rely on user-generated content to inform their choice selection process. By demonstrating what UGC features are important for a source to be credible and thus more persuasive, this provides new evidence for what cues are salient to individuals formulating a choice and which are negligible. It affirms what qualities of online messages prompt inferences that UGC is produced by valid sources and thus why recommendations are worthy of consideration.

CHAPTER 5. FUTURE DIRECTIONS

Taken together, these results suggest that individuals are cognizant of cases when features of user-generated content co-occur. This co-occurrence can function as a useful criterion to identify online deception by implicating a higher probability that different content messages are not comprised of unique entities. It further illustrates the process through which these inferences are formulated by demonstrating that independence judgments are contingent on the accessibility of inculcating cues. Whereas prior online research adopting information-processing approaches has provided typologies concerning what online information may impact online credibility judgments, many fall short in articulating the cognitive processes through which social cues prompt credibility judgments or describe the strategies of information search that people utilize. Thus, we provide novel evidence for the importance of independence to user-generated content as well as describe how and when idiosyncratic indicators can detect if independence has been violated. These studies contribute to understanding about the processes underlying information retrieval and integration for credibility judgments of online content. Despite these findings, further inquiry is warranted concerning the scope of idiosyncratic indicator utility to UGC, the conditions under which judgments of duplicity are activated, and how independence judgments affect the online selection processes.

Further studies may explore how idiosyncratic indicators function in the deceit identification process by determining how the co-occurrence criterion is utilized with other information inferred from UGC. The present studies investigated if co-occurrence along content message features (attributes, text duplication) and source cues (timestamp, username similarity) prompt lower independence judgments. However, it is highly likely that co-occurrence along many other content features may prompt suspicion of duplicity insofar as they implicate a higher

probability of dependency between content entries. Research would benefit from exploring features of communication messages such as parts of speech or sentiment analysis to understand how they influence independence judgments. An important boundary condition governing what features constitute co-occurrence and thus prompt duplicity suspicion consists in the extent to which these features are specific to two or more content entries while being distinguished from other entries within the content sample.

Across both studies, we found that lower independence judgments were observed for indicators embedded in message content as well as source dimensions inferred through platforms. However, one limitation concerns the relatively small effect in independence ratings for certain indicators with a non-significant difference being observed for timestamp similarity across studies. The different magnitude in the effects on independence ratings signaled through these co-occurring features suggest that idiosyncratic indicators differ significantly in the extent to which observers infer that content is duplicitous. If the accessibility of cues is required to prompt attention to co-occurrence, this could explain the lack of an effect of timestamp similarity as it may not be a content attribute to which UGC users frequently attended. Future studies could explore how the relative prominence of certain features impacts the magnitude of the effect on independence judgments.

Whereas the present studies provide strong evidence that varying the stimulus-to-noise ratio may moderate the impact of idiosyncratic indicators on independence judgments, the effect of review number should be interpreted with caution as other factors may have influenced the results. Numerous features of user-generated content have been investigated in the context of consumer choice (see Cheung et al., 2009). For example, the valence of reviews has been shown to influence evaluations of review credibility (Lim & Van der Heide, 2014). We controlled for

review valence as this manipulation was beyond the scope of the present investigation and each review exhibited high ratings. The higher number of reviews may have increased the accepted of the advocated product as it indicated more people considered it favorably, thus impacting purchase intention. To provide further evidence for the assumption of default independence, further studies may pursue alternative manipulations to increase the noise-to-signal ratio while holding review number constant. One approach may entail varying the amount of information provided within content entries or altering features of the online platform to obscure or highlight co-occurrence. Alternatively, future studies may test potential interactions between the number of reviews and the valence of reviews by systematically varying the valence of reviews in conjunction with idiosyncratic indicators. It is reasonable to suggest that co-occurrence may not predict lower independence judgments when they exhibit low valence as there is less incentive to duplicate content and thus observers may infer a lower probability of duplicity.

In addition to further articulating the processes and boundary conditions governing idiosyncratic indicators, researchers would benefit from clarifying how independence judgments relate to selection processes online. One limitation of Study 2 is that not all idiosyncratic indicators affected lower message credibility and purchase intention ratings when present between review pairs. This suggests that idiosyncratic indicators likely vary in the degree to which they impact credibility evaluations and choice directly. This provides some initial support for the notion that idiosyncratic indicators may be germane to online credibility and choice selection processes. However, further investigation in to how idiosyncratic indicators and independence judgments relate to online credibility evaluation is warranted.

Future research would benefit from testing how idiosyncratic indicators and other features of digital technology impact online credibility evaluations and explore how

independence judgments can complement existing online credibility research. Recent credibility research has emphasized the emergent role of cognitively inexpensive, mental shortcuts through which the veracity of online content can be assessed (Metzger & Flanagin, 2013, Sundar, 2008). These approaches suggest that emergent features of digital spaces problematize existing credibility tools and endeavor to describe how individuals utilize information under these conditions. Consistently, idiosyncratic indicators offer a novel information processing strategy by delineating how individuals can adapt judgments under conditions of high information volume and source obscurity. Thus, research may benefit from testing how idiosyncratic indicators may interact with other online information processing strategies. Component in this task is identifying how and the conditions under which individuals may integrate multiple strategies to operate under the constraints of different online information environments. Specifically, if individuals assume content independence by default, the utility of idiosyncratic indicators is contingent of the awareness of indicators which may be embedded in a high volume of information. Fogg (2003) argued that visual design features of a website exert a greater influence on credibility judgments compared to content. Future research may thus test how design features may facilitate or constrain attention toward indicators and thus impact inferences about online credibility.

The role of idiosyncratic indicators in choice selection was further limited by the inability to test how indicators impact the propensity for participants to formulate a choice. Our judgment task restricted the number of sampled alternatives and the review samples which referenced them to two per task. This design was able to provide participants with a choice while minimizing the potential that any review samples would exhibit unintended variation in similarity and thus bias independence ratings. While this design may have increased the internal consistency of Study 2,

it is likely that this choice simultaneously reduced its ecological validity. Prior meta-analyses on choice deferral have found that the number of alternatives available and the complexity of the choice task are important conditions for deciders to avoid formulating a choice (Chernev, Böckenholt, & Goodman, 2014).

As this task was designed to reduce the number of alternatives and provide recommendations, it is reasonable that this facilitated the decisions of participants and prompted nearly all of them to formulate a choice. Future studies would benefit from providing more choice alternatives to afford a more naturalistic condition under which people would be more likely to defer choice. If there is sufficient variation in choice deferral, the extent to which independence explains choice deferral could be investigated. Alternatively, future experiments could independently vary whether the reviews referencing alternatives exhibit co-occurrence to investigate if participants are more likely to select alternatives that are advocated by heterogenous reviews. This would provide direct evidence concerning the persuasive impact of independence judgments on choice behavior and further illuminate the cognitive processes underlying the selection and integration of online choice information.

REFERENCES

- Anderson, E. T., & Simester, D. I. (2014). Reviews without a purchase: Low ratings, loyal customers, and deception. *Journal of Marketing Research*, *51*(3), 249-269.
- Bawden, D., & Robinson, L. (2009). The dark side of information: overload, anxiety and other paradoxes and pathologies. *Journal of Information Science*, *35*(2), 180-191.
- Blair, J. P., Levine, T. R., & Shaw, A. S. (2010). Content in context improves deception detection accuracy. *Human Communication Research*, *36*(3), 423-442.
- Blair, J. P., Reimer, T. O., & Levine, T. R. (2018). The Role of Consistency in Detecting Deception: The Superiority of Correspondence over Coherence. *Communication Studies*, *69*(5) 1-16.
- Bond Jr, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, *10*(3), 214-234.
- Brandt, A. C., Vonk, R., & van Knippenberg, A. (2009). The source effect: Person descriptions by self versus others have differential effects on impression formation. *Personality and Social Psychology Bulletin*, *35*(7), 965-977.
- Brandt, A. C., Vonk, R., & van Knippenberg, A. (2011). Augmentation and discounting in impressions of targets described by third parties with ulterior motives. *Social Cognition*, *29*(2), 210-220.
- Brown, J., Broderick, A. J., & Lee, N. (2007). Word of mouth communication within online communities: Conceptualizing the online social network. *Journal of Interactive Marketing*, *21*(3), 2-20.
- 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.
- Carr, C. T., Hall, R. D., Mason, A. J., & Varney, E. J. (2017). Cueing employability in the gig economy: effects of task-relevant and task-irrelevant information on Fiverr. *Management Communication Quarterly*, *31*(3), 409-428.
- Chernev, A., Böckenholt, U., & Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, *25*(2), 333-358.

- Cheung, M. Y., Luo, C., Sia, C. L., & Chen, H. (2009). Credibility of electronic word-of-mouth: Informational and normative determinants of on-line consumer recommendations. *International Journal of Electronic Commerce*, 13(4), 9-38.
- Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, 43(3), 345-354.
- Chung, C., Nam, Y., & Stefanone, M. (2012). Exploring online news credibility: The relative influence of traditional and technological factors. *Journal of Computer-Mediated Communication*, 17, 171-186
- Cosley, D., Lam, S. K., Albert, I., Konstan, J. A., & Riedl, J. (2003, April). Is seeing believing?: how recommender system interfaces affect users' opinions. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 585-592). ACM.
- DeAndrea, D. C. (2012). Participatory social media and the evaluation of online behavior. *Human Communication Research*, 38(4), 510-528.
- DeAndrea, D. C., & Carpenter, C. J. (2016). Measuring the construct of warranting value and testing warranting theory. *Communication Research*, 0093650216644022.
- DeAndrea, D. C., Tong, S. T., & Lim, Y. S. (2018). What causes more mistrust: profile owners deleting user-generated content or website contributors masking their identities?. *Information, Communication & Society*, 21(8), 1068-1080.
- DeAndrea, D. C., Van Der Heide, B., & Easley, N. (2014). How modifying third-party information affects interpersonal impressions and the evaluation of collaborative online media. *Journal of Communication*, 65(1), 62-78.
- DeAndrea, D. C., Van Der Heide, B., Vendemia, M. A., & Vang, M. H. (2015). How people evaluate online reviews. *Communication Research*.
<http://dx.doi.org/10.1177/0093650215573862>.
- Dellarocas, C. (2006). Strategic manipulation of internet opinion forums: Implications for consumers and firms. *Management Science*, 52(10), 1577-1593.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129(1), 74-118.

- Doh, S. J., & Hwang, J. S. (2009). How consumers evaluate eWOM (electronic word-of-mouth) messages. *CyberPsychology & Behavior*, *12*(2), 193-197.
- Donath, J. (2007). Signals in social supernets. *Journal of Computer-Mediated Communication*, *13*(1), 231-251.
- Ellison, N., Heino, R., & Gibbs, J. (2006). Managing impressions online: Self-presentation processes in the online dating environment. *Journal of Computer-Mediated Communication*, *11*(2), 415-441.
- Flanagin, A. J., & Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society*, *9*(2), 319-342.
- Flanagin, A. J., & Metzger, M. J. (2008). Digital media and youth: Unparalleled opportunity and unprecedented responsibility. In M. J. Metzger and A Flanagin (Eds.), *Digital media, youth, and credibility*, 5-27. Cambridge, MA: MIT Press.
- Flanagin, A. J., & Metzger, M. J. (2013). Trusting expert-versus user-generated ratings online: The role of information volume, valence, and consumer characteristics. *Computers in Human Behavior*, *29*(4), 1626-1634.
- Flanagin, A. J., Metzger, M. J., Pure, R., Markov, A., & Hartsell, E. (2014). Mitigating risk in ecommerce transactions: perceptions of information credibility and the role of user-generated ratings in product quality and purchase intention. *Electronic Commerce Research*, *14*(1), 1-23.
- Fogg, B. J. (2003). Prominence-interpretation theory: Explaining how people assess credibility online. *Proceedings of CHI'03, Extended Abstracts on Human Factors in Computing Systems* (pp. 722-723). New York, NY: ACM.
- Forrest, E., & Cao, Y. (2010). Opinions, recommendations and endorsements: The new regulatory framework for social media. *Journal of Business and Policy Research*, *5*, 88-99.
- Fox, J., & Moreland, J. J. (2015). The dark side of social networking sites: An exploration of the relational and psychological stressors associated with Facebook use and affordances. *Computers in Human Behavior*, *45*, 168-176.

- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review*, *103*(4), 650-669.
- Gigerenzer, G., & Kurz, E. M. (2001). Vicarious functioning reconsidered: A fast and frugal lens model. In K. R. Hammond, & T. R Stewart (eds.). *The essential Brunswik: Beginnings, explications, applications*, 342-347, Oxford, UK: Oxford University Press.
- Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.
- Hancock, J. T., & Guillory, J. (2015). Deception with technology. In S. Sundar's (Ed.) *The handbook of the psychology of communication technology*, 270-289. Malden, MA: Wiley.
- Hartwig, M., & Bond Jr, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin*, *137*(4), 643-659.
- Heydari, A., ali Tavakoli, M., Salim, N., & Heydari, Z. (2015). Detection of review spam: A survey. *Expert Systems with Applications*, *42*(7), 3634-3642.
- Huang, J. H., & Chen, Y. F. (2006). Herding in online product choice. *Psychology & Marketing*, *23*(5), 413-428.
- Huh, Y. E., Vosgerau, J., & Morewedge, C. K. (2014). Social defaults: Observed choices become choice defaults. *Journal of Consumer Research*, *41*(3), 746-760.
- Jindal, N., & Liu, B. (2008). Opinion spam and analysis. In *Proceedings of the 2008 international conference on web search and data mining* (pp. 219-230). ACM.
- Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives?. *Science*, *302*(5649), 1338-1339.
- Kowai-Bell, N., Guadagno, R. E., Little, T., Preiss, N., & Hensley, R. (2011). Rate my expectations: How online evaluations of professors impact students' perceived control. *Computers in Human Behavior*, *27*(5), 1862-1867.
- Lankes, R. D. (2008). Credibility on the internet: shifting from authority to reliability. *Journal of Documentation*, *64*(5), 667-686.
- Lee, B. K., & Lee, W. N. (2004). The effect of information overload on consumer choice quality in an on-line environment. *Psychology & Marketing*, *21*(3), 159-183.

- Lee, E. J., & Shin, S. Y. (2014). When do consumers buy online product reviews? Effects of review quality, product type, and reviewer's photo. *Computers in Human Behavior*, *31*, 356-366.
- Lee, M., & Youn, S. (2009). Electronic word of mouth (eWOM) How eWOM platforms influence consumer product judgement. *International Journal of Advertising*, *28*(3), 473-499.
- Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect." *Communication Monographs*, *66*, 125-144.
- Liu, B. *Web data mining: Exploring hyperlinks, contents, and usage data*: Springer.
- Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science*, *62*(12), 3412-3427.
- Ma, Y., & Li, F. (2012). Detecting review spam: Challenges and opportunities. In *Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, *8th International Conference on* (pp. 651-654). IEEE.
- Malbon, J. (2013). Taking fake online consumer reviews seriously. *Journal of Consumer Policy*, *36*(2), 139-157.
- McCroskey, J. C., & Richmond, V. P. (1996). *Fundamental of human communication*. Prospect Heights, IL: Waveland Press.
- Metzger, M. J., & Flanagin, A. J. (Eds.). (2008). *Digital media, youth, and credibility*. Cambridge, MA: MIT Press.
- Metzger, M. J., & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, *59*, 210-220.
- Neubaum, G., & Krämer, N. C. (2017). Opinion climates in social media: Blending mass and interpersonal communication. *Human Communication Research*, *43*(4), 464-476.
- Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867-872.

- Paolacci, G., Chandler, J., & Ipeirotis, P. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision Making*, 5(5), 411-419.
- Patil, M. S., & Bagade, A. M. (2012). Online review spam detection using language model and feature selection. *International Journal of Computer Applications*, 59(7), 33-36.
- Park, C., & Lee, T. M. (2009). Information direction, website reputation and eWOM effect: A moderating role of product type. *Journal of Business research*, 62(1), 61-67.
- Park, H. S., Levine, T. R., McCornack, S. A., Morrison, K., & Ferrara, M. (2002). How people really detect lies. *Communication Monographs*, 69(2), 144-157.
- Putrevu, S., & Lord, K. R. (1994). Comparative and noncomparative advertising: Attitudinal effects under cognitive and affective involvement conditions. *Journal of Advertising*, 23(2), 77-91.
- Reimer, T., Hertwig, R., & Sipek, S. (2012). Probabilistic persuasion: A Brunswikian theory of argumentation. In R. Hertwig, U. Hoffrage, & the ABC Research Group (Eds.), *Simple heuristics in a social world* (pp. 33-55). New York: Oxford University Press.
- Reimer, T., & Katsikopoulos, K. V. (2004). The use of recognition in group decision-making. *Cognitive Science*, 28(6), 1009-1029.
- Reinhard, M. A., Sporer, S. L., Scharmach, M., & Marksteiner, T. (2011). Listening, not watching: Situational familiarity and the ability to detect deception. *Journal of Personality and Social Psychology*, 101(3), 467-484.
- Sen, S., & Lerman, D. (2007). Why are you telling me this? An examination into negative consumer reviews on the web. *Journal of Interactive Marketing*, 21(4), 76-94.
- Sheehan, K. B. (2018). Crowdsourcing research: data collection with Amazon's Mechanical Turk. *Communication Monographs*, 85(1), 140-156.
- Sprague, R., & Wells, M. E. (2010). Regulating online buzz marketing: Untangling a web of deceit. *American Business Law Journal*, 47(3), 415-454.
- Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger and A. Flanagin (Eds.), *Digital media, youth, and credibility*, (pp. 73-100). Cambridge, MA: MIT Press.

- Sundar, S. S. (1999). Exploring receivers' criteria for perception of print and online news. *Journalism & Mass Communication Quarterly*, 76, 373-386.
- Sundar, S. S., & Nass, C. (2000). Source orientation in human-computer interaction: Programmer, networker, or independent social actor. *Communication Research*, 27(6), 683-703.
- Tavakoli, M., Heydari, A., Ismail, Z., & Salim, N. (2015). A Framework for Review Spam Detection Research. *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 10(1), 67-71.
- Utz, S., Kerkhof, P., & Van Den Bos, J. (2012). Consumers rule: How consumer reviews influence perceived trustworthiness of online stores. *Electronic Commerce Research and Applications*, 11(1), 49-58.
- Walther, J. B., & Jang, J. W. (2012). Communication processes in participatory websites. *Journal of Computer-Mediated Communication*, 18(1), 2-15.
- Walther, J. B., & Parks, M. R. (2002). Cues filtered out, cues filtered in. *Handbook of interpersonal communication*, 3, 529-563.
- Walther, J. B., Van Der Heide, B., Hamel, L. M., & Shulman, H. C. (2009). Self-generated versus other-generated statements and impressions in computer-mediated communication: A test of warranting theory using Facebook. *Communication Research*, 36(2), 229-253.

APPENDIX A. SAMPLE REVIEW PAIR MESSAGE CONDITION

With-Attribute Co-occurrence	John H 7:03pm, November 9 th , 2018 I already knew before I bought this camera that it is a good camera, but that didn't stop me from being surprisingly impressed with it once I got it. This camera seemed to have all of the things that I was looking for in a purchase. It features built in Wi-Fi capability, high-resolution video recording, and a long-lasting battery . I have no complaints.
With-Attribute Co-occurrence	S Richardson1 10:37am, October 4th, 2018 I've been using this camera for quite a while now and it has exceeded my expectations. This bundle has everything you need if you're into photography. I appreciate that it comes with built in Wi-Fi capability, high-resolution video recording, and a long-lasting battery . Overall, the camera is awesome.
With-Text duplication	M.A.R.94 2:19 pm, August 7 th , 2018 I bought this camera for my personal use two months ago and it has been great so far. The camera and everything it comes with is top quality. It has an Optical viewfinder, automatic picture settings, and an easily usable digital screen. I'm really glad that I purchased this.
With-Text duplication	Heidi M 1:10 pm, January 5th, 2018 I bought the camera for my personal use two months ago and it has been great so far. The product and everything it comes with is high quality. It has UV filters, a photo video quality tripod, and is good in low light. I'm really glad that I bought this.
With-Timestamp similarity	Jessica Ortiz 4:17 pm, April 22nd, 2018 For the past few months, I've been keeping tabs on this camera and finally decided to spend the cash. I bought it because I was really impressed with the features in the camera. I like the big 32 GB storage card, the high-speed USB card reader, and the cleaning kit that comes with it. I'm not sure why I didn't just buy it sooner.
With-Timestamp similarity	Matt1009 4:18 pm, April 22nd, 2018 Other than having a smart phone, I didn't have much experience with cameras before buying this. Recently, I've learned what really stands out about it. It includes automatic picture modes, the durable carrying bag, and the slight zoom function. I really learned to appreciate this camera.
With-Source username similarity	T Clark4 6:03 pm, May 27 th , 2018 I got this camera when my older camera stopped working well and this seems to be an improvement. The list of features made it really stand out to other brands. It is affordable, has great picture clarity, and great clarity in the photos. I'd say it is a great buy for personal use.

With-Source username similarity	T Clark2 3:52 pm, April 22 nd , 2018 Over the last year or so, I have been wanting to get a ‘real’ camera and finally got one. All the things the camera comes with made it an easy choice. I particularly like the lens filters, intuitive on-screen controls, and the memory card bag. I’d definitely recommend buying it.
---------------------------------------	---

APPENDIX B. STUDY 1 SCENARIO INFORMATION

Imagine that you are searching for different, hypothetical electronic products to purchase online for yourself. You want to look up more information about the products on a popular product reviewing website. You should consult and evaluate product information and reviews to ensure that you make an informed opinion about what is best for you. In what follows, you will be asked to evaluate electronic product information, review information, and make choices about the products. Next, you will be asked a number of survey questions relating to your perceptions of this information. Please bear in mind there are no right or wrong answers. We are only interested in your personal preferences and opinions.

APPENDIX C. STUDY 2 SCENARIO INFORMATION

Imagine that you are searching for different electronic products to purchase online for yourself. You want to look up more information about the products on a popular product reviewing website. You don't have a lot of knowledge about the specific products you are purchasing so you should consult and evaluate additional information to ensure that you make an informed opinion about what is best for you.

In what follows, you will be asked to review information about electronic products, make choices about them, and respond to a number of survey questions relating to your perceptions about them. Please bear in mind there are no right or wrong answers. We are only interested in your personal preferences and opinions.

APPENDIX D. HYPOTHETICAL PRODUCT PAGE

Electronics > Cameras & Photo



Add to Cart

Buy now

S 41793 EOS 5D 23.1MP Black Digital-SLR Camera

★★★★★ [see customer reviews](#)

Price: \$419.00

- Extended 2-year warranty - DSLR Auto Power Slave Flash 2 - UV Filters - 4 X Optical Zoom - Optical viewfinder - Auto picture review setting - Screen Protectors - High framerate 180Fps
- 58" Camera Video Tripod - Low light sensitive - Durable material - Intelligent auto mode - Kit utility bag - LCD - Editable digital screen - Lens cleaner kit - UV-CPL-FLD Filter kit - Fiberglass construction
- Wifi-connection capability - Full HD 1080p Video Recording - LP-E10 Lithium-Ion Battery - Digital flash controls - Memory Card Hardcase - High Definition 58mm Wide Angle Lens - Splash-resistant case - USB Card Reader - 32 GB SD

APPENDIX E. SAMPLE STUDY 2 FEATURED REVIEWS

Below are some of the featured reviews for the Camera product. When you have finished evaluating them, move on to the next stage to make a choice about the products.



Timothy H

★★★★★ 10:25 pm, October 19th, 2018

I have been browsing around for a new camera for a while and finally settled on this one and it feels like a good choice so far. It has all the important things that I wanted: weather resistant material, a built-in flash lamp, and a 2-year warranty. All in all, it seems like I made a good choice.



Matt 1009

★★★★★ 7:03 pm, November 9th, 2018

I bought this camera for my personal use two months ago and it has been great so far. The camera and everything it comes with is top quality. It has automatic picture modes, the convenient carrying bag, and the slight zoom function. I'm really glad that I purchased this.

APPENDIX F. STUDY 2 INDEPENDENCE DESCRIPTIVES

Dependent Measures		
	Review Low	Review High
Attribute Co-occurrence - With		
<i>M</i>	4.28*	5.67
<i>SD</i>	2.01	.98
Attribute Co-occurrence - Without		
<i>M</i>	4.73*	5.67
<i>SD</i>	1.84	1.14
Text Duplication – With		
<i>M</i>	3.73**	5.57
<i>SD</i>	2.03	1.08
Text Duplication – Without		
<i>M</i>	5.27**	5.79
<i>SD</i>	1.49	1.03
Timestamp Similarity – With		
<i>M</i>	4.42	5.63
<i>SD</i>	1.93	1.04
Timestamp Similarity - Without		
<i>M</i>	4.58	5.73
<i>SD</i>	1.95	1.08
Username Similarity – With		
<i>M</i>	4.18**	5.61
<i>SD</i>	1.97	1.07
Username Similarity - Without		
<i>M</i>	4.81**	5.74
<i>SD</i>	1.86	1.05

Note. Means with * subscript between indicator conditions denotes significant differences at the .05, whereas ** subscripts indicate significance at the .001 level.