

5-2018

Perceived Credibility of Information on Internet Health Forums

Madison Sauls

Clemson University, msauls@clemson.edu

Follow this and additional works at: https://tigerprints.clemson.edu/all_dissertations

Recommended Citation

Sauls, Madison, "Perceived Credibility of Information on Internet Health Forums" (2018). *All Dissertations*. 2110.
https://tigerprints.clemson.edu/all_dissertations/2110

This Dissertation is brought to you for free and open access by the Dissertations at TigerPrints. It has been accepted for inclusion in All Dissertations by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.

PERCEIVED CREDIBILITY OF INFORMATION ON INTERNET HEALTH
FORUMS

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Human Factors Psychology

by
Madison Elizabeth Sauls
May 2018

Accepted by:
Dr. Laura Whitlock, Committee Chair
Dr. Lee Gugerty
Dr. Rich Pak
Dr. Fred Switzer

ABSTRACT

Internet forums (Yahoo! Answers, Reddit, etc.) have become highly utilized resources that provide informational support on diverse topics. Nearly anyone can contribute information to forums, regardless of their expertise on the topic. Thus, forum users are responsible for evaluating the advice they receive. This raises questions of how information credibility is assessed by users, particularly those seeking health information in forums. There are many explicit and implicit cues that may influence how users evaluate information credibility on health forums, such as spelling accuracy and community star ratings. However, many of these cues have only been examined through interview techniques and not studied experimentally. The present study used the Elaboration Likelihood Model (ELM; Petty & Cacioppo, 1986) as a theoretical framework to explain how forum users evaluated health information under different circumstances. A factorial design was used to examine how perceived credibility of forum advice was affected by community star ratings, the presence or absence of spelling errors, the level of participant involvement, and the context of the severity of the health topic. Results indicated that posts with high star ratings were perceived as significantly more credible than posts with low star ratings, and posts without spelling errors were perceived as significantly more credible than posts with spelling errors. However, results did not support participants' credibility evaluations of advice through the ELM framework. The lack of support for this framework may have implications for how much effort forum users take to evaluate advice credibility. However, limitations of the study may have played a role in the findings and thus are discussed.

DEDICATION

This dissertation is dedicated to those who have supported me over the course of my educational journey – those who have loved me, even when they didn’t like me. To my Mom, for answering nearly every phone call over the last ten years, and for every big or small gesture that made it easier for me to focus and achieve my goals – especially over the last three years. To my Dad, for supporting me and trusting me every time I changed directions, for giving me wisdom in times of need, and for your attempt at keeping me level headed. To my Sisters, Jessica, Rachel, and Macy, for being my best friends and for always believing in me and pumping me up when I needed it. Finally, to Spencer, for your love and commitment from afar. Thank you for keeping me grounded, for traveling hours to experience my stress with me, and for always making me laugh, even when I wanted to cry. I love you all very much, and feel honored and blessed that you have shared this journey with me.

ACKNOWLEDGEMENTS

I would like to share my appreciation with those who have advised me through all stages of graduate school. First, I want to give a special thanks to Dr. William Overman for mentoring me in my first two years of graduate school and laying a strong foundation that made me confident in pursuing a doctorate. Next, thank you to Dr. Laura Whitlock for advising me through my dissertation and for making me a better researcher. I'd like to thank Dr. Lee Gugerty, Dr. Rich Pak, and Dr. Fred Switzer for providing thoughtful and challenging feedback during this process, and for helping shape a strong dissertation that I am proud of. Finally, I'd like to thank Dr. Janet Donnelly for providing guidance throughout the many iterations of this document that eventually lead to the final product.

TABLE OF CONTENTS

	Page
TITLE PAGE	i
ABSTRACT	ii
DEDICATION	iii
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	v
LIST OF FIGURES	vi
CHAPTER	
I. INTRODUCTION	1
Internet Health Forums	1
Health Forum Uses and Users.....	3
Credibility Assessments	6
How credibility theories map onto internet health forums	17
Conclusion	26
Present Study	28
II. METHOD	34
Pilot Testing	34
Participants.....	34
Materials	37
Design	42
Procedure	42
Variable Coding and Organizing	44
III. RESULTS	47

Table of Contents (Continued)

	Page
Manipulation Check.....	47
Demographics	47
Perceived Credibility	48
Perceived Intellectual Ability	50
Task Completion Time	51
Information Retention.....	53
Motivation.....	55
Subjective Evaluation of Forum Post	56
Expertise	58
Individual Differences	59
 IV. DISCUSSION.....	 62
Star Ratings.....	62
Spelling Errors	64
Context.....	66
Involvement	68
Subjective Evaluation Criteria	69
Individual Differences	70
Overall.....	71
 V. LIMITATIONS AND FUTURE DIRECTIONS.....	 74
 VI. CONCLUSIONS.....	 79
 VII. APPENDICES	 80
A. Experimental Text from Forum Posts and Responses	81
B. Credibility Assessment Scales	84
C. Intellectual Scale	85
D. Expertise Items.....	86
E. Instructional Manipulation Check.....	87

Table of Contents (Continued)

	Page
F. E Health Literacy Scale (e-HEALS)	88
G. Short Test of Functional Health Literacy (S-TOFHLA).....	89
H. Short Forum 36 Item Health Questionnaire (SF-36)	90
I. Internet Usage Survey	96
J. Retention Check Quiz	98
K. Pilot Testing Results	101
 VIII. REFERENCES	 103

LIST OF TABLES

Table		Page
1.	Overall demographics and demographic comparisons by context	47
2.	Displayed means for main dependent variables by main independent Variables	48
3.	Frequencies of free response data sorted by topic of comment	56
4.	Results of questions regarding experience with health scenario, with frequencies of developed expertise variable	59
5.	Pilot results for importance of question	102
6.	Pilot results for forum responses.....	103

LIST OF FIGURES

Figure	Page
1. Sample thread adapted from www.patient.info/forum	2
2. Exclusion criteria during data collection and cleaning	37
3. Life-threatening forum post followed by a forum response with low star ratings and spelling errors	38
4. Flow of forum post task evaluation	42
5. Task flow for entire study beginning with informed consent and ending with debriefing	44
6. A sample response to the subjective evaluation criteria item, shown with coded criteria	46
7. Differences in perceived credibility for each independent variable.....	49
8. Differenced in perceived intellectual ability between independent variables.....	51
9. Differences in task completion time between independent variables	52
10. Interactions between star ratings and spelling errors for task completion time	53
11. Differences in information retention between independent variables	54
12. Interactions between star ratings and spelling errors for information retention	55
13. Results from Chi-square test of association between context and ‘see a doctor’ criterion	57
14. Results from Chi-square test of association for context and ‘own personal knowledge’ criterion	58

CHAPTER 1

INTRODUCTION

Internet Health Forums

The internet is a platform that offers individuals access to information covering nearly every topic imaginable. While the internet was once only accessible to few people, it is now used daily by 62% of adults in the United States (Zickuhr & Madden, 2012). Users can access information or social interaction from home or anywhere internet access is available. Internet forums, also referred to as discussion or bulletin boards, are a common source of both information and social interaction online. Popular examples of internet forums include Yahoo! Answers, Reddit, and WebMD forums. Forums are composed of conversations that follow a hierarchical structure that originate with an initial post followed by multiple replies, often referred to as a “thread” (Figure 1;C). Most forums have interactive mechanisms that are meant to help readers evaluate information within a post. For example, some forums assign credentials to each contributor based on personal information such as level of education, number of posts or number of replies to unique questions (Figure 1;A). Also, some forums allow active users to rate posts, typically shown through symbols such as star ratings or likes (Figure 1;B).

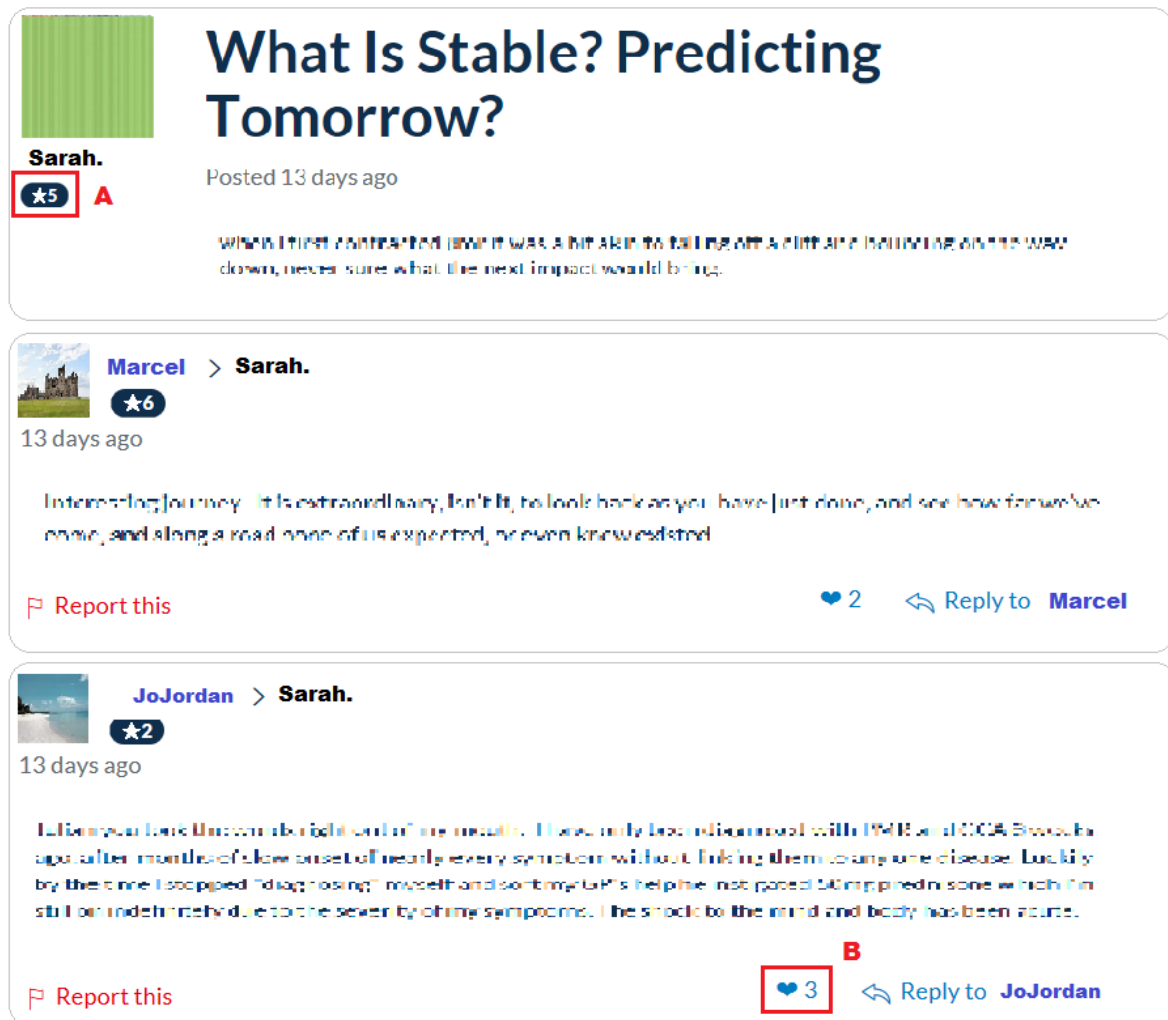


Figure 1. Sample thread adapted from www.patient.info/forum. (A) Credentials assigned to the contributor by the forum, “Level 5: Guru”. (B) Active voting system for the posts, 3 people have indicated it is a good answer. (C) Example of thread: an initial question followed by responses.

Forums allow users to participate in social exchanges with other members through replying to other users’ questions or asking questions of their own. Alternatively, users can take a more passive role and read through questions and replies that other users have posted without posting themselves. In one review of health forum behavior, researchers found fewer than 15% of users who read information on health forums actively contributed information to the forum by writing posts and providing health information (Thackeray, Crookston, & West, 2013). Internet forums stand in contrast to traditional

internet websites that publish content by official authors. Rather, forums are made up of a community of users that drive the flow of information, and users are not typically required to be professionals in the specific topic. In the case of traditional websites, users can view the author and take further steps to verify the author's credentials. In the case of forums, however, the author may not be a known or easily verifiable source.

Health Forum Uses and Users

Health forums are consulted for various reasons by user groups looking for different types of information. Behavior within a health forum depends on both the needs and individual qualities of users. Health forum posting behavior is influenced by factors such as type of illness, type of support needed, and gender of the poster. One study that examined health forum posting behavior looked at and coded posts from two separate WebMD forum groups: a breast cancer forum and a prostate cancer forum (Blank & Adams-Blodnieks, 2007). The most common topics of discussion overall for both forums was the solicitation of either informational or emotional support. Separate analyses of the two forums revealed that the breast cancer group followed the overall trend, with support posts being the most common. However, in the prostate cancer group, medical and treatment posts were the most common. Additionally, while breast cancer posts were more likely to be posted on one's own behalf, prostate cancer posts were almost equally likely to be posted by the individual with cancer or by a loved one. This finding pointed to traditional gender roles existing within health forums, where women were more likely to seek support while men were more likely to seek medical/treatment information, and women were more likely than men to act as a caregiver for others (Blank & Adams-Blodnieks, 2007).

Similar trends involving both demographics and support type have been observed in other internet health forums. In a review of three different health-related forums, Uden-Kraan and colleagues (2008) found that 91% of the posters were middle-aged females. Community members offered different types of informational and emotional support to advice seekers, including personal experience sharing, information provision, and comments of empathy and support. These posts suggest that the type and structure of information shared on forums varies based on individual users. Still, regardless of how users contribute information, it is possible that the desired format of information received in forums is dependent on the recipient, making it difficult to assess how users as a whole evaluate each type of information.

Demographic differences may exist between those who consult traditional internet websites for health information versus those who utilize internet health forums (Magnezi et al., 2015). Responses to a survey measuring health forum use showed that those who consulted traditional websites were more likely to be male, while those who consulted health forums for information were more likely to be female. Also, health forum users were typically older than individuals who only used traditional websites. The study found no significant differences in race, ethnicity, household income, education level, or living environment between health forum users, suggesting a broad appeal of such sites. Individuals with lower income reported using the internet to acquire informational as well as emotional support. Therefore, health forums may be an appropriate fit for those of lower income, as health forums provide both types of support in one location.

In contrast to other findings, some evidence suggests that lower income individuals may use social health sites more frequently than those with higher incomes

(Atkinson, Saperstein, & Pleis, 2009). One potential interpretation of this finding is that those with lower incomes may be replacing formal medical care with advice from online health forums. There is evidence that for some populations, low-income individuals may have low health literacy levels (Schillinger, Barton, Karter, Wang, & Adler, 2006). Likewise, adults covered by Medicare, Medicaid, or uninsured adults are more likely to have basic or below basic health literacy skills (Vernon, Trujillo, Rosenbaum, & DeBuono, 2007). This lack of health literacy skills could potentially hinder their ability to fully comprehend advice given online, which would be especially problematic when using the internet as a primary source of health information. Since past research has shown that lower income individuals are more likely to use the internet for social support (Magnezi et al., 2015), health forums have the potential to serve their needs, but only if the information is presented in a way that can be understood and retained.

Previous connections have been made between internet use and individuals with low health literacy. One ethnographic study of low literate adults revealed certain user characteristics regarding internet use and computers (Zarcadoolas, Blanco, Boyer, & Pleasant, 2002). When researchers asked participants what they would use the internet for provided they had access, health information was the prevailing response. The researchers noted users' online information search was often hindered by literacy-related barriers, such as misspelling search terms and misunderstanding information categories. Therefore, an effort must be taken to make online health information and content both accessible and easy to understand. Unfortunately, there are few academic studies that have tried to bridge the gap between low health literacy and health information online (Zarcadoolas et al., 2002). If low-income individuals are searching for both informational

and social support, internet health forums may be a valuable tool, and research should incorporate ways to reach these populations and appropriately serve their needs, thus bridging the gap in that domain.

Credibility Assessments

While internet forums can offer benefits to users looking for informational and emotional support, there are potential drawbacks as well. In most health forums there are no restrictions on who is allowed to contribute information, which raises questions of information credibility (Metzger, 2007). An individual's assessment of online health information is influenced by the perceived credibility of the speaker and his/her information. Credibility, often defined as the extent to which information is believable, is a perceived quality that results from many different dimensions of information (Fogg & Tseng, 1999). The perception of credibility is further complicated by differences that are perceived differently depending on the individual. In the field of information technology, credibility matters when computer products are used as knowledge sources, to instruct users, and/or act as decision aids, among other reasons (Fogg & Tseng, 1999). Therefore, internet health forum users should be concerned with information credibility as they look to those sources for advice in decision-making moments. Theories of perceived credibility and information evaluation were developed far sooner than the development of information technologies, and can serve as the groundwork for understanding the specific case of credibility online.

Source credibility typically involves a communicator's argument, the level of persuasion that may result from the argument, and the partial or total acceptance of the argument by a listener. While there may be many characteristics that work together to

affect persuasion, researchers have described source credibility as the combined effect of a speaker's perceived expertise and trustworthiness (Hovland, Janis, & Kelley, 1953). To gain a full understanding of credibility, the precursors of expertise and trustworthiness must also be understood. Overall, an individual is perceived as an expert when they seem to be knowledgeable, experienced, competent, and intelligent (Fogg & Tseng, 1999). Perceived trustworthiness between individuals is based on attributes of perceived expertise, reliability, intentions, activeness, personal attractiveness, and the majority opinion of other listeners towards the speaker (Giffin, 1967). It is evident by these definitions that there are many attributes that are combined to result in the credibility judgment. Also, because expertise is one of the attributes that comprises the construct of trustworthiness, it is clear that expertise and trustworthiness are not easily disentangled. Central to both expertise and trustworthiness, however, is the fact that each attribute is usually perceived, not objectively obvious or stated, and the resulting credibility rating is based on the information receiver's own perception (Berlo, Lemert, & Mertz, 1969). This point of perception is important in online forum environments for two reasons. First, objective markers of credibility are often not readily available in internet forums. For example, some attributes such as intention and reliability may not be readily known by the receiver, and will be perceived and subjectively interpreted. Second, each receiver's perception is, by nature, based on personal interpretation, suggesting that credibility ratings of information on internet forums may be influenced by individual differences.

Elaboration Likelihood Model and Credibility. As previously stated, theories of credibility tend to focus on how a communicator's argument may result in persuasion of an audience to partially or fully accept information as stated or a lack of persuasion

that leads to the rejection of information. Therefore, attributes of perceived credibility and theories of persuasion may work hand in hand to explain how individuals evaluate and act on information. One theory of persuasion, the Elaboration Likelihood Model (ELM; Petty & Cacioppo, 1986), focuses on how multiple attributes attached to a source of information shape or impact how the message is evaluated. Examples of these attributes are source expertise, number of arguments for a point, and the mood of the receiver. The ELM is based on a continuum of elaboration, whereby the degree of motivation to elaborate, or closely and carefully attend to and evaluate the information, varies. At the low end of the elaboration continuum, individuals are less engaged, and their attitudes toward information are formed or changed via processes that take less cognitive effort, which results in quicker judgment. This is classified as the peripheral route of evaluation. At the high end of the elaboration continuum, attitudes are based on the results of extensive and effortful information-processing. This is classified as the central route of evaluation. The motivation of an individual to take either the peripheral route or the central route acts on a continuum. That is, the peripheral and central route should not be thought of as dichotomous, but rather as a range of cognitive effort and information processing that depends on ever changing circumstances.

One variable that may influence whether an individual takes a more central or peripheral route of elaboration is awareness of one's knowledge and capabilities within the domain (Petty & Wegener, 1999). For example, those who know nothing about an issue or do not believe in their own analytic capabilities may rely on others' knowledge to make their decision. In the case of health information, this could be detrimental if the user does not utilize all of the cues within the advice to calibrate an evaluation. For

example, some internet health forums allow users to publicly rate information, which enables posts to be given an overall rating. If an individual lacks analytic capability or takes the peripheral route for other reasons, they may see the community rating and decide to take the attached advice as appropriate, even if other forum members have contributed information that may match better to what the user needs. In this case, the individual is discounting other contextual clues within the information that could potentially lead to reading the advice more carefully, reading more sources, calibrating an appropriate level of trust towards the overall advice, and thus more carefully shaping the credibility assessment.

One of the biggest predictors of motivation and the subsequent route of elaboration is how important the issue is to the person, which is often affected by how invested an individual is in finding the right information. For important issues, individuals typically have a greater desire to be correct in their evaluation of information that may affect them if they act on it. Therefore, they are more likely to take the central route and carefully scrutinize the information before making a judgment (Petty & Wegener, 1999). For less important issues, individuals are more likely to take the peripheral route and review the information through low-effort processes. This personal investment is sometimes referred to as the level of involvement in the decision situation, and in some cases the level of involvement tends to affect the processing strategy used during information verification.

Researchers have used the ELM to understand online consumer behavior by manipulating the level of involvement a consumer has with the product and then identifying whether the central or peripheral route is taken. In one study, participants

viewed a web camera shopping site and made judgments about specific products (Yang, Hung, Sung, & Farn, 2006). Researchers manipulated the level of involvement of the participants. First, they measured the level of experience each individual had with web cameras. Participants who had little experience with web cameras were told that they had an opportunity to win an unrelated item at the end of the experiment, and were marked as low involvement participants. Participants with a lot of web camera experience were told they had the opportunity to win a web camera at the end of the experiment, and were marked as high involvement participants. The results showed that individuals with low involvement relied more on peripheral cues, such as company seals, to formulate their level of trust towards the product. However, individuals with high involvement took the central route and relied on product information quality to formulate their level of trust towards the product.

In another study applying the ELM to online consumerism, participants with varying levels of involvement read customer reviews of a new MP3 player and evaluated the review writers, the products, and any issues associated with the product (Lee, Park, & Han, 2008). Product reviews were experimentally manipulated such that some reviews were low quality and some were high quality. In general, participants were in higher agreement with reviews that were written at a high quality. However, the impact of quality on evaluations was significantly higher for the high involvement individuals than the low involvement individuals, suggesting that high involvement individuals read the content of the reviews more thoroughly. On the other hand, the results showed that individuals in the low involvement group were more persuaded by the titles of the reviews, rather than the actual content. While this study was not conducted within the

health domain, it is similar to health forums in that reviews were written by other members in the same community. As in health forums, the author was not necessarily an expert in the field, and the user had to assess the review based on cues within and surrounding the information.

Although the ELM has been applied within the context of online health information, research is somewhat limited. One study tested how the ELM fit credibility assessments of online information in traditional health websites and found that individual differences played a role in which route users took to evaluate the information (Freeman & Spyridakis, 2004). Those who were interested in the topic or had familiarity with the topic had a higher motivation to evaluate the information more thoroughly. Additionally, those who found the information easy or very easy to read were more likely to take the central route of processing. Results also revealed a positive correlation between the age of the reader and the perceived credibility rating, and a negative correlation between perceived article difficulty and perceived author expertise. Ultimately, these results illustrate the complex relationship between elaboration and credibility ratings. For example, simply written posts may result in participants taking time and effort to read through the information, yet these particular passages are perceived as less expert due to being viewed as simple (Freeman & Spyridakis, 2004). Taken together with the studies of online consumerism, these results suggest that the ELM is able to account for many of the complex judgments internet users make when evaluating information online in multiple domains.

Importance and involvement have both been studied in multiple types of information exchanges to better understand which route individuals take in evaluating

information. Chaiken (1980) conducted a study evaluating written content, during which participants read and evaluated persuasive messages. Some participants were told they would discuss the same topic in a future session (high involvement condition), while others were told they would discuss a different topic in a future session (low involvement condition). Those in the low involvement condition tended to utilize heuristic strategies to quickly process the message content, while those in the high involvement condition tended to engage in more in-depth processing. The results of this study are important to consider when examining information evaluation by forum users who may have different levels of involvement. Many individuals view forums solely as a means to find free information, treating it like a common good (Soroka & Rafaeli, 2006). Since these individuals are not utilizing forums for social support or relationship building, they may only engage in short-term use of the forum until their question is satisfied. Theories of trust are helpful in explaining how short-term and long-term forum use may result in different types of credibility assessments.

The concept of swift trust versus slow trust (Meyerson, Weick, & Kramer, 1995) can be used to explain how humans develop social rules online (Corritore, Kracher, & Wiedenbeck, 2003). The classical concept of slow trust suggests an underlying trust that develops over a long-term working relationship. For example, an individual who works with the same functional team at a job over months or years will slowly develop an underlying trust of the people and mechanisms that drive the team. In contrast, swift trust is specific to a particular action and is swiftly created during temporary involvement and quickly ends. For example, an individual who works with a cross-functional team to accomplish a finite task over multiple days develops a swift trust that is sufficient to work

with the team to accomplish the goal. In online internet forum situations, the same types of trust may be developed based on the individual's goal. This goal may be a finite, time-based goal where a solution is needed by a certain deadline. In contrast, the goal may be less clear, and may involve the need or desire to build a relationship or knowledge base over an infinite amount of time.

For internet health forums, the nature of the illnesses that individuals want to address may map onto the goals associated with swift and slow trust. For example, users with an acute and non-life-threatening medical question typically have an overall goal to find a solution with a finite timeline. For this type of question, the user may seek quick information that will not have lasting or repeating consequences or applications--in essence, a quick fix. In this case, the level of trust the user will develop is very specific to the action, as it will serve the purpose of evaluating credibility, and will swiftly end after the user leaves the site. However, users with chronic and life-threatening medical issues may need more in-depth information that could be applied for a longer period of time. In this case, a long-term working relationship may be slowly cultivated, which may generate some level of underlying trust. Understanding the different use scenarios for health forums and how they map onto theories of trust and credibility is helpful in identifying cases in which individuals are more prone to rely on a certain cue. These theories also suggest that assessments are made on a continuum, and change based on the individual's goals and involvement at the time of the evaluation.

Heuristics to Evaluate Credibility. Other theories may be used separately or in conjunction with the ELM to explain what makes a listener perceive the speaker to be an expert, what makes a listener perceive the speaker to be trustworthy, and what explicit

and implicit cues drive these perceptions. Simon's (1955) theory of bounded rationality suggests that people are not always able to act perfectly rationally due to human cognitive limitations. That is, humans seek to find an optimal combination between cognitive effort and efficiency, and there are often constraints on one's ability to process information (Metzger et al., 2010). Therefore, people are likely to rely on heuristic processing to assess information credibility. Heuristic processing involves the selective use of information to solve decision problems simply (Payne & Bettman, 2004). There are many reasons that individuals engage in heuristic processing in decision-making. One reason is that the individual may have a limited cognitive capacity based on intrinsic cognitive properties or environmental properties, such as time constraints or distractions. Heuristics may also be used in cases in which limited capacities in cognition, effort, or time need to be reserved. Lastly, individuals may simply choose heuristics because the use of such has resulted in satisfactory results in the past (Payne & Bettman, 2004).

Payne and Bettman (2004) proposed that humans have a decision strategy toolbox that is pulled from in order to make the best decision under working circumstances. As stated previously, humans switch between the more effortful decision-making strategies and appropriate heuristics as circumstances change. One example of strategy switching is the expectancy violation heuristic. The expectancy violation heuristic suggests that when information fails to meet the user's expectation, arousal and distraction occurs, causing individuals to systematically review the discrepancy and evaluate what that means for the credibility of the source (Burgoon, Stern, & Dillman, 1995). This heuristic was observed in focus groups during website assessments (Metzger & Flanagin, 2013). Some participants expressed that issues such as poor grammar or typographical errors

contradicted the expectancy that someone put time and money into a website. This evidence suggests that heuristics, typically associated with the peripheral route, may result in the individual taking a more central route of evaluation. Likewise, this switch occurs due to individual expectations of how information should be delivered, further supporting the role of individual differences in credibility evaluations.

Another heuristic found to play a role in traditional website evaluation is the anchoring heuristic. In a traditional website, cues embedded in the structure and layout are more likely to be salient, and as such drive the initial judgment of the website (Sundar, 2008). That is, the first piece of information acts as an anchor, setting the tone for how the rest of the information is received, which eventually informs the credibility evaluation and thus decision-making. Even if later information has the potential to shape the initial judgment, there are circumstances where the original judgment persists through the decision task. For example, according to the systematic-heuristic hypothesis, in an ambiguous situation, the first cue will likely bias processing even when it is systematic and thorough (Chaiken & Maheswaran, 1994).

Based on the present understanding of credibility judgments, it is clear that the process of engaging in information evaluation is nuanced by individual differences. While individual differences have not been studied within the internet forum domain, trends in traditional online information evaluation have been identified. One article highlighted individual attributes that influence susceptibility to phishing attempts online. For example, the technique of imparting a sense of urgency to make a judgment may capitalize on attentional resources that vary by person (Williams, Beardmore, & Joinson, 2017).

Other factors are less straightforward, such as self-awareness. Individuals who are more self-aware tend to consider their own knowledge and how they normally behave when making decisions, leading them to resist acts of persuasion (Williams et al., 2017). In addition to self-awareness, individual trust plays a role in credibility judgments. For instance, due to the large amount of information humans process each day, the disposition to trust information online is generally biased by how truthful previous encounters were (Elaad, 2003). This is likely because scrutinizing each piece of information in contrast to the baseline level of trust would be more cognitively demanding. Overall, when it comes to evaluating potentially false information online, individual differences in attributes such as self-control, self-deception, expertise, and motivation may be present (see Williams et al., 2017 for full review). Therefore, while individual differences have not been clearly identified in evaluating internet forum advice, it stands to reason that they may play a role in decision-making due to the similarity of the two domains.

Research on information evaluation has shown that information judgments are complex and can be affected by many factors, such as involvement in the issue, motivation, perceived importance of the issue, and knowledge within the domain. Some studies suggest that the ELM maps on well to how consumers evaluate information online in regards to products and purchasing (e.g. Yang et al., 2006; Lee et al., 2008). However, it is not clear how those factors affect information judgment within the internet health forum paradigm. The relationship between certain cues within information postings and resulting credibility ratings seems to have more than one layer. First, detected cues influence whether the individual will take a more central or peripheral route

in evaluation. Then, once that path is implicitly or unconsciously selected, those same cues play a further route in how credibility is judged.

In the case of internet health forums, this process is likely to be complicated by the fact that attributes such as expertise are usually not explicitly stated or known, and so this information must be inferred by the reader. Therefore, it is important to understand what aspects of online information relate to certain cues of information quality and credibility. To answer that question, it is necessary to review what cues are identified in internet forums, how they map onto information attributes (e.g., author expertise), and how they implicitly or explicitly influence credibility ratings.

How credibility theories map onto internet health forums

While the evaluation of health information in internet forums is not currently mapped on to specific theories of persuasion, researchers have identified several implicit and explicit cues that may inform behavior when reviewing online information and determining its respective credibility. Researchers in one study reviewed different types of online information and identified five main attributes of information that should be used to verify credibility: accuracy, authority, objectivity, currency, and coverage (Metzger, 2007). That is, users should verify the identity and expertise of the author, whether the site or author has conflicts of interest, if the information is current, and if the information appropriately reflects the most crucial aspects of the topic. Unfortunately, one study showed that internet users were more likely to verify information that was easy to find, such as if the information reflected current knowledge on the topic, and less likely to verify information that was more difficult to find, such as author qualifications or credentials (Flanagin & Metzger, 2000).

Arguably the most dangerous issue with seeking information on internet health forums, especially with the lack of verification skills, is the opportunity for error by contributors. Owing to the open nature of internet health forums and the ability for anyone to contribute regardless of their credentials or expertise in the area, the information posted by users can be unconventional or inaccurate (Culver, Gerr, & Frumkin, 1997). In a standardized review of 1,658 forum messages regarding hand and arm conditions, 79% of the posts offering medical information were found to be authored by individuals without any formal medical training. Furthermore, roughly 36% of those posts written by non-professionals were labeled as unconventional by real medical professionals. Additionally, 61% of the advice was based on personal experience. While advice based on personal experience is not necessarily harmful advice, prior research has found that users tend to trust experiential evidence online (Golbeck & Fleischmann, 2010). If users give inappropriate weight to experiential evidence on a health forum, they may overestimate its credibility.

The ways in which information is portrayed in internet forums is not standardized, and formatting, displays, and the inclusion of certain details depends on the individual forum. Furthermore, the way in which the displayed information is used is influenced by individual and forum specific factors or cues (Brady et al., 2016). Individual cues are based on the writing of the contributor, such as word choices or writing styles that elicit higher perceived trust and credibility. In contrast, forum specific cues are those that the forum itself provides for the users to interact with, such as the ability to rate a post publicly or vote on which post best answers a question within a thread. Whether an individual perceives health forum information to be trustworthy or credible may depend

on the presence of these individual or forum specific cues, though it is still not clear which cues play the most important role in this evaluation.

Individual Cues. Individual cues are those that are inherent within the poster, and as such vary across contributors. Examples include the individuals writing style, background, and experience. Some evidence suggests that internet users tend to trust information that comes from contributors with similar perspectives and backgrounds. In one health forum dedicated to chronic conditions, individuals sought out other users that shared the same opinions about the disease and then built online relationships with them to perpetuate knowledge sharing (Brady, Segar, & Sanders, 2016). Similarly, the background of those individuals became a driving factor as to whether users would find those contributors credible. The same effect was found when examining differences in perceived credibility and homophily, or the extent to which the reader of the content shares personal characteristics with the writer of the content. (Wang, Walther, Pingree, & Hawkins, 2008). In this study, participants were exposed to mock-ups of both a government-sponsored cancer website and a cancer health forum, and were asked to respond to items of credibility, homophily, and overall utility of the information. For the traditional website, the main driving force of evaluation was the perceived credibility of the information, where higher ratings of credibility suggested a greater likelihood of using the advice or recommending it to a friend. For the health forum, homophily was the main driving force for how individuals evaluated the health information, suggesting they would heed the advice or recommend it to a friend because someone similar to them suggested it.

Other individual cues may be influential in credibility assessments. For example, one study found that experiential evidence affected information perception. In a mocked-up internet forum, posts were manipulated to include either general advice (“a golden retriever...”) or advice based on personal experience (“*my* golden retriever...”)(Golbeck & Fleischmann, 2010). Perceived trust in posts was higher when personal experience was implied within the post. Similarly, forum contributors were assigned profile images, either of the individual alone or of the individual with a golden retriever. The presence of the golden retriever in the photo, which signified personal experience with the topic, yielded higher levels of trust towards the advice. Another manipulated factor, length of response, was also positively related to higher levels of trust (Golbeck & Fleischmann, 2010). This relationship has been found in other internet forum work as well (Harper, Raban, Rafaeli, & Konstan, 2008).

Other research has examined exchanges in real-world forums to draw conclusions about how users evaluate information on them. In a review of 4,739 posts to an existing forum, many publicly written judgments of information quality and credibility were found in response to other users’ advice (Savolainen, 2011). Specifically, users made direct comments to contributors about the quality of their information (i.e., “You cannot provide any facts to strengthen your claims because you know nothing about this issue.”). For a post to be judged as high quality, the post needed to be perceived as useful, correct and specific. For a post to be judged as credible, the author of the post needed to be perceived as honest, an expert in the field, and to have a good reputation in the online community. Again, these results suggest perceived qualities, which are difficult to measure and dissociate from other cues.

In a separate study, internet health forum users were interviewed about their perception of health information provided by other users (Lederman, Fan, Smith, & Chang, 2014). Interviews revealed multiple criteria that forum users followed when evaluating advice. For example, users took into account the perceived argument quality of the post and stated that they found it more valuable if the argument seemed like it made sense, was rational, and followed logical reasoning. The degree to which the advice was verifiable also influenced how information was evaluated. For example, a post was verifiable if it referenced scientific links or other evidence backing up its claims. Also, the post was verifiable if other contributors agreed with the advice or backed it up with their own experience, demonstrating the use of the endorsement heuristic. In this particular forum, some comments made by contributors suggested individual differences in how individual posts were evaluated. For some members, information was evaluated positively only if the contributor was literate, typically demonstrated through a competently written post. However, other members explicitly mentioned that the presence of grammatical errors did not take away from the overall value of the information of a post, especially if it brought light to a health situation that the user needed information on.

Grammar and spelling has been identified as a factor affecting perceived information credibility in other domains. In a series of interviews, Wikipedia users revealed that among other cues, poor grammar and spelling made the information less trustworthy, while good grammar and spelling made the writer seem knowledgeable (Rowley & Johnson, 2013). Poor grammar and spelling may affect perceived credibility because it makes the writer look like an amateur rather than a professional, indirectly

addressing the expertise requirement for a high credibility judgment (Metzger & Flanagin, 2010). Participants generally agreed that information should be presented clearly and professionally as a reflection of their expertise and attention to detail. When writers have not conformed to this expectation, the information provided is often discounted almost automatically as non-credible.

Still, the exact effect that poor grammar and spelling may have on one's perception of an author is not completely understood. Kreiner, Schnakenberg, Green, Costello, and McClin (2002) found that both phonological and typographical spelling errors negatively affected the perception of authors' writing ability. Here, participants read sample readings from supposed published authors, and then rated the authors on their intelligence and their writing ability. Evidence suggested that phonological spelling errors had a negative effect on the perceived intellectual ability of the author, though this was dependent on the amount of errors in the passage. In a separate study, typographical errors were inserted into emails to measure readers' perceptions of the authors under the absence of explicit situational information (Vignovic & Thompson, 2010). Typographical errors of both spelling (homophones and incorrect contractions) and grammar (singular/plural mix-ups, comma splicing, and incorrect word endings) were placed within emails written from a potential work collaborator, and participants rated the email sender on perceived conscientiousness, intelligence, and trustworthiness. Email authors received lower ratings on all three attributes when such errors were included in the writing. Thus, when other explicit information is absent, judgments may be made based on factors such as grammar and spelling. This finding may have implications for internet

health forums where attributes such as expertise and authorship are not always clearly defined.

Forum Specific Cues. Forum specific cues are cues resulting from mechanisms implemented by the forum itself to portray information to the user, such as community star ratings. In qualitative studies of forum post evaluations, individuals value the community as a whole and how it collectively perceives certain information. In one study, researchers conducted interviews in order to assess the importance of ratings in information evaluation (Lederman, Fan, Smith, & Chang, 2014). In one interview, a participant said that if a post received good ratings from the community (e.g., 60 “likes”), then the information seemed more reliable. This example of the endorsement heuristic, referred to in forum literature as the bandwagon effect, calls into question the importance of forum specific factors that influence perceived trust and credibility. Few studies have attempted to understand how the bandwagon effect influences the perception of online information, and the results of these studies have produced conflicting evidence. In most internet forums, a question is posed and is in turn answered by multiple users. In some forums community members can rate each answer using established forum mechanisms (e.g., star ratings). In forums utilizing the star rating method, a member can read an answer and rate it anywhere from 1 to 5 stars, depending on how credible he/she thinks the advice is. These star ratings are averaged and displayed above the post, often revealing which post is thought to be the best by the community. This method of allowing the community to publicly rate advice may activate the endorsement heuristic, whereby individuals are inclined to believe information that others also believe. In fact, participants in focus groups revealed that they sometimes placed more emphasis on the

opinions of other community members than on their own firsthand knowledge (Metzger & Flanagin, 2010). Therefore, it is important to understand the degree to which bandwagon cues are reliable signs of credibility on forums. Then, credibility evaluation behavior can be studied under the context of bandwagon cues.

In order to address the degree to which bandwagon cues are reliable indicators of credibility, it is necessary to determine if community ratings are accurate in identifying the accuracy of information available. Some studies show that using the endorsement heuristic is potentially problematic. For example, in one study of an autism health forum (Ben-Sasson, Pelleg, & Yom-Tov, 2016), researchers found that the community tended to rate forum posts differently than medical professionals. Clinicians reviewed each question and answer and selected the best answer, only agreeing with the community vote for best answer on approximately 28.6% of the questions. Parents of children with autism were also asked to rate each reply, and of all the answers voted “best” by the parents, only 31% were rated as “best” by clinicians (Ben-Sasson, Pelleg, & Yom-Tov, 2016). These results suggest that if individuals are making judgments based on the endorsement heuristic, they may be relying on suboptimal advice. However, other reviews of health forum posts have found evidence that most posts containing medical information are seen as conventional by professionals. In fact, in a separate but similar study, none of the medical information recommended in the forum was seen as dangerous to others (Uden-Kraan et al., 2008). These mixed results suggest that the accuracy of information may be forum specific, and that users should not rely solely on community ratings to make a credibility assessment.

Because the accuracy of information provided on health forums seems to vary, it is important to review how community ratings influence evaluation behavior. One study asked participants to read real advice that had been posted on a programming forum but manipulated whether a community rating was visible alongside the advice or not (Matthews, 2015). Some posts included the advice accompanied by a community rating as shown by up and down votes, while other posts only included the advice. Overall, there was a correlation between the participants' ratings of the advice and the community ratings of the advice, even when the participants did not see the community ratings. However, that correlation was stronger when the community ratings were present, suggesting that perceived credibility was affected by what others in the community thought. This finding strengthens the argument that bandwagon effects play a role on information perception.

Kim and Sundar (2011) examined the relationship between bandwagon cues and author expertise ratings in internet forums. Mocked-up forum threads were manipulated to have varying degrees of community endorsement as shown through multiple bandwagon cues, as well as varying levels of authority shown through expertise cues. Forum contributors were given four bandwagon cues (number of views by other members, number of replies from other members, number of shares, and a star rating of the thread's helpfulness) and two authority cues (membership status and number of posts made by the contributor). Each cue was manipulated to represent either a high or low bandwagon effect and either a high or low authority level. The results suggested that bandwagon cues only had an effect on attitudes towards message content while in the presence of high authority cues. However, this finding was reported as only marginally

significant ($p = .07$). Overall, it is unclear how much weight community ratings are given by those searching for the best advice. Understanding evaluation behavior within the context of community ratings and bandwagon cues is an important issue, specifically when the advice concerns important health issues. Therefore, this mechanism should be studied to understand the implications of community ratings, especially given the evidence that not all community ratings are aligned with professional opinion (Ben-Sasson, Pelleg, & Yom-Tov, 2016).

Conclusion

Internet health forums are highly utilized sources of information that provide social support and community involvement without the barriers associated with in-person resources. However, the credibility of information on forums is not always outwardly apparent. Given the amount of information being exchanged through health forums and its potential benefit or detriment to the user, some internet health forums have instituted policies to help users gather the most helpful and least harmful advice, while still being able to benefit from the social support that many value. Thus, it is important to understand how cues within health forums influence users' evaluation of the advice. Classic credibility research has shown that information is evaluated differently based on attributes specific to the individual. While information is sometimes reviewed thoroughly and critically, other times it is reviewed quickly with the use of cognitive heuristics. Some of these processes have been identified through experimental research, while others need clarification.

As individuals turn toward internet forums for health information, it is important to determine how these evaluation processes map onto the structure of internet forums,

and how they can be used to calibrate the appropriate level of trust. Forum users have reported that they trust information in posts written by people like them (Brady, Segar, & Sanders, 2016) and that homophily drives whether a user will adopt the advice (Wang, Walther, Pingree, & Hawkins, 2008). However, some research suggests that forum specific cues, such as an indicator of the contributor's expertise on the topic, may override initial biases towards posts written by like-minded individuals (Liao & Fu, 2014). This suggests that forum specific cues are valuable tools to help calibrate an appropriate level of trust towards information. Other forum specific cues, such as the bandwagon effect, have yielded mixed results concerning their role in the evaluation of information on internet forums. In one study, high community ratings did not result in a significantly higher perceived credibility (Kim & Sundar, 2011). However, in a different study, the presence of high community ratings strengthened the perceived credibility of information in posts (Matthews, 2015). Community ratings may map onto the endorsement heuristic, suggesting the potential to have a bigger effect on the judgment if the individual takes the peripheral route in evaluation. This is especially important to clarify as it has been shown that medical professionals do not always agree with community endorsements. Still, this effect has yet to be measured in health forums, where the advice offered has the potential to affect the health and safety of the information seeker.

It is clear that there are further complications to understanding information evaluation on internet forums. While grammar and spelling errors within the written content matter to some individuals looking for health advice, others need the information enough that they look past those errors in hopes of finding valuable advice (Lederman et

al., 2014). This highlights the need for further understanding of how individual differences play a role in health forum evaluation. Due to the differences in forum user characteristics as well as individual differences in susceptibility to internet fraud (Williams et al., 2017), it is likely that evaluation techniques will differ based on individual qualities of forum users. This may include differences in health literacy, internet use, general health status, and demographic variables. Overall, there is a lack of clarity of individual and forum specific cues and how they affect perceived credibility and thus future behavior. It is important to clarify these issues experimentally in order to expand knowledge of the nature of information evaluation within the context of internet health forums.

Present Study

The purpose of the present study was to clarify how individuals evaluate forum information and the role of credibility cues and context in evaluation mechanisms, applying an understanding of heuristics and the ELM as a theoretical framework. In order to test this framework, this study experimentally varied four variables. First, community star ratings were varied to better understand how heuristics, such as the endorsement heuristic, play a role in forum evaluation. The presence or absence of spelling errors was varied to better understand how forum users interpret flawed internal content written by other users. This study also varied whether the post was life-threatening or non-life-threatening to manipulate importance of the post since posts like these often co-occur in forums. Lastly, the level of involvement each participant took in evaluation was manipulated to better understand the evaluation route taken by the participant. The following dependent variables were measured: credibility ratings of the advice, retention

of the information in the forum post and response, task completion time, and perceived intellectual ability of the forum post writer.

Qualitative responses were also collected to help clarify the role of the ELM in evaluation behavior. All participants were asked to describe what information they relied on when evaluating the post. Responses were coded to detect comments about internal argument quality, external cues such as star ratings, or both, and these responses served as another dependent variable. Age, health literacy, forum use, internet use, and health status were measured to study individual differences in evaluation behavior.

The following hypotheses were tested in the present study:

Forum Specific and Individual Cues

Hypothesis 1 - Post ratings: According to the endorsement heuristic, individuals are more likely to believe information and sources in which others have expressed belief (e.g. Metzger & Flanagin, 2013). In the case of internet health forums, star ratings and other public rating systems represent community judgments of the post. It has been suggested that internet users tend to automatically trust content from unknown persons if it has been previously recommended by others in the community (Metzger, Flanagin, & Medders, 2010; Metzger & Flanagin, 2013). *Therefore, highly rated posts are expected to result in higher credibility ratings than low-rated posts.*

Hypothesis 2 - Spelling errors: Spelling errors within website text may affect information credibility (Rowley & Johnson, 2013; Metzger & Flanagin, 2010), while spelling errors within interpersonal communication (i.e., email messages) may influence perceived conscientiousness, intelligence and trustworthiness of the email writer (Vignovic & Thompson, 2010). *Therefore, forum posts with accurate spelling will result*

in higher credibility ratings than forum posts with spelling errors. A post with accurate spelling will be judged to be written by an author with higher intellectual ability, while the same post with spelling errors will be judged to be written by an author with lower intellectual ability.

Trust Cue Conflict

Hypothesis 3 - Star ratings and spelling: Individuals may have an initial trust towards highly-rated posts in accordance with the endorsement heuristic. However, the expectancy violation heuristic posits that items that violate an individual's expectations concerning a source compel them to more thoroughly appraise the source of violation (Burgoon, Stern, & Dillman, 1995). A highly rated forum post that is spelled accurately or a poorly rated forum post that has spelling errors may not activate the expectancy violation heuristic because in either case the star ratings and the message content both convey the same message about the post quality.

By contrast, highly rated posts that have spelling errors may activate the expectancy violation heuristic because the second cue suggests a different level of credibility than the first cue. According to the ELM there are many reasons why an individual elaborates more or less during evaluation, and certain cues within the information are treated as either central or peripheral cues depending on the level of elaboration (Petty & Wegener, 1999). In the case of internet forum cues, the potential expectancy violation that comes from highly rated posts with spelling errors may act as one reason individuals elaborate more during evaluation. *Therefore, when participants view posts that violate the expectancy violation, they will take a more central route of evaluation. This will be*

shown through an interaction between star ratings and spelling errors for task completion time and information retention.

Context - Severity of Health Issue

Hypothesis 4 - Life-threatening vs. non-life-threatening post: One issue that may affect the level of elaboration during information evaluation is perceived importance of the issue (Petty & Wegener, 1999). Thus, life-threatening and non-life-threatening scenarios are likely to be evaluated differently. Posts that include a potentially life-threatening condition may result in a higher level of perceived importance due to the more serious consequences associated with offering advice. *Therefore, individuals who read questions and advice about life-threatening health situations will likely take a more central route of processing. This effect is expected to be stronger under conditions of higher involvement. Therefore, an interaction between context and involvement will be observed for task completion time and information retention scores.*

Involvement

Hypothesis 5 - Low Involvement vs. High Involvement: Previous studies have manipulated participants' level of involvement in the evaluation of consumer-written reviews. Results of these studies showed that participants with low involvement relied on peripheral cues such as review titles while participants with high involvement relied on central cues such as internal argument quality (Yang, Hung, Sung, & Farn, 2006; Lee, Park, & Han, 2008). While this effect has not been examined in the internet forum context, the domain is similar in that reviews and forum advice are written by the user, not necessarily by a professional in that domain. *Therefore, individuals with a higher level of involvement are expected to take a more central route of processing when*

evaluating questions and advice on a health forum shown through a longer task completion time and higher information retention scores than individuals with low involvement.

Individual Differences. Individual differences between people such as variations in internet use, health status and health literacy were examined for effects on evaluation behavior during the study. The role of individual differences in the context of health forum use is less studied than traditional internet use, in part due to the sample sizes required to investigate these questions. Consequently, there was little existing research in the internet forum domain to directly guide hypotheses. Therefore, the role of individual differences in the present study was guided by findings in related areas but took more of an exploratory approach.

Internet Use. Prior research suggests that trust in online information is calibrated over time with experience, and that users with more experience using the internet tend to be more confident in the information and information owners (Dutton & Shepherd, 2006). At the same time, those who have more internet experience are likely to have experienced trust-testing online events such as receiving fraudulent or obscene email messages, and are consequently more aware of the potential for encountering information of low credibility. *Therefore, individuals with more internet use may be more trusting of forum posts, potentially resulting in higher perceived credibility of the posts. Because individuals with more internet experience are more aware of potentially fraudulent experiences they may take the central route and evaluate the forum post more thoroughly, shown through a longer task completion time.*

Health Status. Those with health concerns may find health information to be more personally relevant, and therefore may be more interested in it. Prior work has shown that individuals who are personally interested in a topic may take a more central route when evaluating information associated with the topic (Yang, et al., 2006).

Therefore, individuals with worse health status may be more interested in reading health forum information and consequently take a more central route of processing while evaluating it, shown through a longer task completion time.

Health Literacy. Previous work has found that lower income individuals use social health sites more than higher income individuals (Atkinson, Saperstein, & Pleis, 2009), perhaps as a supplement to or even replacement of formal health care. Unfortunately, evidence suggests that persons with lower incomes may have lower health literacy levels (Schillinger et al., 2006), which may hinder their ability to adequately comprehend online health information. Because of the greater effort involved in processing written information when literacy skills are lower, evaluation techniques may vary across literacy levels. Specifically, *individuals with lower health literacy levels may take a more peripheral route of evaluation, and external cues such as star ratings will moderate credibility ratings. This will be shown through an interaction between star ratings and health literacy scores when predicting credibility ratings.*

CHAPTER II

METHOD

Pilot Testing

Pilot testing was conducted to ensure that the forum posts and responses differed significantly in importance by context, and that the spelling errors were salient enough to be detected. Data were collected from 40 participants using Mechanical Turk (MTurk). First, participants were randomly assigned to either the life-threatening or non-life-threatening question. After reading the forum question, participants responded to items assessing the importance of the question. The life-threatening question received significantly higher ratings on every importance item than the non-life-threatening question. Therefore, it was concluded that importance was manipulated successfully by context.

Participants were then randomly assigned to the appropriate response with either correct spelling or with spelling errors. The results indicated that participants detected the spelling errors, and that they thought the spelling errors took away from the overall advice quality. Therefore, we concluded that the responses with spelling errors were calibrated correctly for the final study. Pilot results can be viewed in Appendix K.

Participants

Initial data were collected from 418 participants (228 male, 187 female, 3 prefer not to answer) who were recruited via MTurk. The average age of participants was 36.72 years old ($SD = 10.52$). Data were collected in batches and distributed across various times of day over the course of two weeks to reduce the likelihood of time-of-day or day-of-week effects on participant demographics (Ipeirotis, 2010). The survey was timed and

following each batch of data, completion times were compared to determine whether participants spent sufficient time responding to the survey. Participants who clicked through items without spending sufficient time to read the materials were excluded from the data set, and more data was collected to ensure the correct sample size. Over the course of data collection, twenty-six participants were excluded because they spent significantly less time than other participants on multiple items in the survey. Ten participants were excluded because they were from duplicate IP addresses, and eight participants were excluded due to being outside of the United States.

After data collection was complete, additional measures were taken to ensure data quality. Twelve participants were excluded because their forum post task completion time was greater than three standard deviations above the mean, suggesting they stayed on the page while doing other activities (i.e., a participant spending 11 minutes on a page when the mean response time was 2 minutes and 2 seconds). These participants were excluded due to the possibility of stepping away from the task or doing other tasks while completing the survey.

Additionally, leverage values for each regression model were individually calculated using Mahalanobis Distance and these values were plotted and contrasted against established cutoff values (Tabachnick & Fidell, 2001). Leverage values determine how far observed values are from mean values in selected scales to detect discrepancies in data, ultimately allowing for the identification of outliers. Following this procedure, ten participants were excluded from the Internet Usage Scale, and one participant was excluded from the eHEALS. Otherwise, all values fell within acceptable standards.

Finally, a subset of data (n=221) was coded for the information retention variable and for the free response subjective evaluation item regarding what information each participant used to rate the advice. During data collection, the Subjective Evaluation Criteria item was initially erroneously excluded from the full survey. After collecting data for 158 participants, the question was inserted and data collection continued to ensure that the proposed sample size was collected for this variable. Therefore, all of the data coding for both the information retention variable and the free response variable was conducted for the last 221 participants. Comparison analyses were conducted to ensure no significant differences in data resulted from the non-random nature of data collection for this item. There were no statistically significant differences between participants before and after the item was added in age, perceived credibility, perceived intellectual ability, or task completion time. Therefore, the full dataset was used for main comparisons using these variables, and the subset of data was used to analyze research questions regarding information retention and subjective evaluation techniques. See Figure 2 for a full description of data exclusion and resulting sample size for analyses.

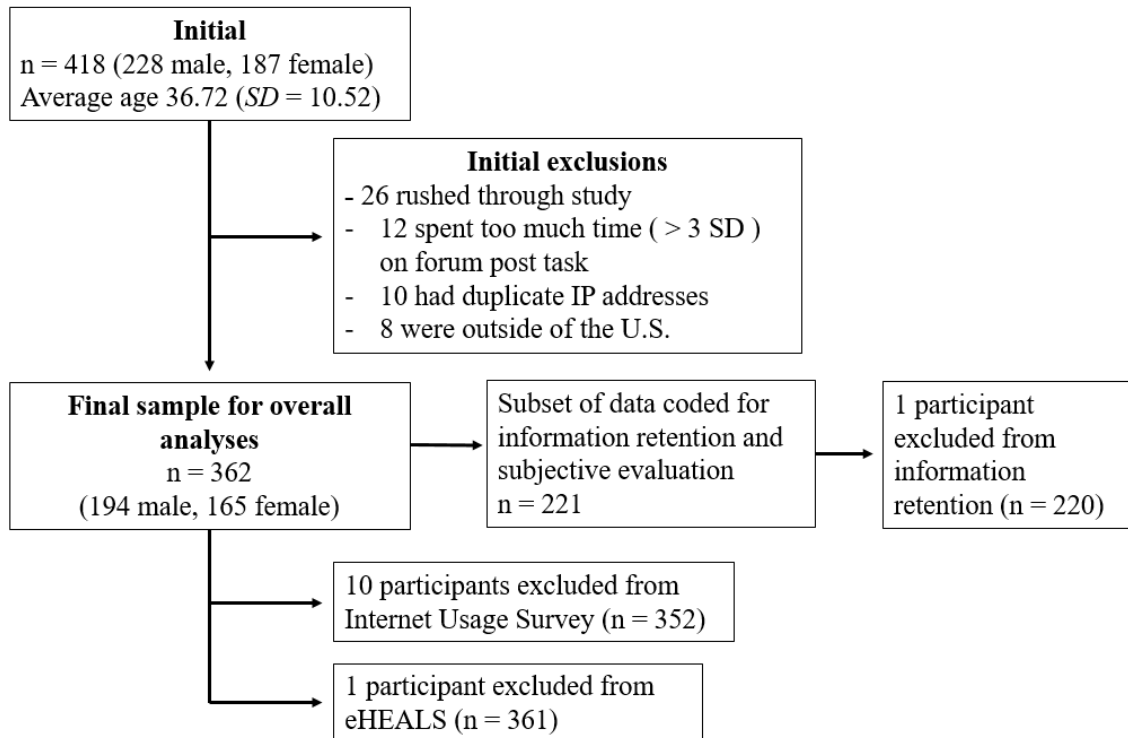


Figure 2. Exclusion criteria during data collection and cleaning. All sample sizes reported are for analyses with respective scale/item.

Materials

Demographic Questionnaire. Demographic information including age, gender, racial or ethnic identification, location (state, country), and primary and secondary language was collected. Other questions regarding income range, employment status, student status, and Mechanical Turk (MTurk) status were asked.

Forum Post Evaluation Task. The Forum Post Evaluation Task served as the experimental portion of the study. An example forum post and response is shown in Figure 3 (for larger version and full text see Appendix A).

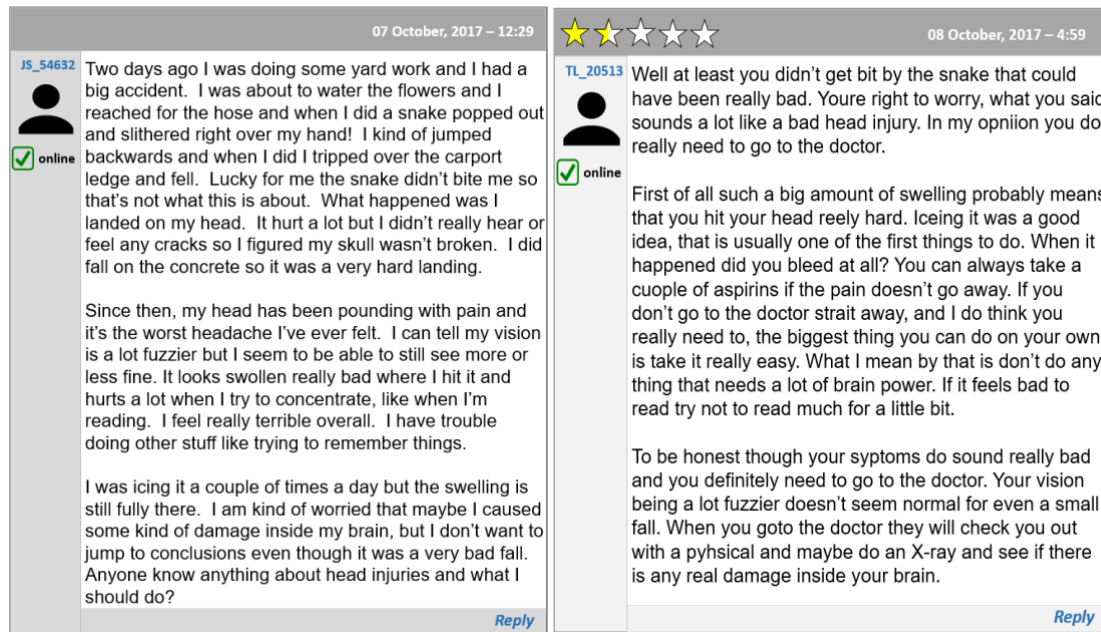


Figure 3. Life-threatening forum post followed by a forum response with low star ratings and spelling errors. See appendix A for larger images.

Each participant was presented with a mocked up forum thread consisting of an initial question posed by a user (left side of Figure 3) and a single response given by another user (right side of Figure 3). Depending on the assigned condition, participants saw a forum post with high or low star ratings, text with correct spelling or spelling errors, and a life-threatening or non-life-threatening question. Both forum posts were designed to be almost identical except for the subject matter. The life-threatening question was regarding a head injury, while the non-life-threatening question was regarding a wrist injury. When possible, all words were kept the exact same, except those that changed the context of the injury. For example, the sentence, “What happened was I landed on my wrist”, was written as “What happened was I landed on my head” in the life-threatening version. The response to each question was structured the same way, in that the content was the same except for injury specific words. The spelling errors were the same in both conditions.

Perceived Credibility Scale. Participants responded to a five-item perceived credibility scale (Flanagin & Metzger, 2000; Appendix B) developed to examine internet message credibility. This scale measured the main components of credibility by assessing believability, accuracy, trustworthiness, bias, and completeness of information. Prior work suggests that expertise is an important factor in determining credibility (Hovland, Janis, & Kelley, 1953). Therefore, an additional item assessing the expertise of the advice was included.

Intellectual Scale. The Intellectual Scale (Warner & Sugarman, 1986; Appendix C) included five semantic differential items (incompetent/competent, ignorant/knowledgeable, irresponsible/responsible, unintelligent/intelligent, foolish/sensible) used to assess the perceived intellectual ability of the advice author. Participants rated the advice writer on a scale using each item's anchors.

Importance Items. Importance items included four questions assessing how important participants perceived the forum post to be. Participants rated the questions on a seven-point Likert style scale.

Subjective Evaluation Criteria item. The Subjective Evaluation Criteria item was a subjective measure developed to collect self-reports about the factors influencing participants' credibility ratings. This measure was a free response question that asked participants to describe what they relied on to evaluate the health forum post.

Motivation Item. One question asked participants to rate how motivated they were to read the forum post and response and answer the questions that followed.

Expertise Items. Expertise items were questions regarding participants' expertise within the context of the forum post task. Specifically, participants were asked whether

they or someone close to them had ever experienced the injury they saw in the post, and if they had any official medical or health training or education (Appendix D).

Instructional Manipulation Check (IMC). An IMC (Oppenheimer, Meyvis, & Davidenko, 2009) is a measure to detect participant satisficing in a particular study, and to remove these participants from the study early on. The IMC developed for this study was a paragraph of instructions with check boxes of various health issues below it. In the middle of the paragraph, text asked participants not to select any of the options, but rather to click the continue button (Appendix E).

The eHealth Literacy Scale (eHEALS). The eHEALS (Norman & Skinner, 2006; Appendix F) measured the ability to both find and use health information online. The eHEALS is a self-report tool that provides an estimate of participants' eHealth skills, based on the individuals' perception of their own skills and knowledge. The scale was composed of eight items and used a 5-point Likert scale ranging from "strongly agree" to "strongly disagree." The questions targeted six core literacies generally associated with eHealth literacy: traditional literacy, health literacy, information literacy, scientific literacy, media literacy, and computer literacy.

Short Test of Functional Health Literacy in Adults (S-TOFHLA). The S-TOFHLA (Baker, Williams, Parker, Gazmararian, & Nurss, 1999; Appendix G) is designed to test the individual's ability to read and understand real materials from the health care setting. It was created using passages from pre-surgery protocol and Medicaid application materials. The S-TOFHLA is made up of various paragraphs with fill in the blank questions throughout. Participants were asked to fill in each blank with a word provided in a list of multiple choice items.

SF-36 Health Survey. The 36-item version of the Short Form Health Survey (SF-36; Ware & Sherbourne, 1992; Appendix H) was used to measure health. The SF-36 is a validated 36 item tool that measures quality of life by assessing physical, mental, and general health. It has been used widely for a broad range of purposes by both health care organizations and researchers. The SF-36 returned eight total scores: physical functioning, role limitations due to physical health, role limitations due to emotional problems, energy/fatigue, emotional well-being, social functioning, pain, and general health.

Internet Usage Survey. The internet usage survey developed for this study consisted of a subscale drawn from a larger Technology Experience Survey (Pak, Rovira, McLaughlin, & Leidheiser, 2017; Appendix I). Items on this subscale measured how much the participant uses different platforms on the internet, such as social media, weather, shopping, and health sites. Additional questions addressing if and how often users get health information from others online as well as their level of participation in internet forums were asked and analyzed separately.

Retention Check Quiz. The retention check quiz consisted of a free recall question and an eleven-item quiz. The free recall questions instructed participants to write down as much as they could remember from the forum post and the forum response in as much detail as possible. The eleven-item quiz then assessed participants' memory of the content by asking questions such as "What type of medicine did the advice giver recommend?" The quiz included both multiple choice and fill in the blank items (Appendix J). Both items were administered to ensure a wide range of retention data was collected.

Design

The present study was a 2 x 2 x 2 x 2 between subjects experiment. The experimentally manipulated independent variables of the study and their levels were as follows: 1) Community star ratings: Highly rated posts (4.5 stars) versus low-rated posts (1.5 stars); 2) Spelling: Accurate spelling versus spelling errors; 3) Health context: Life-threatening versus non-life-threatening; 4) Involvement: Higher versus lower involvement. Other independent variables measured were age, health literacy, internet use, and personal health. The main dependent variables of the study were credibility ratings assigned to the forum posts (Perceived Credibility Scale), the perceived intellectual ability of the advice giver (Intellectual Scale), task completion times (time taken to read both the forum question and response), self-reports of the credibility assessment process (Subjective Evaluation Criteria item), self-reported task motivation, and information retention (Retention Check Quiz).

Procedure

After opting into the survey and giving informed consent, qualified MTurk workers completed all portions of the study through Qualtrics. First, they completed the demographics questionnaire, followed by the Forum Post Evaluation Task (Figure 4).

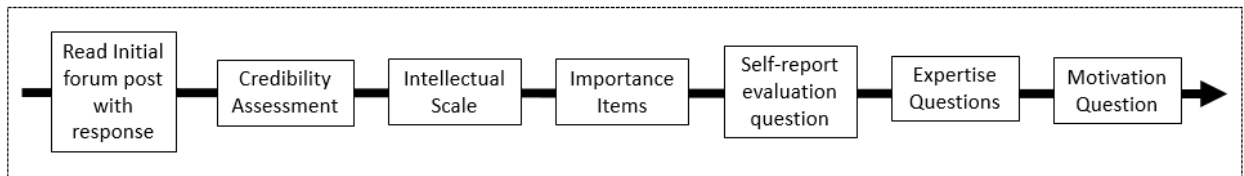


Figure 4. Flow of Forum Post Evaluation Task. This figure illustrates the flow that participants will work through during the Experimental task, starting with reading an initial forum post with response.

During this task, participants first read instructions that explained the nature of the forum post they would see. Specifically, they were told that that the post was made by

someone seeking help online who received advice from a forum community member. After reading the general instructions, each participant was randomly assigned to either the low involvement or high involvement condition. Those in the low involvement condition began the task directly after the general instructions. Those in the high involvement condition received further instructions that were meant to suggest a more serious nature of the task and thus a higher involvement in reading and evaluating the question and advice. They were asked to imagine that a friend had the same medical condition and needed help evaluating whether it was good advice or not. The friend wanted the participant to thoroughly evaluate the advice and write out an evaluation to aid them in making the decision. Participants were told that if they provided a good evaluation to their hypothetical friend they would be given an extra incentive after participating. In actuality, participants in the high involvement condition were not asked to write the hypothetical email at the end of the study, but were debriefed and given the extra dollar.

After reading all instructions, participants in both conditions continued to the next screen with their assigned version of the forum post and response. After reading the post, they responded to the Credibility Assessment and the Intellectual Scale. Next, participants completed the Self-Report Evaluation Measure and the Expertise Questions. Directly following the task, all participants indicated how motivated they were to evaluate the forum post. All scales and questions given during the experimental task can be found in Appendix B, and the Forum Post Evaluation Task flow is shown in Figure 3.

Following the Forum Post Evaluation Task, participants were routed to the IMC, where they read a large paragraph of instructions that told them not to select any items on

the screen, but to continue to the next page. At this point, any participants who failed to follow these instructions were re-routed to a custom end of survey message explaining that they did not follow the guidelines of the study and why they would not receive payment. Participants who followed instructions in the IMC then completed the eHEALS and the s-TOFHLA, the SF-36, and the Internet Usage Survey. The time it took to complete these questionnaires served as an inter-trial retention interval before the Retention Check Quiz, during which participants responded to items testing their memory of the information from the forum post. Upon completion of the study, participants were debriefed and awarded monetary compensation. Figure 5 depicts the task flow for the entire study.

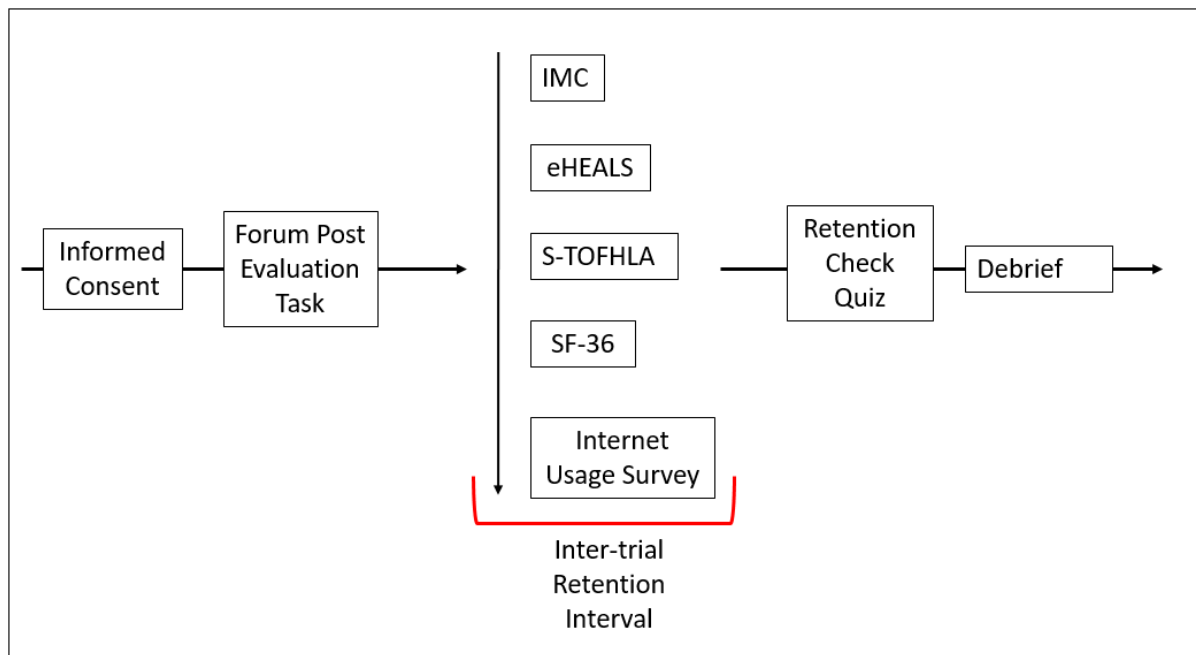


Figure 5. Task flow for entire study beginning with informed consent and ending with debriefing.

Variable coding and organizing

Retention Check Quiz. Prior to collecting data, both forum posts and responses were broken down into idea units consisting of a subject and a predicate. During the

experiment, participants were instructed to recall as much as they could about the forum question and respond in as much detail as possible. A team of researchers were trained to code each free recall response according to the idea units previously developed. An idea unit was counted if the “gist” of the statement was recalled (Turner & Greene, 1978; Dixon, Hultsch, Simon, & Von Eye, 1984). A random selection of 20 questions and responses were pulled from the data set and scored for accuracy. The number of agreements was divided by the total number of idea units to ensure that each scorer maintained an inter-rater reliability of at least .90. The final information retention score for each question and response was expressed by the number of idea units correctly recalled. Following the overall exclusions, one participant was excluded for writing an unrelated paragraph in response to the question.

Subjective Evaluation Criteria item. Directly after completing the forum post evaluation task, participants were asked to list, in detail, any piece of information they used to evaluate the credibility of the advice in the forum post. First, a random selection of responses was reviewed by two researchers. Then, the researchers agreed on seven themes that emerged as evaluation criteria: 1) Recommending a doctor – participant mentioned that advising to go to the doctor, or not go to the doctor, played a role in their evaluation; 2) Source expertise – participant mentioned the credentials of the post writer, stating the presence or absence of expertise or knowledge on the subject; 3) Spelling errors – participant noted the spelling and grammar of the post; 4) Star ratings – participants noted the star ratings assigned to the post; 5) Experiential evidence – the participant assumed the advice was written based off of personal experience; 6) Personal knowledge – the participant compared the advice to their own knowledge on the subject

and made a judgment accordingly. All responses were then coded using those themes as dichotomous (present or absent) criteria. A sample response with the resulting coding is shown in Figure 6.

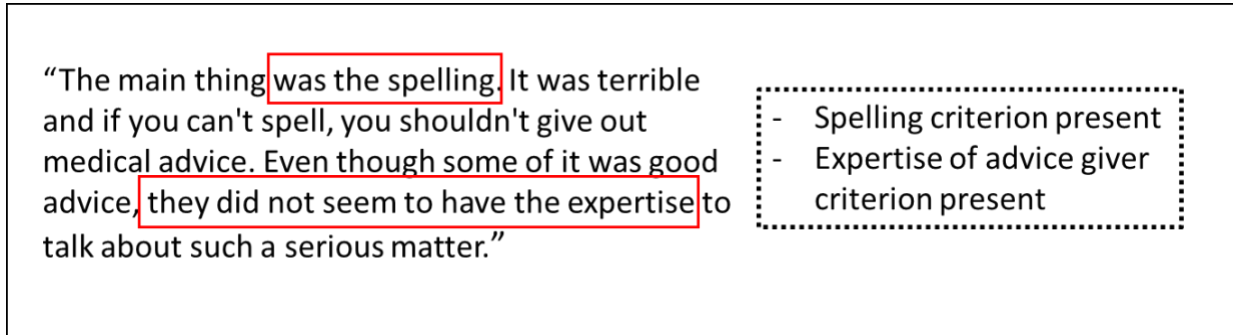


Figure 6. A sample response to the subjective evaluation criteria item, shown with coded criteria.

Credibility scale. The adapted credibility scale included an extra item evaluating the perceived expertise of the advice. A reliability test of the original scale revealed a Cronbach's alpha of .856, and a follow up test with the additional item revealed a Cronbach's alpha of .867. Therefore, the item was averaged with the other scale items to yield one credibility score.

Importance scale. The importance scale was created for the study to see how important participants thought each question was. The four items yielded high reliability ($\alpha = .911$). Therefore, these items were averaged to create one measure for importance of question.

CHAPTER III

RESULTS

Manipulation Check

An initial analysis was conducted to determine whether participants rated the life-threatening medical question as more important than the non-life-threatening medical question. An independent samples t-test revealed that the life-threatening question was rated as statistically more important ($M = 6.27$, $SD = .69$) than the non-life-threatening question ($M = 3.18$, $SD = 1.00$), $t(360) = 34.15$, $p < .001$, with an effect size of $d = 3.09$.

Demographics

Because the experimental material varied by context, demographics were compared between the life-threatening and non-life-threatening groups to ensure group characteristics were not a confounding factor on the present study. All results were insignificant, and the p values can be seen in Table 1 along with the overall demographics.

Table 1

Overall demographics and demographic comparisons by context.

	Overall	Life-threatening	Non-life-threatening	
Sample Size	362 ^a	183	179	
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>p</i>
Age	36.72 (10.52)	36.97 (10.79)	36.47 (10.25)	.65
Education ^b	3.49 (1.10)	3.46 (1.07)	3.53 (1.12)	.60
Years working for MTurk	2.65 (1.76)	2.59 (1.70)	2.71 (1.82)	.53
Hours per week on MTurk	21.71 (12.63)	22.22 (12.73)	21.18 (12.55)	.43

^aAge $n = 360$.
^bRanging from 1 'High school incomplete' to 6 'Postgraduate degree'

Total means were calculated for each dependent variable for star ratings, spelling, context, and involvement, and can be viewed in Table 2.

Table 2.

Displayed means for main dependent variables by main independent variables.

	Perceived Credibility		Perceived Intellectual Ability		Task Completion Time			Information Retention ^a	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>
Star Ratings									
Low (n = 186)	3.10	.90	4.44	1.45	102.74	47.85	n = 112	11.60	4.85
High (n = 176)	3.32	.88	4.90	1.50	102.20	45.43	n = 108	12.07	4.50
Spelling									
Good (n = 186)	3.40	.85	5.01	1.45	103.87	45.06	n = 111	11.86	4.18
Errors (n = 176)	3.00	.90	4.30	1.45	101.01	48.31	n = 109	11.81	5.15
Context									
Life-threatening (n = 183)	3.24	.86	4.66	1.47	97.66	42.30	n = 113	12.41	4.78
Non-life-threatening (n = 179)	3.17	.93	4.66	1.51	107.42	50.31	n = 107	11.22	4.51
Involvement									
Low (n = 182)	3.17	.90	4.64	1.50	99.98	45.27	n = 111	11.60	4.72
High (n = 180)	3.24	.89	4.68	1.48	105.01	47.95	n = 109	12.06	4.65

^aA subset of data was coded for information retention, respective sample sizes denoted in table.

Perceived Credibility

A 2 (context) x 2 (involvement) x 2 (star ratings) x 2 (spelling) factorial ANOVA was run to examine any main effects of the independent variables on perceived credibility

of information. The means and standard deviations of credibility ratings for each level of each independent variable can be seen in the second and third columns of Table 2, and the results of the ANOVA can be viewed in Figure 7. As seen in Figure 7, post with high star ratings yielded higher credibility ratings than posts with low star ratings. Also, posts written without spelling errors received higher credibility ratings than posts written with spelling errors. The results inferred from Figure 7 were confirmed by the factorial ANOVA, which yielded a main effect of star ratings, $F(1,346) = 5.66, p = .02, \eta_p^2 = .016$, with posts with high stars receiving higher credibility ratings ($M = 3.32, SD = .88$) than posts with low star ratings ($M = 3.10, SD = .90$). There was also a main effect of spelling, $F(1,346) = 17.67, p < .001, \eta_p^2 = .049$, with posts without spelling errors receiving higher credibility ratings ($M = 3.40, SD = .85$) than posts with spelling errors ($M = 3.00, SD = .90$). No significant interactions were found for perceived credibility between independent variables.

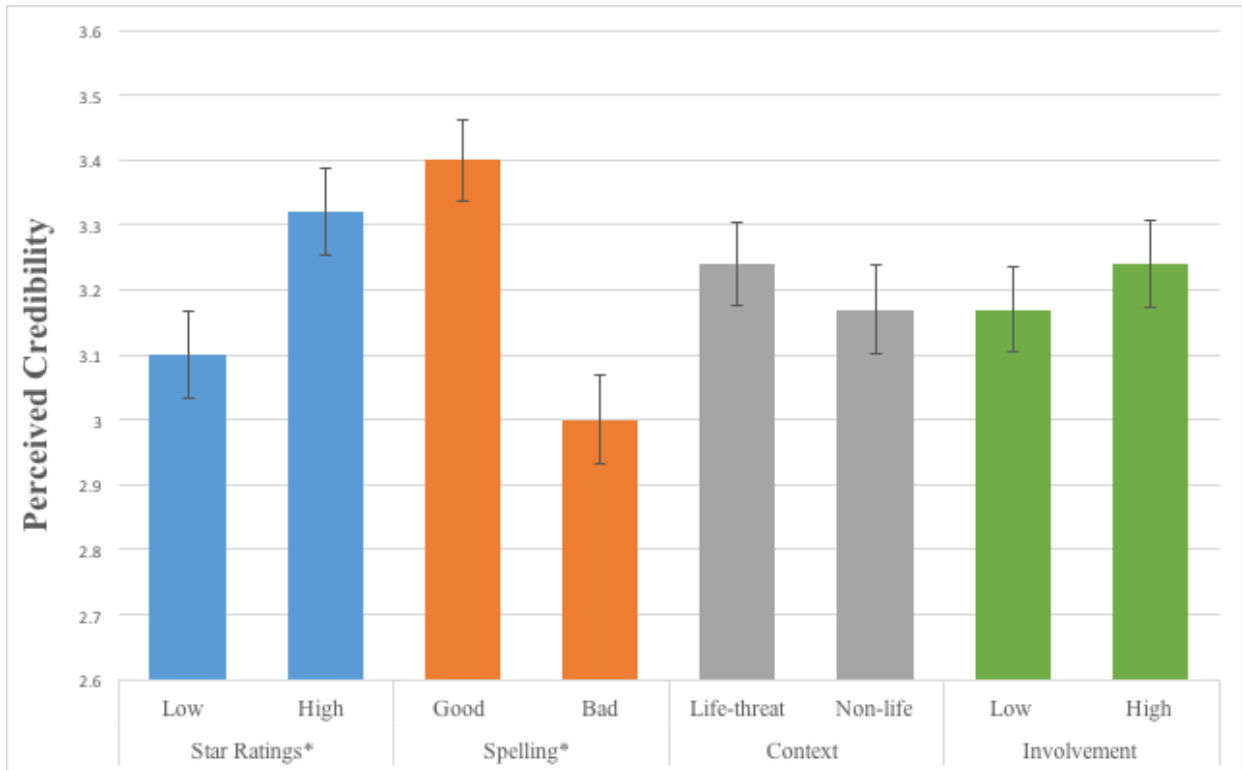


Figure 7. Differences in perceived credibility for each independent variable. *Significant at the $p < .05$ level.

Perceived Intellectual Ability

To examine main effects of each independent variable on perceived intellectual ability of the author, a 2 (spelling) x 2 (star ratings) x 2 (context) x 2 (involvement) factorial ANOVA was run. Mean ratings of perceived intellectual ability can be seen in the fourth column of Table 2, and the comparisons of each independent variable can be viewed in Figure 8. It appears that authors of posts with high star ratings received higher intellectual ability ratings than authors of posts with low star ratings, and authors of posts without spelling errors received higher intellectual ability ratings than authors of posts with spelling errors. These impressions were confirmed in the factorial ANOVA, which indicated a significant main effect of spelling, $F(1,346) = 20.45, p < .01, \eta_p^2 = .056$, where authors of posts written without spelling errors receiving higher ratings of

intellectual ability ($M = 5.01$, $SD = 1.45$) than authors of posts with spelling errors ($M = 4.30$, $SD = 1.45$). Additionally, a main effect of star ratings was observed, $F(1,346) = 8.94$, $p < .01$, $\eta_p^2 = .025$, where authors of highly rated posts received higher ratings of intellectual ability ($M = 4.90$, $SD = 1.50$) than authors of low-rated posts ($M = 4.44$, $SD = 1.45$). No other significant main effects or interactions were observed. Out of 110 participants who saw a post with spelling errors, 46 (41.8%) noted the spelling errors in their evaluation.

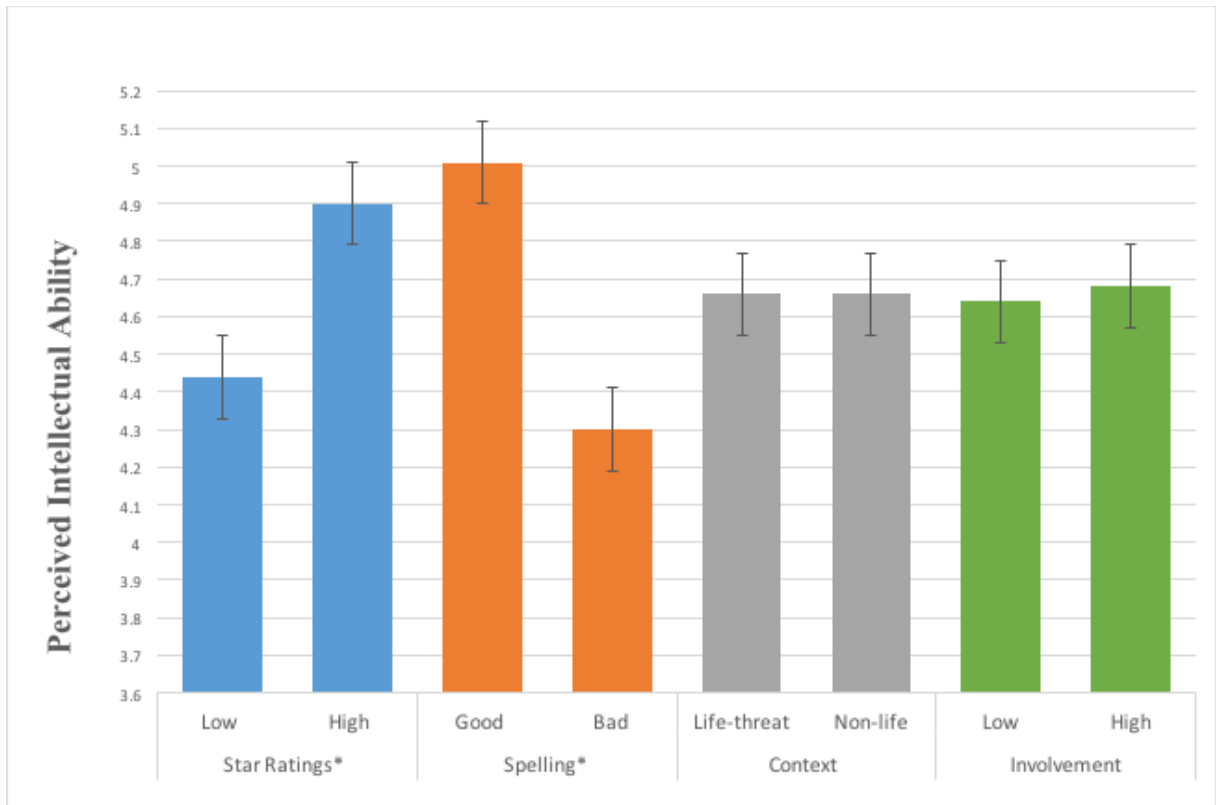


Figure 8. Differences in perceived intellectual ability between independent variables.
*Significant at the $p < .05$ level.

Task Completion Time

To examine the effect of condition on task completion time, a 2 (spelling) x 2 (star ratings) x 2 (context) x 2 (involvement) factorial ANOVA was run. As seen in

Figure 9, participants assigned to the non-life-threatening posts spent significantly longer reading the post than participants assigned to the life-threatening post. No other differences in task completion time can be seen.

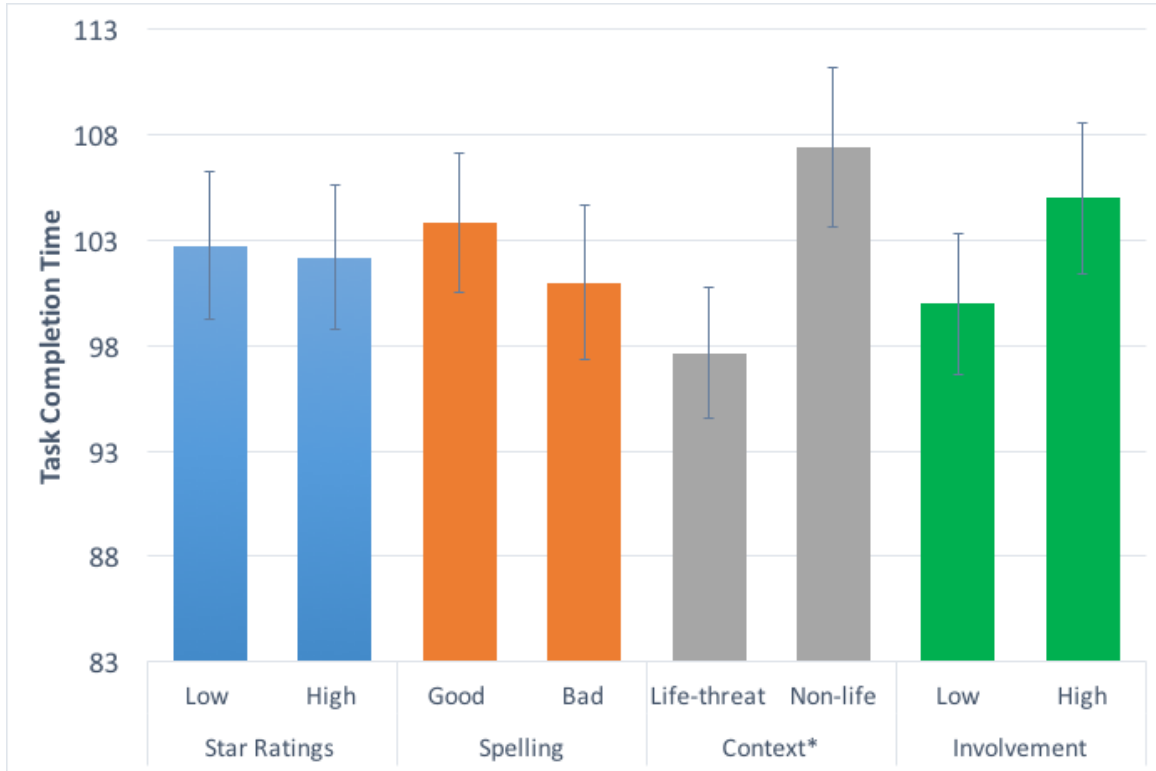


Figure 9. Differences in task completion time between independent variables.

*Significant at the $p < .05$ level.

The results of the ANOVA confirmed the significant main effect of context, $F(1,346) = 3.89, p = .049, \eta_p^2 = .011$, where participants took significantly more time (in seconds) to read the non-life-threatening post ($M = 107.41, SD = 3.76$) than the life-threatening post ($M = 97.66, SD = 3.13$). No other significant main effects or interactions were observed for task completion time. Figure 10 shows a lack of an interaction between star ratings and spelling for task completion time.

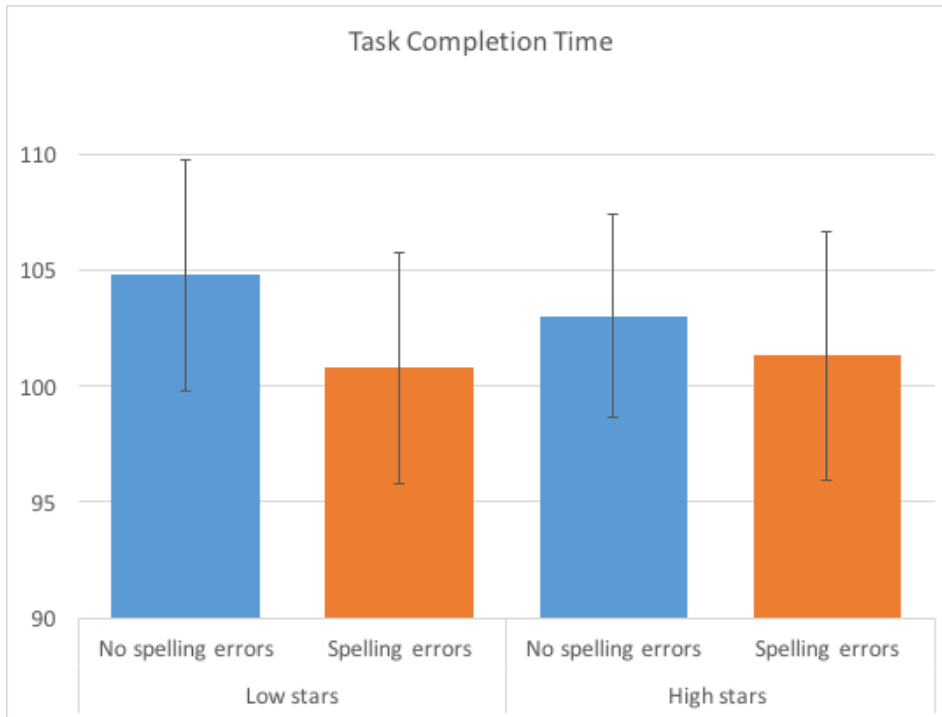


Figure 10. Interactions between star ratings and spelling errors for task completion time.

Information Retention

Following a visual inspection of both the eleven-item quiz and the free recall data, tests of normality were conducted. Data from the eleven-item quiz were not normally distributed, as assessed by Shapiro-Wilk's test ($p < .01$), due to a negatively skewed distribution (-1.0 , $SE = .13$). Because free recall data were normally distributed ($p = .10$), these data were used as the information retention variable. A 2 (spelling) x 2 (star ratings) x 2 (context) x 2 (involvement) factorial ANOVA was run to determine any differences in information retention.

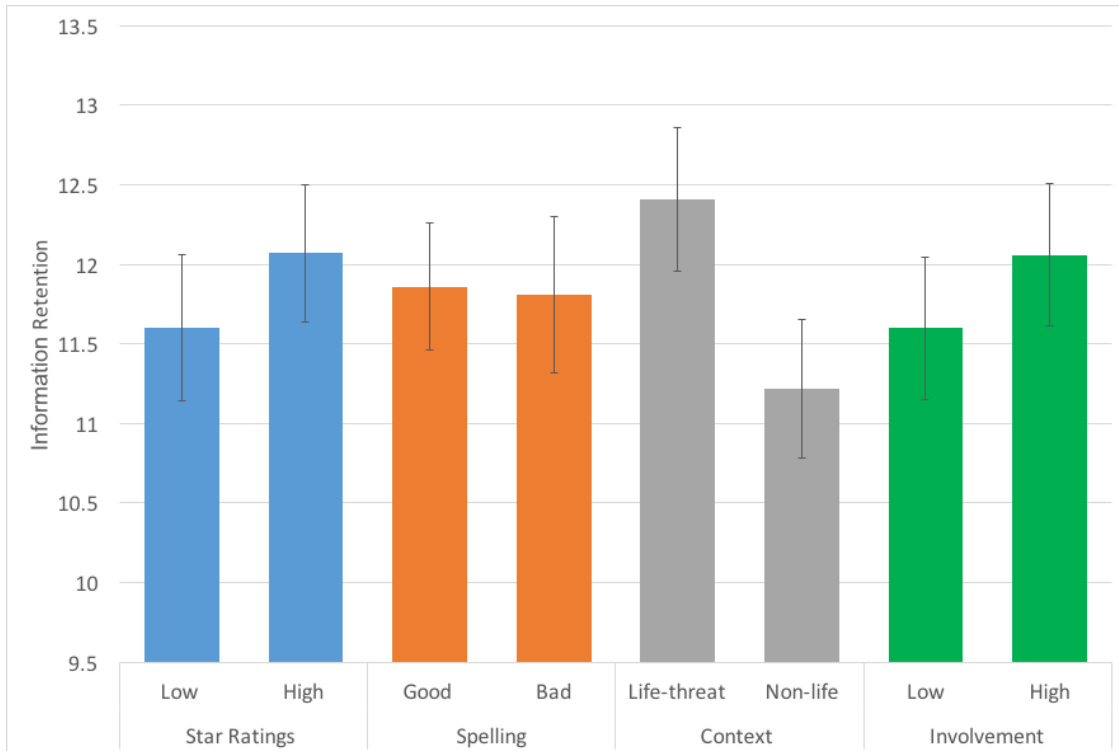


Figure 11. Differences in information retention between independent variables.

In Figure 11, it appears that those reading the life-threatening post retained significantly more information than those reading the non-life-threatening post. However, the ANOVA revealed no significant main effects or interactions between levels of each independent variable for information retention. Figure 12 shows a lack of an interaction between star ratings and spelling for information retention.

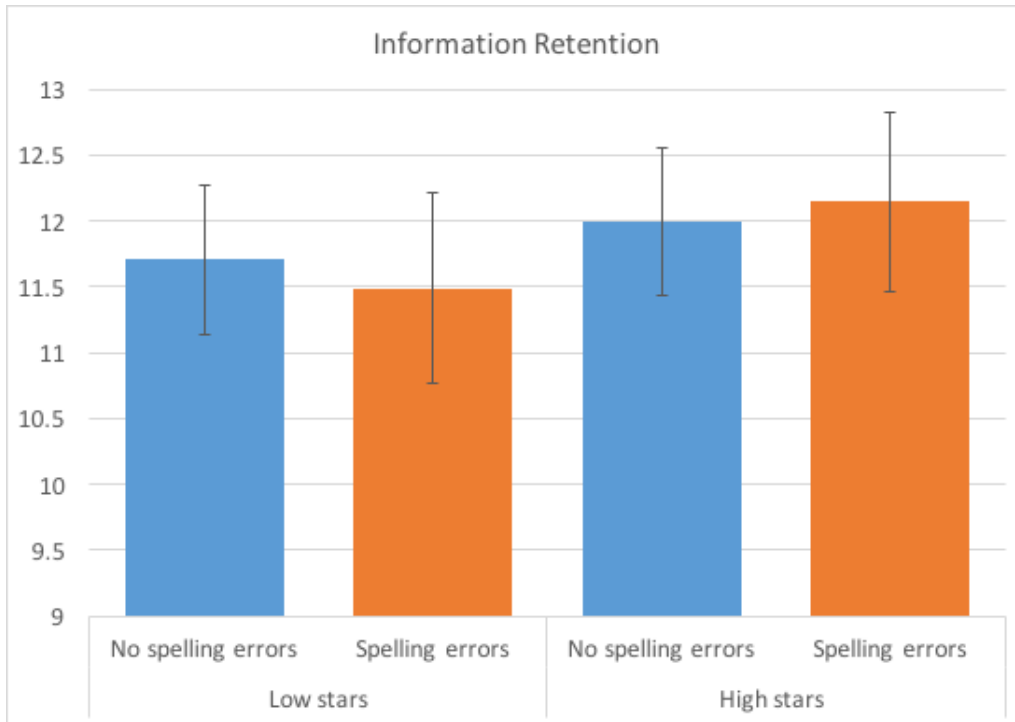


Figure 12. Interactions between star ratings and spelling errors for information retention.

Motivation

Motivation scores were not normally distributed, as assessed by Shapiro-Wilk's test ($p < .01$), due to a negatively skewed distribution (-2.28 , $SE = .13$). While an exponential data transformation reduced skewness ($-.93$, $SE = .18$), the new data were not normally distributed as assessed by a Shapiro-Wilk's test ($p < .01$). Therefore, comparisons of motivation between groups must be observed with a critical eye. A 2 (context) by 2 (involvement) factorial ANOVA was used to assess differences in motivation to complete the forum post evaluation task. The results indicated no significant difference between those in the life-threatening and non-life-threatening group, $F = 3.36$, $p = .07$, or between those in the high versus low involvement conditions, $F = .03$, $p = .87$.

Subjective evaluation of forum post

The frequencies of each subjective evaluation criterion were calculated and can be seen in Table 3

Table 3.

Frequencies of free response data sorted by topic of comment.

Topic Group	Present	Absent	% present
Doctor	110	110	50.00%
Source Expertise	39	181	17.73%
Spelling errors	47	173	21.36%
Star ratings	7	213	3.18%
Experiential evidence	5	215	2.27%
Personal knowledge	41	179	18.64%

Chi-square tests for association were run to determine whether observed dichotomous free response variables were associated with manipulated independent variables. The results indicated two significant associations. First, there was a statistically significant association between context and the ‘recommending a doctor’ criterion, $\chi^2(1) = 25.48, p < .01$. The association between the two variables was moderately strong, $\Phi = .34, p < .01$. The results of this test can be graphically viewed in Figure 13.

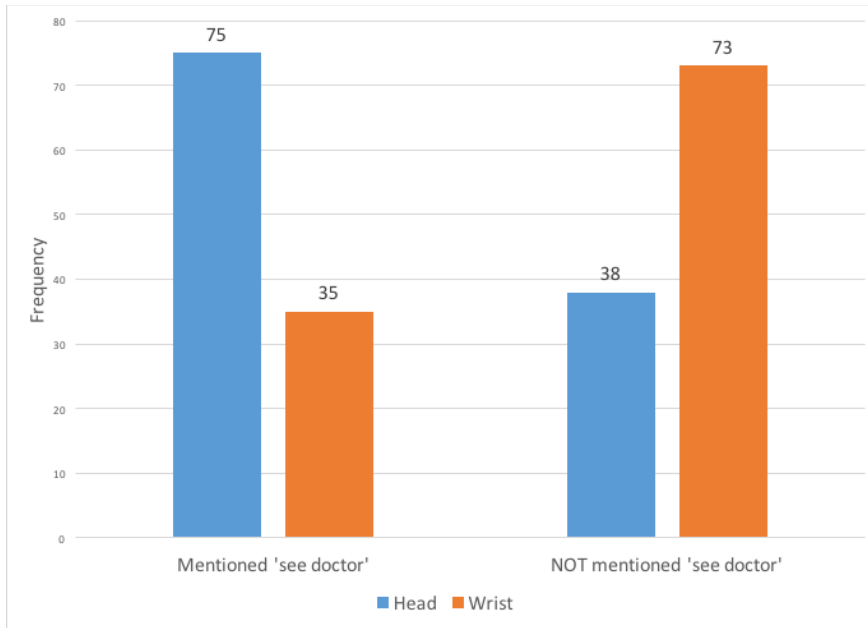


Figure 13. Results from Chi-square test of association between context and ‘see a doctor’ criterion.

Furthermore, there was a statistically significant association between context and personal knowledge criterion, $\chi^2(1) = 4.26, p = .04$, and the observed association between the two was small, $\Phi = .14, p = .039$. The results of this test can be graphically viewed in Figure 14.

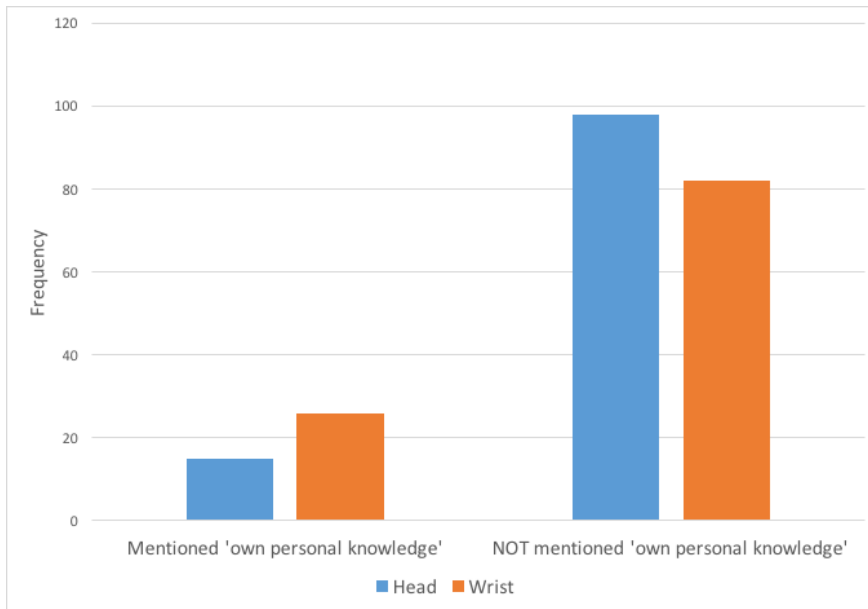


Figure 14. Results from Chi-square test of association for context and ‘own personal knowledge’ criterion.

Expertise

Participants were asked about how much experience they had with the injury they read about in the post, and were also asked if they had any formal medical education or training. The results from these questions can be seen in Table 4. Due to the lack of participants falling into the medical education or training groups, those two groups were combined with personal experience to create a dichotomous expert variable. Independent sample t-tests were run to examine the differences between experts and non-experts in credibility ratings, perceived intellectual ability, task completion time and information retention. The analyses failed to reveal any significant differences.

Table 4.

Results of questions regarding experience with health scenario, with frequencies of developed expertise variable.

	Head	Wrist	New Expert Variable	
^a Personal Experience with injury	68 (37.16%)	96 (53.63%)	Non Expert	184
^b Medical education or training	28 (15.3%)	21 (11.73%)	Expert	178

Note: Head n = 183, Wrist n = 179

a. Have you or someone close to you experience a head/wrist injury?

b. Have you had any formal medical education or training?

Individual Differences

Due to the large number of variables in the present study, initial correlations were run to identify any significant relationships between the main independent and dependent variables. Then, specialized models were created to measure the moderation effects of the categorical independent variables on the relationship between the most important predictor variables (i.e., age, health variables, internet variables, and internet health literacy) and dependent variables (i.e. perceived credibility, perceived intellectual ability, and perceived question importance). Due to the exploratory nature of the moderation analyses, only significant results were reported for analyses outside of the explicit hypotheses.

Internet Use. An initial correlation failed to reveal a significant relationship between internet use and perceived credibility, $r(350) = .01, p = .86$. Similarly, there was no significant relationship between overall internet use and task completion time, $r(350) = .05, p = .39$. Further analyses were then run to assess how different behaviors relating to internet forums impacted perceived credibility. There was a significant positive

relationship between posting to internet forums and perceived credibility of the forum posts, $r(351) = .15, p < .01$, as well as a significant relationship between posting to forums and perceived intellectual ability of the advice writer, $r(351) = .17, p < .01$.

Health Status. An initial analysis failed to find a significant relationship between general health and task completion time, $r(351) = .06, p = .29$. However, those with higher scores on the emotional well-being variable spent longer on the forum post task, $r(351) = .13, p = .02$, as well as those scoring higher in energy, $r(351) = .12, p = .03$. Due to the size of the correlations and due to their overall relevance to the present study, these variables were not investigated further.

Health Literacy. Prior to analyses, data for the s-TOFHLA were examined. These data were highly negatively skewed as the majority of participants scored nearly perfectly on the task; thus, the task was excluded from analyses. Therefore, the eHEALS for internet health literacy scale was used in all health literacy analyses. Initial analyses failed to reveal any relationship between internet health literacy and the main dependent variables of the study.

A custom model was built to assess the statistical significance of the interaction term between internet health literacy and star ratings when predicting perceived credibility. The results failed to reveal a significant interaction between internet health literacy and star ratings, $F(1,356) = .24, p = .64$. A separate custom model was built to assess the statistical significance of the interaction term between internet health literacy and forum post context when predicting perceived credibility. There was a main effect of context, $F(1,356) = 4.76, p = .03, \eta_p^2 = .013$, but no main effect of internet health literacy ($p = .14$). There was also a statistically significant interaction between context and internet

health literacy, $F(1,356) = 5.24, p = .02$, which accounted for 1.5% of the variance in perceived credibility scores. Simple slopes analysis revealed a statistically significant positive linear relationship between internet health literacy and perceived credibility of the advice under the life-threatening condition, ($b = .32, SE = 0.11, p < .01$) but not under the non-life-threatening condition ($b = -.050, SE = 0.116, p = .66$).

Age. A multiple regression was run to assess the interaction term between age and star ratings when predicting perceived credibility of the advice. There was a statistically significant moderator effect of star ratings, which explained an additional 1.3% of the total variance of the model, $F(1,356) = 4.66, p = .03$. However, simple slopes analysis revealed no significant relationships between age and perceived credibility of the advice for the low star condition ($b = 0.01, SE = 0.01, p = .12$) or the high star condition ($b = -0.01, SE = 0.01, p = .13$).

The same analysis was run to examine the interaction between age and star ratings when predicting perceived intellectual ability of the advice writer. There was a statistically significant moderator effect of star ratings, explaining an additional 1.6% of the total variance, $F(1,356) = 5.54, p = .02$. Simple slopes analysis revealed that there was a significantly positive linear relationship ($b = 0.021, SE = 0.01, p = .04$) between age and perceived intellectual ability of the advice writer for the low star condition but not for the high star condition ($b = -0.015, SE = 0.011, p = .17$).

CHAPTER IV

DISCUSSION

The present study sought to examine several variables that may play a role in how individuals make credibility assessments on internet health forums. The effect of high or low star ratings, spelling errors, context of the health issue, and the involvement of the individual evaluating the advice were examined. Overall, there was an effect of star ratings and spelling errors on perceived credibility of the advice, but there was no effect of health context or involvement. In the upcoming sections, each independent variable will be discussed in further detail, outlining the results and potential implications, as well as how the results support the theoretical framework of the Elaboration Likelihood Model.

Star Ratings

The results of the present study revealed a significant main effect of star ratings on perceived credibility of advice, as posts with high star ratings received higher credibility ratings than posts with low star ratings, and the authors of the highly rated posts receiving higher intellectual ability ratings than authors of low-rated posts. This finding provides support for Hypothesis 1, which predicted that the endorsement heuristic would be used in evaluating information. This experimental finding supports qualitative data from multiple studies in which forum users outlined the importance of the overall community opinion of online advice (Lederman, Fan, Smith, & Chang, 2014; Metzger & Flanagin, 2013). This finding also provides support for the use of the anchoring heuristic when evaluating forum advice. Prior work has shown that cues embedded within the structure of a website tend to set the tone for how the overall page is evaluated (Sundar,

2008). In the present study, star ratings were placed at the top left side of the post to make the cue salient to participants. Thus, higher credibility ratings for posts with high star ratings may suggest that the star ratings were viewed initially and set the tone for how the post would be rated.

To further examine the effect of star ratings on perceived credibility, free response data were coded to identify if and when participants mentioned high or low ratings on the post. Although there was an experimental difference of perceived credibility between high and low star ratings, these ratings were only mentioned in seven out of 221 responses (four times for low stars (3.5%) and three times for high stars (2.8%). One potential reason that star ratings were rarely mentioned in the free response data is because participants may not be fully aware of their evaluation techniques. Prior work has shown a discrepancy between explicitly stated information verification techniques and observed information verification techniques, whereby stated techniques are often not executed (Flanagin & Metzger, 2000). It could be the case that the phenomenon is occurring in reverse, where individuals are using specific pieces of information to determine credibility but are not explicitly aware of using the cue as an evaluation tool.

Taken together, the anchoring heuristic and the endorsement heuristic support participants' use of up front information provided by other community members when making a credibility assessment. Prior studies have produced different findings in regards to whether star ratings are reliable in identifying appropriate advice. While one review of an autism forum found that star ratings were rarely indicative of the most medically sound advice (Ben-Sasson et al., 2016), another review of a medical forum found star

ratings to be reliable in identifying good advice (Uden-Kraan et al., 2008). These findings, taken together with the findings from the present study, have important implications for forums in both the decision to implement community ratings as well as the placement of the ratings. The key issue with deciding whether to implement community ratings into a forum is that regardless of how the community rates the advice, there is usually no guarantee that the advice is accurate or is the absolute best advice for the user. Therefore, the present findings suggest that the decision to adopt community ratings should be made separately for each forum, based on the nature of the forum and advice typically given. After making the decision to implement community ratings, administrators should consider the placement of the cue. If forum administrators want users to be aware of the cue when evaluating the advice, placing it up front may be the best option. However, it is also possible that people should have the option to evaluate advice first and then adjust their initial credibility assessment with the opinion of the community. Overall, forum administrators should understand the influence of these cues on evaluation behavior when making the decision to use community ratings and deciding on their best placement.

Spelling Errors

The results of the present study revealed a main effect of spelling errors on both perceived credibility of the advice and perceived intellectual ability of the writer. This finding fully supported Hypothesis 2, which predicted that spelling errors would affect how participants perceived both the post and the poster. Prior work has provided mixed evidence regarding how spelling and grammar affect perceived credibility of information. First, interviews with health forum users revealed that some people find poor spelling and

grammar to be indicative of poor advice, while other people thought advice written with poor spelling and grammar could still be useful (Lederman et al., 2014). However, studies have provided evidence that poor spelling and grammar affect perceived information quality in other online domains (Rowley & Johnson, 2013; Metzger & Flanagin, 2013) and that phonological spelling errors affect perceived intelligence of the writer (Vignovic & Thompson, 2010). In the case of the present study, posts that contained spelling errors (both phonological errors and typographical errors) resulted in significantly lower ratings of perceived credibility, and the writer of the post received significantly lower ratings of intellectual ability. These experimental results were supported by a number of free-response comments regarding spelling. Out of the 110 participants who saw a post with spelling errors, 41.8% ($n = 46$) noted the spelling errors in their evaluation, and a majority of these participants used poor spelling as a supporting factor in their ratings. For example, responders made comments such as, “If you can’t spell, you shouldn’t give out medical advice”, or “You don’t want to take medical advice from someone who cannot even spell medical.”

While results from the present study support spelling as an important factor in determining credibility, they do not support the use of the ELM as a framework to evaluate advice with regards to how spelling errors may interact with other variables (i.e., star ratings). It was hypothesized that high star ratings assigned to a post with multiple spelling errors would activate the expectancy violation heuristic, whereby the spelling errors would disrupt the initial tone set by the high star ratings, leading the participant to take a more thorough processing strategy when evaluating the advice. However, there was no interaction between star ratings and spelling errors for task completion time and

information retention. Therefore, quantitative data did not support Hypothesis 3, which predicted the use of the expectancy violation heuristic within the ELM. However, seven participants made comments that somewhat supported the expectancy violation heuristic. For example, one person commented, “It was difficult for me to look past the spelling errors of the person giving advice, but not being able to spell correctly does not mean that the person may not be able to advise on injuries.” Put more simply, “the spelling was terrible, but the idea of the advice was solid.” These comments are potentially indicative of a more central route of elaboration, whereby participants are noting a cue that may suggest lower credibility, and are evaluating other factors to build a full assessment.

The present results provide support to previous studies in two ways. The results indicate that spelling errors significantly affect the way forum posts are evaluated. However, the overall mean for credibility ratings of post with spelling errors indicated that the advice was ‘somewhat credible’ ($M = 3.00$). Therefore, it may be that spelling errors were taken into account, but in most cases they did not cause the participant to completely discount the advice. Still, none of the participants mentioned using both spelling and star ratings as cues in evaluating credibility, failing to support Hypothesis 3.

Context

It was hypothesized that those reading a life-threatening post would likely take a more central route of processing, shown through longer task completion times and higher information retention scores. This hypothesis was developed from the premise that the importance of the issue is a determining factor of elaboration (Petty & Wegener, 1999). Indeed, present findings support that the life-threatening post was perceived as significantly more important than the non-life-threatening post. However, despite being

instructed that their job would be to evaluate the quality of the advice, those reading the life-threatening post spent significantly *less* time reading the forum post and resulting advice than those reading the non-life-threatening post. Additionally, information retention did not significantly differ between the two groups. This finding is counter to Hypothesis 4, which predicted that the life-threatening scenario would result in a more central route of processing. This hypothesis was developed from the theory that importance may drive the level of elaboration by the individual (Petty & Wegener, 1999). However, prior work suggests that the level of processing may be affected by other factors, such as personal relevance of the topic (Petty & Cacioppo, 1986). In the case of the present study, 53.63% of those in the non-life-threatening condition reported having personal experience with a wrist injury, whereas only 37.16% of those in the life-threatening condition had personal experience. It is possible that having personal experience with the health scenario made the information more relevant, which lead to a more central route of processing.

Another relevant factor that influences how much a person elaborates during information evaluation is the desire to be correct (Petty & Wegener, 1999). The present study indirectly instilled a need for correctness in the involvement manipulation, whereby rewards would be paid out only if the post evaluation was “good.” According to prior work, individuals may believe that the best way to make a correct judgment is to engage in higher information-processing activity, or to take a more central route of elaboration (Petty & Cacioppo, 1979). However, there may be some individuals who, in their desire to be correct, may take a more peripheral route of elaboration if they have little knowledge of the domain (Petty & Wegener, 1999). This alternative view could provide a

partial explanation for why individuals in the life-threatening condition spent significantly less time reading the post and response.

Involvement

Prior work has suggested that the level of involvement an individual has with an evaluation process may determine the level of elaboration taken during the task (Chaiken, 1980). In Chaiken's study, students read text, and high involvement students were told that they would need to discuss the text in the future, whereas low involvement students were told that they would discuss a different topic in the future. The present study added additional instructions to half of the participants. In addition to rating the advice, they were told that they would need to evaluate the advice carefully in order to provide a written evaluation to a hypothetical friend needing advice. Furthermore, they were told that they would receive a one dollar bonus for writing a good evaluation of the advice. However, no differences between high and low involvement participants were found for task completion time or information retention.

To understand why these results may have occurred, it is important to understand incentives through MTurk. The average MTurk worker makes less than five dollars per hour (Hitlin, 2016), so the bonus one dollar for the proposed task was substantial, suggesting a high incentive for doing the task well. Therefore, it was plausible that this manipulation would result in higher involvement. At the end of the study, participants were not actually required to write this evaluation to a friend. Rather, they were debriefed and given the dollar anyway. It is possible that if the participants had written an official evaluation of the advice, it would have included thorough comments indicating close scrutiny of advice. However, results from this study indicate that either the involvement

manipulation did not increase participant involvement sufficiently, the dependent variables in place were not accurate enough measurements of involvement, or that the level of involvement did not affect the level of processing participants engaged in.

Subjective Evaluation Criteria

There were very few comments by participants regarding the star ratings when evaluating the forum post advice. Conversely, nearly 42% of participants who saw a post with poor spelling mentioned that it played a role in their evaluation process. The most common subjective criteria to evaluate advice was whether or not the advice writer recommended – or did not recommend – consulting a doctor. Chi-square tests of association revealed some interesting trends in evaluation behavior depending on the context of the health question. First, results showed that more participants in the life-threatening condition used consulting a doctor as criteria than participants in the non-life-threatening condition. At the same time, results revealed that more participants in the non-life-threatening condition mentioned drawing on their own knowledge of the situation to determine if the advice was good. These two findings together suggest that the severity of the issue plays an important role in evaluation behavior. For a situation that is interpreted as highly important, participants are aware that an internet health forum is not the ideal way to care for the situation, possibly because expertise is needed. However, for a situation that is interpreted as less important and non-life-threatening, people may be more open to the idea of seeking help from peers to solve the problem. Overall, the frequency of suggesting to see a doctor versus the frequency of relying on one's own personal knowledge suggests that individuals may attempt to balance seeking medical attention versus deferring to an internet health forum for medical advice.

Individual Differences

The present study provided the opportunity to examine many individual differences due to the large sample size. However, it should be noted that due to the number of relationships being measured for individual differences, there was a higher chance of committing a Type I error and erroneously rejecting the null hypothesis. The present results yielded few significant relationships of interest to this study. First, overall internet use was not a significant predictor of perceived credibility or intellectual ability. However, there was a small significant relationship between posting to internet forums and perceived credibility as well as intellectual ability. This finding suggests that those who post to internet health forums may have a different overall mental model of internet forums and the unique form of information exchange. Further analyses were run to investigate whether any of the forum cues moderated the relationship. For example, analyses were run to determine if individuals with more posting experience would be more likely to rely on star ratings due to their experience with that type of metric. However, no further relationships were uncovered.

Health literacy data suggests that the s-TOFHLA was not a valid test to administer remotely. The ceiling effect observed for this test was likely because participants had as much time as they needed to read the sentences and match the correct word to the sentence. Therefore, administering this test online may not be advisable. The eHEALS, however, is a scale developed for online administration and is used to measure perceived skills at using information technology for personal health (internet health literacy). Analyses revealed a significant interaction between context and internet health literacy, such that internet health literacy predicted credibility ratings when moderated by

context. Tests of simple slopes revealed that this effect was only seen in those in the life-threatening context. One potential explanation for this finding is that because following advice in response to a life-threatening issue has potentially higher repercussions, it was more important to make a credibility assessment based off of one's own confidence in making health-related decisions online. However, it is possible that the non-life-threatening condition had far less important implications, and therefore the credibility decision was less demanding. However, this effect only accounted for 1.7% of the variance and cannot be used to make any strong claims about internet health literacy.

Further analyses revealed a significant interaction between star ratings and age predicting credibility. However, this effect only accounted for 1.3% of the variance of the model, and tests of simple slopes failed to reveal an individual effect under conditions of low or high star ratings. Similarly, star ratings moderated the relationship between age and intellectual ability of the advice writer, where age was a significant predictor under conditions of low star ratings. This effect only accounted for 1.6% of the variance. As described earlier, high star ratings resulted in significantly higher ratings of perceived credibility. These results suggest that star ratings may be more influential for younger adults when taking into account credibility. It is possible that older adults are more trusting of the information in general.

Overall

Overall, the findings from the present study support the bandwagon effect. The results show that individuals take into account the community opinion when evaluating health information. This finding has practical implications, especially in connection with other previous research. For example, medical professionals reviewing advice on one

forum found that most posts offering health advice were seen as conventional, and none of them were viewed as dangerous (Uden-Kraan et al., 2008). Therefore, community ratings may help to affirm a decision to follow advice by individuals who have questions. However, other evidence suggests that medical professionals do not always agree with the most highly rated advice on a forum (Ben-Sasson, Pelleg, & Yom-Tov, 2016), suggesting that star ratings or other bandwagon mechanisms should not be the only tool to monitor whether advice is credible or not.

The findings from the study also support the claim that internal cues (i.e., spelling) matter when evaluating credibility. Identical advice written without spelling errors received higher credibility ratings, which not only raises the question of how important spelling is to peoples' perception of advice quality, but also of whether this judgment is a good thing or not. Writers with spelling errors were rated as having a significantly lower intellectual ability, an effect that could potentially be erased with something as simple as a spell checker. However, if the claim that spelling errors is a result of being less knowledgeable is true, then perhaps spelling should remain as typed by the advice writer in order to provide more information to the person evaluating the advice. Overall, the lack of an interaction between star ratings and spelling errors results in the question of what other cues may help calibrate trust.

The findings from the present study indicate that the ELM is not able to account for information evaluation behavior on internet health forums. The ELM posits that individuals at the low end of elaboration are more likely to rely on heuristics to make judgments. While a main effect of star ratings was observed, there was no interaction between star ratings and involvement. Furthermore, the ELM posits that factors such as

the level of importance or motivation to evaluate information tends to result in higher elaboration. However, in the present study, regardless of health condition, experience with the particular illness, or opportunity to make more money, participants spent the same amount of time evaluating the advice and reported being equally motivated to participate in the task. While the present study failed to find support for the ELM as a framework for forum post evaluation, limitations of the study will be addressed, and future directions will be recommended.

CHAPTER V

LIMITATIONS AND FUTURE DIRECTIONS

The present study utilized a between-subjects design with forum posts that were carefully crafted to have minimal differences. This was done to ensure that posts with varying subject matter could be directly compared to each other without extraneous information confounding the comparison. However, due to the big difference in the health context, it was nearly impossible to include all of the intricate details that may exist for only a wrist injury or only a head injury. Therefore, these posts may have been less representative of their respective injuries. To this same point, each post had the exact same spelling errors. In order to maintain this consistency, spelling errors were not made on words directly related to the health problem. A few subjective evaluation responses indicated that the spelling errors brought into question how knowledgeable the advice giver was on the subject. It is possible that the impact of spelling errors would be different if the words were directly related to the injury and not the surrounding text.

In the future, studies could concentrate on fewer variables to have more control over content in the forum posts. For example, a study could focus on one health context, such as only a life-threatening post, to compare the other individual forum cues, enabling the post to be written with fewer constraints. In this case, the post could be written to more accurately reflect comments made about the respective injury. Furthermore, the spelling manipulation could be reflected in words specific to the injury. While some comments in the subjective evaluation item reflected potentially deeper processing with regards to the spelling mistakes, it is difficult to interpret those findings post hoc, especially with only forty-six participants mentioning it. It is clear that spelling has an

effect on credibility, but how poor spelling interacts with other variables is still a question that should be investigated further.

One potential limitation of the present study was the effectiveness of the involvement condition. The lack of significant differences in the involvement condition could potentially suggest that the ELM is not an appropriate framework for understanding health forum evaluation behavior. However, it is possible that the involvement manipulation was not strong enough. There were also no differences in the subjective evaluation behavior among individuals in either of the involvement conditions, which suggests that any difference in evaluation criteria resulted from the other manipulations. Future studies aiming to examine the role of the ELM on evaluating health forum credibility should create a stronger manipulation for involvement. At the same time, it is important that there are good measures in place to confirm that the manipulation was successful. In the present study, participants were asked to rate their level of motivation to read and rate the forum question and response, and there was no difference between levels of involvement. However, there was a ceiling effect for the motivation variable and so those data were not helpful as a manipulation check.

The dependent variable of task completion time was used to measure differences between high and low involvement participants, which may not have been an effective measure. In the present study, participants were instructed that they would read and then evaluate advice. Then, they read the post and the response, and clicked to the next page to complete the credibility evaluation. The study was set up in this way so that credibility questions would not lead participants to re-read the advice, potentially biasing their thought processes. However, in doing this, the process of evaluating the information was

potentially low in external validity because participants were unable to return to the advice and spend more time evaluating it. In a typical internet forum, users are able to read the advice as many times as need to make a judgment. While participants were able to spend as much time as desired on the page, the nature of an online survey is to read and then click through to the next page. Therefore, task completion time may not have been the most accurate indicator of elaboration while evaluating the advice.

In the future, studies should have other forms of measurement in place to detect a state of elaboration that is potentially affected by so many variables that heavily rely on personal experience (e.g., relevance or interest). In the present study, participants in the high involvement condition were told that they would have to write an email to a close friend evaluating the advice. However, at the end of the study they were not required to write this email and were instead debriefed and compensated. It is possible that if the participants had written an official evaluation of the advice, some indication of involvement would have been observed. This is an issue that could be addressed in future studies.

In order to further clarify the level of involvement, future studies could test a more realistic and interactive forum environment. A typical forum question is often followed by multiple responses. Furthermore, there is often the option to click on parts of the post, such as the author's icon, and learn more about the author's credentials. In the future, studies could create more realistic forum threads to provide participants more information, and thus a better opportunity to be more involved in the evaluation process.

Another limitation of the present study was the health literacy variable and the way in which it was measured. The s-TOFHLA has not been validated as an online

measure, and the results indicated that either the sample had an overall high literacy level or that the test should not be administered online. Regardless, using this test as a main health literacy measurement limited the analyses and results for this variable. In the future, if researchers want to detect changes in health literacy, it is important to identify the right measurement tools. Currently, there are few options that can be administered online, and so health literacy is a variable that may only be effectively observed through in-person administration.

One limitation that is inherent with data collection on Mechanical Turk is that regardless of the number of elements put in place to control data collection, it is impossible to know exactly what participants were doing during this study. Because of this, some individuals were excluded due to longer task times, and some individuals were excluded due to duplicate IP addresses. While it is possible that participants could have other family members who participated on the same computer, this doesn't rule out the possibility of sharing information or of someone having two separate MTurk IDs to complete high paying HITs. Therefore, some data were eliminated in order to preserve the quality of the data, even though in some circumstances the data could have been provided by legitimate reasons. However, all data were excluded before analyses were run to avoid bias.

Finally, the present study selected multiple variables to manipulate and measure to better understand forum post evaluation behavior. However, these variables were selected from a longer list of variables that are yet to be fully understood within this context. It has been shown that explicit source expertise of the post writer plays a role in evaluation on forums (Kim & Sundar, 2011). Source expertise was not chosen as a main

variable in the present study due to the fact that evidence already supports its role in evaluating credibility, as well as to the fact that source expertise is often unknown in the context of health forums. However, ELM literature suggests that source expertise is an attribute that impacts how messages are evaluated (Petty & Cacioppo, 1986), and therefore it is a variable that should be addressed in future research.

CHAPTER VI

CONCLUSIONS

The present study examined the role of star ratings, spelling errors, context of health problem, and participant involvement as factors that may influence the evaluation of internet health forum advice. The results of this study supported the endorsement heuristic and the anchoring heuristic, as well as prior work that suggests spelling plays a role in how credibility is perceived. The present study did not provide evidence that the ELM is an appropriate framework to understand perceived credibility of forum advice, but provided ways in which this framework can be studied more effectively. Overall, advice on internet health forums is evaluated in a complicated manner, and the present study supports the need to further investigate evaluation behavior.

APPENDICES

APPENDIX A

Experimental Text from Forum Posts and Responses

Life threatening Question: Two days ago I was doing some yard work and I had a big accident. I was about to water the flowers and I reached for the hose and when I did a snake popped out and slithered right over my hand! I kind of jumped backwards and when I did I tripped over the carport ledge and fell. Lucky for me the snake didn't bite me so that's not what this is about. What happened was I landed on my head. It hurt a lot but I didn't really hear or feel any cracks so I figured my skull wasn't broken. I did fall on the concrete so it was a very hard landing.

Since then, my head has been pounding with pain and it's the worst headache I've ever felt. I can tell my vision is a lot fuzzier but I seem to be able to still see more or less fine. It looks swollen really bad where I hit it and hurts a lot when I try to concentrate, like when I'm reading. I feel really terrible overall. I have trouble doing other stuff like trying to remember things.

I was icing it a couple of times a day but the swelling is still fully there. I am kind of worried that maybe I caused some kind of damage inside my brain, but I don't want to jump to conclusions even though it was a very bad fall. Anyone know anything about head injuries and what I should do?

Response with correct spelling: Well at least you didn't get bit by the snake that could have been really bad. You're right to worry, what you said sounds a lot like a bad head injury. In my opinion you do really need to go to the doctor.

First of all such a big amount of swelling probably means that you hit your head really hard. Icing it was a good idea, that is usually one of the first things to do. When it happened did you bleed at all? You can always take a couple of aspirins if the pain doesn't go away. If you don't go to the doctor straight away, and I do think you really need to, the biggest thing you can do on your own is take it really easy. What I mean by that is don't do anything that needs a lot of brain power. If it feels bad to read try not to read much for a little bit..

To be honest though your symptoms do sound really bad and you definitely need to go to the doctor. Your vision being a lot fuzzier doesn't seem normal for even a small fall. When you go to the doctor they will check you out with a physical and maybe do an X-ray and see if there is any real damage inside your brain.

Response with spelling errors: Well at least you didn't get bit by the snake that could have been really bad. **Youre** right to worry, what you said sounds a lot like a bad head injury. In my **opniion** you do really need to go to the doctor.

First of all such a big amount of swelling probably means that you hit your head **reely** hard. **Iceing** it was a good idea, that is usually one of the first things to do. When it happened did you bleed at all? You can always take a **cuople** of aspirins if the pain doesn't go away. If you don't go to the doctor strait away, and I do think you really need to, the biggest thing you can do on your own is take it really easy. What I mean by that is don't do **any thing** that needs a lot of brain power. If it feels bad to read try not to read

much for a little bit.

To be honest though your **syptoms** do sound really bad and you definitely need to go to the doctor. Your vision being a lot fuzzier doesn't seem normal for even a small fall. When you **goto** the doctor they will check you out with a **pyhsical** and maybe do an X-ray and see if there is any real damage inside your brain.

Non-Life-Threatening Question: Two days ago I was doing some yard work and I had a little accident. I was about to water the flowers and I reached for the hose and when I did a snake popped out and slithered right over my hand! I kind of jumped backwards and when I did I tripped over the carport ledge and fell. Lucky for me the snake didn't bite me so that's not what this is about. What happened was I landed on my wrist. It hurt a little but I didn't really hear or feel any cracks so I figured my wrist wasn't sprained. I did fall in the grass so it was a very soft landing.

Since then, my wrist has hurt a tiny bit but it's not the worst bruise I've ever felt. I can tell my joint is a little stiffer but I seem to be able to move it more or less fine. It doesn't look swollen where I hit it and it only seems to hurt when I'm using it a lot, like when I'm writing. I feel mostly OK overall. I don't have trouble doing other stuff like trying to hold things.

I was icing it a couple of times a day and the swelling is now fully gone. I am kind of worried that maybe I caused some kind of damage inside my wrist, but I don't want to jump to any conclusions because it was a pretty light fall. Anyone know anything about wrist injuries and what I should do?

Response with correct spelling: Well at least you didn't get bit by the snake that could have been really bad. You're right to not worry, what you said sounds a lot like a little wrist sprain. In my opinion you don't really need to go to the doctor.

First of all such a small bruise probably means that you didn't hit your wrist really hard. Icing it was a good idea, that is usually one of the first things to do. When it happened did you bleed at all? You can always take a couple of aspirins if the pain doesn't go away. If you don't go to the doctor straight away, and I don't think you really need to, the biggest thing you can do on your own is take it really easy. What I mean by that is don't do anything that needs a lot of wrist movement. If it feels bad to write try not to write much for a little bit.

To be honest though your symptoms don't sound that bad and you probably don't need to go to the doctor. Your joint being a little stiffer seems normal for even a small fall. If you go to the doctor they would check you out with a physical and maybe do an X-ray and see if there is any real damage inside your wrist.

Response with spelling errors: Well at least you didn't get bit by the snake that could have been really bad. **Youre** right to not worry, what you said sounds a lot like a little wrist sprain. In my **opniion** you don't really need to go to the doctor.

First of all such a small bruise probably means that you didn't hit your wrist **reely** hard. **Iceing** it was a good idea, that is usually one of the first things to do. When it happened did you bleed at all? You can always take a **cuople** of aspirins if the pain doesn't go away. If you don't go to the doctor strait away, and I don't think you really need to, the biggest thing you can do on your own is take it really easy. What I mean by

that is don't do **any thing** that needs a lot of wrist movement. If it feels bad to write try not to write much for a little bit.

To be honest though your **syptoms** don't sound that bad and you probably don't need to go to the doctor. Your joint being a little stiffer seems normal for even a small fall. If you **goto** the doctor they would check you out with a **pyhsical** and maybe do an X-ray and see if there is any real damage inside your wrist.

APPENDIX B

Credibility Assessment Scales

Credibility Scale (Flanagin & Metzger, 2000)

(1-5 likert; not at all - extremely)

Rate the degree that you find information in the forum advice to be:

- Believable
- Accurate
- Trustworthy
- Biased (reverse score)
- Complete

Expert Question

(1-5 likert; not at all – extremely)

Rate the degree that you find the author of the advice to be an Expert

APPENDIX C

Intellectual Scale

Intellectual Scale (Warner & Sugarman, 1986)

Five seven-point semantic differential items:

- Incompetent/Competent
- Ignorant/ Knowledgeable
- Irresponsible/Responsible
- Unintelligent/ Intelligent
- Foolish/Sensible

APPENDIX D

Expertise Items

Expertise Items

1. Do you have any experience dealing with a wrist/head [depending on condition] injury - either with you or someone you know? (Select all that apply)
 - No
 - Yes, I have had a wrist/head [depending on condition] injury myself
 - Yes, someone close to me has had a wrist/head [depending on condition] injury
2. Do you have any formal education or training in health or medicine? (select highest level that applies)
 - Certificate in health or medicine
 - If yes please specify
 - Associate's degree in health field (nursing, radiology, physical therapy assistant, etc.)
 - If yes, please specify
 - Bachelor's degree in health field
 - If yes, please specify
 - Advanced degree in health field (MD, DNP, PA, PT, PhD in health field, etc.)
 - If yes, please specify
3. Other than formal training, have you ever had any health or medical training?
 - No
 - Yes
 - Please specify

APPENDIX E

Instructional Manipulation Check

Instructions: “Most of the typical posts involving health information ask questions that many different people have probably experienced. Individuals from different backgrounds with different skills usually have different opinions about how good online information is. In order to facilitate our research on health forums we are interested in knowing certain factors about you, the consumer. Specifically, we are interested in whether you actually take the time to read the directions; if not, then some of our manipulations that rely on changes in the instructions will be ineffective. So, in order to demonstrate that you have read the instructions, please ignore the health items below. Instead, simply click on the continue button to proceed to the next screen. You are not required to click anything else on this page. Thank you very much.”

Which of these health activities have you had experience with?

(Click on all that apply)

Diabetes, Lung disease, Allergies, High blood pressure, High cholesterol,
Skin rash, Coughing, Broken bones, Sore throat, Heart disease, None of these

APPENDIX F

E Health Literacy Scale (e-HEALS)

1. I know what health resources are available on the Internet
2. I know where to find helpful health resources on the Internet
3. I know how to find helpful health resources on the Internet
4. I know how to use the Internet to answer my health questions
5. I know how to use the health information I find on the Internet to help me
6. I have the skills I need to evaluate the health resources I find on the Internet
7. I can tell high quality from low quality health resources on the Internet
8. I feel confident in using information from the Internet to make health decisions

Individuals rate items on a 5-point Likert scale (1-strongly disagree, 5-strongly agree)

APPENDIX G

Short Test of Functional Health Literacy (S-TOFHLA)

Your doctor has sent you to have a _____ X-ray

- stomach
- diabetes
- stitches
- germs

You must have an _____ stomach when you come for _____.

- asthma
- empty
- incest
- anemia
- is
- am
- if
- it

The X-ray will _____ from 1 to 3 _____ to do.

- take
- view
- talk
- look
- beds
- brains
- hours
- diets

The Day before the X-ray:

For Supper have only a _____snack of fruit, _____and jelly, with coffee or tea

- little
- broth
- attack
- nausea
- toes
- throat
- toast
- thigh

The Day before the X-ray:

After_____, you must not_____or drink

- minute
- midnight
- during
- before
- easy
- ate
- drank
- eat

The Day before the X-ray:

anything at _____ until after you have_____ the X-ray.

- ill
- all
- each
- any
- are
- has
- had
- was

The day of the X-Ray:

Do not eat_____.

- appointment
- walk-in
- breakfast
- clinic

The day of the X-Ray:

Do not _____, even_____.

- drive
- drink
- dress
- dose
- heart
- breath
- water
- cancer

The day of the X-Ray: If you have any _____, call the X-ray_____ at 616-4500.

- answers
- exercises
- tracts
- questions
- Department
- Sprain
- Pharmacy
- Toothache

I agree to give correct information to _____if I can receive Medicaid.

- hair
- salt
- see
- ache

I _____to provide the country information to _____any

- agree
- probe
- send
- gain
- hide
- risk
- discharge
- prove

...statements given in this _____ and hereby give permission to

- emphysema
- application
- gallblader
- relationship

the_____to get such proof. I _____that for

- inflammation
- religion
- iron
- country
- investigate
- entertain
- understand
- establish

Medicaid I must report any _____ in my circumstances

- changes
- hormones
- antacids
- charges

within _____(10) days of becoming _____of the change.

- three
- one
- five
- ten
- award
- aware
- away
- await

I understand _____ if I DO NOT like the _____ made on my

- thus
- this
- that
- than
- marital
- occupation
- adult
- decision

I have the _____ to a fair hearing. I can _____ a

- bright
- left
- wrong
- right
- request
- refuse
- fail
- mend

hearing by writing or _____ the country where I applied.

- counting
- reading
- calling
- smelling

If you _____ TANF for any family _____, you will have to

- wash
- want
- cover
- tape
- member
- history
- weight
- seatbelt

_____ a different application form. _____, we will use

- relax
- break
- inhale
- sign
- Since
- Whether
- However
- Because

the _____ on this form to determine your _____.

- lung
- date
- meal
- pelvic
- hypoglycemia
- eligibility
- osteoporosis
- chizophrenia

APPENDIX H

Short Form 36 Item Health Questionnaire (SF-36)

In general, would you say your health is:

- Excellent
- Very good
- Good
- Fair
- Poor

Compared to one year ago, how would you rate your health in general now?

- Much better now than one year ago
- Somewhat better now than one year ago
- About the same
- Somewhat worse now than one year ago
- Much worse now than one year ago

The following items are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much? [Yes, limited a lot; Yes, limited a little; No, not limited at all]

- Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports
- Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf
- Lifting or carrying groceries
- Climbing several flights of stairs
- Climbing one flight of stairs
- Bending, kneeling, or stooping
- Walking more than a mile
- Walking several blocks
- Walking one block
- Bathing or dressing yourself

During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health? (Yes or No)

- Cut down the amount of time you spent on work or other activities
- Accomplished less than you would like
- Were limited in the kind of work or other activities
- Had difficulty performing the work or other activities (for example, it took extra effort)

During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)? (Yes or no)

- Cut down the amount of time you spent on work or other activities
- Accomplished less than you would like
- Didn't do work or other activities as carefully as usual

During the past 4 weeks, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbors, or groups?

- Not at all
- Slightly
- Moderately
- Quite a bit
- Extremely

How much bodily pain have you had during the past 4 weeks?

- None
- Very mild
- Mild
- Moderate
- Severe
- Very severe

During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?

- Not at all
- A little bit
- Moderately
- Quite a bit
- Extremely

These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the past 4 weeks [All of the time; Most of the time; A good bit of the time; Some of the time; A little of the time; None of the time]

- Did you feel full of pep?
- Have you been a very nervous person?
- Have you felt so down in the dumps that nothing could cheer you up?
- Have you felt calm and peaceful?
- Did you have a lot of energy?
- Have you felt downhearted and blue?
- Did you feel worn out?
- Have you been a happy person?
- Did you feel tired?

How much of the time during the past 4 weeks... [All of the time; Most of the time; A good bit of the time; Some of the time; A little of the time; None of the time]

- Did you feel full of pep?
- Have you been a very nervous person?
- Have you felt so down in the dumps that nothing could cheer you up?
- Have you felt calm and peaceful?
- Did you have a lot of energy?
- Have you felt downhearted and blue?
- Did you feel worn out?
- Have you been a happy person?
- Did you feel tired?

During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

- All of the time
- Most of the time
- Some of the time
- A little of the time
- None of the time

How TRUE or FALSE is each of the following statements for you. [Definitely True; Somewhat True; Don't know; Mostly False; Definitely False]

- I seem to get sick a little easier than other people
- I am as healthy as anybody I know
- I expect my health to get worse
- My health is excellent

APPENDIX I

Internet Usage Survey

Please indicate **how frequently** you do each of the following:

- 1- Strongly Disagree, 2 – Disagree, 3 – Neither Agree nor Disagree, 4 – Agree, 5 – Strongly Agree

Please respond to each statement below:

1. I get my news (weather, traffic) from the web
2. I get learning information (courses, training) on the web.
3. I get information about leisure events on the web.
4. I get travel information on the web.
5. I get health information from the web.
6. I often use cloud storage (Dropbox, iCloud, Box).
7. I often shop on the web.
8. I often use online video chat (Skype, FaceTime).
9. I often use online banking.
10. I often post to social media.

----- Begin additional items here -----

- a. I post to... check all that apply
 - i. Facebook
 - ii. Twitter
 - iii. Instagram
 - iv. Snapchat
11. I often read social media.
 - a. I read... check all that apply.
 - i. Facebook
 - ii. Twitter
 - iii. Instagram
 - iv. Snapchat
 - v. Other
12. I **read** online forums (also known as message boards or bulletin boards; e.g. Yahoo! Answers, Reddit).
 - i. Yahoo! Answers
 - ii. Reddit
 - iii. WebMD forums
 - iv. Other
13. I **post** to online forums (also known as message boards or bulletin boards; e.g. Yahoo! Answers, Reddit).
 - i. Yahoo! Answers
 - ii. Reddit

- iii. WebMD forums
- iv. Other

14. I work on an online work site (e.g. Mechanical Turk)

“Now we’re going to ask you about getting health information online...”

The Following Scales apply to instrumentation

How often: Never, 1 - 2 times a year, 4 -6 times a year, once a month, 2 - 3 times a month, once a week, 2 - 3 time a week, daily

15. In the past year, when I go online to get health information, I use:

- A search engine, such as Google, Bing, or Yahoo
- A website that specializes in health information, like WebMD
- A more general site like Wikipedia, that contains information on all kinds of topics
- A video site like Youtube
- A social media site like Facebook, Twitter, or Instagram
 - Facebook
 - Twitter
 - Other
- A message board or discussion site like Yahoo! Answers or Reddit
 - Yahoo! Answers
 - Reddit
 - Other specific forum just for that health concern or illness
 - Other
- Other online location for health information (please specify)

16. In the past year, when I get health information from **other people online** I get it from...

- People online who I **do** know in real life (like asking a health question and getting answers from friends or family)
- People online who I **don’t** know in real life (like asking a question and getting answers from anyone who reads it)
- Other [please specify]

17. In the past year I have gone online to:

- a. Read/watch someone else’s health or medical experience online
- b. Find others who might have similar health concerns
- c. Post comments or stories about personal health experiences
- d. Post specific health questions

APPENDIX J

Retention Check Quiz

Retention Check Quiz: Free Recall

Take a second to think about the forum question and response you read earlier in this study. Now, in your own words, write as much as you can remember – in as much detail as possible - about the forum question and response.

Retention Check Quiz: Multiple Choice

1. How many days ago did the person's health problem start?
 - 3 - 5 days ago
 - 6 - 8 days ago
 - 9 - 13 days ago
 - **1 - 2 days ago**
 - Two or more weeks ago

2. What was the person doing when the fall occurred?
 - **Working in the yard**
 - Folding the laundry
 - Washing the dishes
 - Jogging
 - Washing the car

3. What did the person fall on?
 - Pile of leaves
 - Hardwood floor
 - **Grass/Concrete (context dependent)**
 - Stairs
 - Carpeted floor

4. What did the person have trouble doing?
 - Remembering events
 - Taking a shower
 - Making a phone call
 - Hearing the television
 - **Reading/writing (context dependent)**

Fill in the blank:

1. The person was scared by a _____. (snake)
 2. The person was using _____ on the injury. (ice)
9. Wrist: Which of the following were symptoms after the injury? Select all that apply:
- **Pain**
 - **Bruise**
 - Stomach pain
 - **Stiff joint**
 - Bleeding
 - Trouble walking
9. Head: Which of the following were symptoms after the injury? Select all that apply:
- **Headache**
 - **Swelling**
 - Trouble walking
 - **Fuzzy vision**
 - Bleeding
 - Stomach pain

Response Retention Items

10. Wrist: What type of injury did the person think it was?
- Fractured wrist
 - **Wrist sprain**
 - Bruised arm
 - Pulled muscle
 - Deep cut
10. Head: What type of injury did the advice giver think it was?
- Fractured skull
 - **Head injury**
 - Bruised neck
 - Pulled muscle
 - Deep cut
11. If the person goes to the doctor, what does the advice giver think will happen?
(physical and maybe an x-ray)
- The doctor will prescribe antibiotics

- The doctor will prescribe pain medication
- The doctor will keep him/her overnight
- The doctor will put him/her in a cast
- **The doctor will take an x-ray**

12. What type of medicine did the advice giver recommend using?

- Ibuprofen
- Tylenol
- **Aspirin**
- Peroxide
- Advil

13. According to the advice giver, if the person doesn't go to the doctor, what is the best thing they can do on their own?

- Ask a friend to take a look at the injury
- Make a homemade cast
- Take it easy
- Continue to search for information online
- Take a sleep aid

APPENDIX K

Pilot Testing Results

Importance of Question

- The question is important
 - The individual who posted this question is right to worry
 - The individual who posted this question may be dealing with a life threatening issue
 - This person should seek medical attention immediately
- 1- Strongly Disagree
 - 2- Disagree
 - 3- Somewhat Disagree
 - 4- Neither Disagree nor Agree
 - 5- Somewhat Agree
 - 6- Agree
 - 7- Strongly Agree

Table 5.

Pilot results for importance of question.

Question	Mean		p
	Life-Threatening (n=21)	Non-life-threatening (n=19)	
Important	6.76	4.58	< .01
Right to Worry	6.62	3.11	< .01
Issue is life threatening	6.57	1.89	< .01
Immediate medical attention	6.81	3.26	<.01

Spelling Errors

- The advice given was high quality
 - The response contained a lot of spelling/grammar errors
 - Mistakes in the writing took away from the overall quality of the advice
 - The person who wrote the response seemed intelligent
 - The person who wrote the response seemed careless
- 1- Strongly Disagree
 - 2- Disagree
 - 3- Somewhat Disagree
 - 4- Neither Disagree nor Agree
 - 5- Somewhat Agree

- 6- Agree
- 7- Completely Agree

Table 6.

Pilot results for forum responses.

Question	Mean		p
	Correct Spelling (n=19)	Spelling Errors (n=21)	
The advice given in this post is good advice	5.58	5.43	.745
The advice seems like it is written by a typical American adult	5.32	4.90	.404
The person who wrote the advice is careless	2.42	2.24	.700
The person who wrote the advice is intelligent	5.21	4.67	.286
The advice has a lot of spelling errors	1.89	4.57	< .01
Mistakes in the spelling took away from the overall quality of the advice	3.95	2.32	< .01

REFERENCES

- Atkinson, N., Saperstein, S., & Pleis, J. (2009). Using the internet for health-related activities: Findings from a national probability sample. *Journal of Medical Internet Research*, 11(1), e5.
- Baker, D. W., Williams, M. V., Parker, R. M., Gazmararian, J. A., & Nurss, J. (1999). Development of a brief test to measure functional health literacy. *Patient Education and Counseling*, 38(1), 33-42.
- Ben-Sasson, A., Pelleg, D., & Yom-Tov, E. (2016). The quality of online answers to parents who suspect that their child has an autism spectrum disorder. Paper presented at the ICWSM, 543-546.
- Berlo, D. K., Lemert, J. B., & Mertz, R. J. (1969). Dimensions for evaluating the acceptability of message sources. *Public opinion quarterly*, 33(4), 563-576.
- Blank, T. O., & Adams-Blodnieks, M. (2007). The who and the what of usage of two cancer online communities. *Computers in Human Behavior*, 23(3), 1249-1257.
- Brady, E., Segar, J., & Sanders, C. (2016). "You get to know the people and whether they're talking sense or not": Negotiating trust on health-related forums. *Social Science & Medicine*, 162, 151-157.
- Briggs, P., Burford, B., De Angeli, A., & Lynch, P. (2002). Trust in online advice. *Social Science Computer Review*, 20(3), 321-332.
- Burgoon, J. K., Stern L. A., & Dillman, L. (1995). Interpersonal adaptation. Cambridge, U.K.: Cambridge University Press.

- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752.
- Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of personality and social psychology*, 66, 460-460.
- Cheung, J. H., Burns, D. K., Sinclair, R. R., & Sliter, M. (2016). Amazon mechanical turk in organizational psychology: An evaluation and practical recommendations. *Journal of Business and Psychology*, 1-15.
- Chung, S. Y., & Nahm, E. S. (2015). Testing reliability and validity of the eHealth literacy scale (eHEALS) for older adults recruited online. *Computers, Informatics, Nursing*, 33(4). 150-156.
- Corritore, C. L., Kracher, B., & Wiedenbeck, S. (2003). On-line trust: concepts, evolving themes, a model. *International journal of human-computer studies*, 58(6), 737-758.
- Culver, J. D., Gerr, F., & Frumkin, H. (1997). Medical information on the internet. *Journal of General Internal Medicine*, 12(8), 466-470.
- Dixon, R. A., Hultsch, D. F., Simon, E. W., & von Eye, A. (1984). Verbal ability and text structure effects on adult age differences in text recall. *Journal of Verbal Learning and Verbal Behavior*, 23(5), 569-578.

- Dutton, W. H., & Shepherd, A. (2006). Trust in the Internet as an experience technology. *Information, Communication & Society*, 9(4), 433-451.
- Elaad, E. (2003). Effects of feedback on the overestimated capacity to detect lies and the underestimated ability to tell lies. *Applied Cognitive Psychology*, 17(3), 349-363.
- Eysenbach, G., & Kohler, C. (2002). How do consumers search for and appraise health information on the world wide web? qualitative study using focus groups, usability tests, and in-depth interviews. *BMJ (Clinical Research Ed.)*, 324(7337), 573-577.
- Faith, J., Thorburn, S., & Sinky, T. H. (2016). Exploring healthcare experiences among online interactive weight loss forum users. *Computers in Human Behavior*, 57, 326-333.
- Flanagin, A. J., Hocevar, K. P., & Samahito, S. N. (2014). Connecting with the user-generated web: How group identification impacts online information sharing and evaluation. *Information, Communication & Society*, 17(6), 683-694.
- Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of internet information credibility. *Journalism & Mass Communication Quarterly*, 77(3), 515-540.
- Fogg, B. J., & Tseng, H. (1999). The elements of computer credibility. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 80-87).
- Fogg, B. J., Soohoo, C., Danielson, D. R., Marable, L., Stanford, J., & Tauber, E. R. (2003). How do users evaluate the credibility of Web sites? A study with over 2,500 participants. In *Proceedings of the 2003 conference on Designing for user experiences* (pp. 1-15).

- Fox, S. (2011). The social life of health information, 2011. Washington, DC: Pew Internet & American Life Project.
- Freeman, K. S., & Spyridakis, J. H. (2004). An examination of factors that affect the credibility of online health information. *Technical Communication*, 51(2), 239-263.
- Giffin, K. 1967. The contribution of studies of source credibility to a theory of interpersonal trust in the communication process. *Psychological Bulletin*, 68, 104-20.
- Ginossar, T. (2008). Online participation: A content analysis of differences in utilization of two online cancer communities by men and women, patients and family members. *Health Communication*, 23(1), 1-12.
- Golbeck, J., & Fleischmann, K. R. (2010). Trust in social Q&A: The impact of text and photo cues of expertise. *Proceedings of the American Society for Information Science and Technology*, 47(1), 1-10.
- Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of mechanical turk samples. *Journal of Behavioral Decision Making*, 26(3), 213-224.
- Harper, F. M., Raban, D., Rafaeli, S., & Konstan, J. A. (2008). Predictors of answer quality in online Q&A sites. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 865-874.
- Hether, H. J., Murphy, S. T., & Valente, T. W. (2014). It's better to give than to receive: The role of social support, trust, and participation on health-related social networking sites. *Journal of Health Communication*, 19(12), 1424-1439.

- Hitlin, Paul. (2016). 'Research in the Crowdsourcing Age, a Case Study' Pew Research Center. Available at: <http://www.pewinternet.org/2016/07/11/research-in-the-crowdsourcing-age-a-case-study/>.
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). Communication and persuasion; psychological studies of opinion change.
- Huh, J., Yetisgen-Yildiz, M., & Pratt, W. (2013). Text classification for assisting moderators in online health communities. *Journal of Biomedical Informatics*, 46(6), 998-1005.
- Ipeirotis, P. G. (2010). Demographics of mechanical turk.
- Kim, B., & Han, I. (2009). The role of trust belief and its antecedents in a community-driven knowledge environment. *Journal of the American Society for Information Science and Technology*, 60(5), 1012-1026.
- Kim, H., & Sundar, S. S. (2011). Using interface cues in online health community boards to change impressions and encourage user contribution. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 599-608.
- Knapp, C., Madden, V., Wang, H., Sloyer, P., & Shenkman, E. (2011). Internet use and eHealth literacy of low-income parents whose children have special health care needs. *Journal of Medical Internet Research*, 13(3), e75.
- Kreiner, D. S., Schnakenberg, S. D., Green, A. G., Costello, M. J., & McClin, A. F. (2002). Effects of spelling errors on the perception of writers. *The Journal of general psychology*, 129(1), 5-17.

- Lederman, R., Fan, H., Smith, S., & Chang, S. (2014). Who can you trust? Credibility assessment in online health forums. *Health Policy and Technology*, 3(1), 13-25.
- Lee, J., Park, D. H., & Han, I. (2008). The effect of negative online consumer reviews on product attitude: An information processing view. *Electronic commerce research and applications*, 7(3), 341-352.
- Liao, Q. V., & Fu, W. (2014). Expert voices in echo chambers: Effects of source expertise indicators on exposure to diverse opinions. Paper presented at the Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems, 2745-2754.
- Magnezi, R., Grosberg, D., Novikov, I., Ziv, A., Shani, M., & Freedman, L. S. (2015). Characteristics of patients seeking health information online via social health networks versus general internet sites: A comparative study. *Informatics for Health and Social Care*, 40(2), 125-138.
- Marco Leimeister, J., Schweizer, K., Leimeister, S., & Krcmar, H. (2008). Do virtual communities matter for the social support of patients? antecedents and effects of virtual relationships in online communities. *Information Technology & People*, 21(4), 350-374.
- Matthews, P. (2015). Doctoral Dissertation.
- Metzger, M. J. (2007). Making sense of credibility on the web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, 58(13), 2078-2091.

- 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.
- Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of communication*, 60(3), 413-439.
- Meyerson, D., Weick, K. E., & Kramer, R. M. (1996). Swift trust and temporary groups. *Trust in organizations: Frontiers of theory and research*, 166, 195.
- Nahm, E. S., Resnick, B., DeGrazia, M., & Brotemarkle, R. (2009). Use of discussion boards in a theory-based health web site for older adults. *Nursing Research*, 58(6), 419-426.
- Norman, C. D., & Skinner, H. A. (2006). eHealth literacy: Essential skills for consumer health in a networked world. *Journal of Medical Internet Research*, 8(2), e9.
- 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.
- Pak, R., Rovira, E., McLaughlin, A. C., & Leidheiser, W. (2017). Evaluating attitudes and experience with emerging technology in cadets and civilian undergraduates. *Military Psychology*, 29(5), 448.
- Payne, J. W., & Bettman, J. R. (2004). Walking with the scarecrow: The information-processing approach to decision research. *Blackwell handbook of judgment and decision making*, 110-132.

- Petty, R. E., & Cacioppo, J. T. (1979). Effects of forwarning of persuasive intent and involvement on cognitive responses and persuasion. *Personality and Social Psychology Bulletin*, 5(2), 173-176.
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in experimental social psychology*, 19, 123-205.
- Petty, R. E., & Wegener, D. T. (1999). The Elaboration Likelihood Model; Current status and controversies. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 41-72).
- Rowley, J., & Johnson, F. (2013). Understanding trust formation in digital information sources: The case of Wikipedia. *Journal of Information Science*, 39(4), 494-508.
- Savolainen, R. (2011). Judging the quality and credibility of information in internet discussion forums. *Journal of the American Society for Information Science and Technology*, 62(7), 1243-1256.
- Schillinger, D., Barton, L. R., Karter, A. J., Wang, F., & Adler, N. (2006). Does literacy mediate the relationship between education and health outcomes? A study of a low-income population with diabetes. *Public Health Reports*, 121(3), 245-254.
- Sillence, E., Briggs, P., Harris, P. R., & Fishwick, L. (2007). How do patients evaluate and make use of online health information? *Social Science & Medicine*, 64(9), 1853-1862.
- Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, 69(1), 99-118.

- Smith, N. A., Sabat, I. E., Martinez, L. R., Weaver, K., & Xu, S. (2015). A convenient solution: Using MTurk to sample from hard-to-reach populations. *Industrial and Organizational Psychology*, 8(02), 220-228.
- Soroka, V., & Rafaeli, S. (2006). Invisible participants: How cultural capital relates to lurking behavior. Paper presented at the Proceedings of the 15th International Conference on World Wide Web, 163-172.
- Stanford, J., Tauber, E. R., & Marable, L. (2002). Experts vs. online consumers: A comparative credibility study of health and finance Web sites (Consumer WebWatch Research Report). *Journal of Communication*. 42, 73-93.
- Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. *Digital media, youth, and credibility*, 73100.
- Tabachnick, B. G., & Fidell, L. S. (2001). Using Multivariate Statistics (4th ed.). Boston, MA: Allyn and Bacon.
- Thackeray, R., Crookston, B. T., & West, J. H. (2013). Correlates of health-related social media use among adults. *Journal of Medical Internet Research*, 15(1), e21.
- Turner, A., & Greene, E. (1978). The construction and use of a propositional text base. *JSAS Catalog of Selected Documents in Psychology*, 8 (58).
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological review*, 79(4), 281.
- Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making* (pp. 141-162). Springer Netherlands.

- van Uden-Kraan, C. F., Drossaert, C. H., Taal, E., Shaw, B. R., Seydel, E. R., & van de Laar, M. A. (2008). Empowering processes and outcomes of participation in online support groups for patients with breast cancer, arthritis, or fibromyalgia. *Qualitative Health Research*, 18(3), 405-417.
- Vernon, J. A., Trujillo, A., Rosenbaum, S., & DeBuono, B. (2007). Low health literacy: Implications for national health policy. Washington, DC: Department of Health Policy, School of Public Health and Health Services, The George Washington University.
- Vignovic, J. A., & Thompson, L. F. (2010). Computer-mediated cross-cultural collaboration: Attributing communication errors to the person versus the situation. *Journal of Applied Psychology*, 95(2), 265.
- Wang, Z., Walther, J. B., Pingree, S., & Hawkins, R. P. (2008). Health information, credibility, homophily, and influence via the internet: Web sites versus discussion groups. *Health Communication*, 23(4), 358-368.
- Ware Jr, J. E., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical care*, 473-483.
- Warner, R. M., & Sugarman, D. B. (1986). Attributions of personality based on physical appearance, speech, and handwriting. *Journal of Personality and Social Psychology*, 50(4), 792.
- White, M., & Dorman, S. M. (2001). Receiving social support online: Implications for health education. *Health Education Research*, 16(6), 693-707.

- Williams, E. J., Beardmore, A., & Joinson, A. N. (2017). Individual differences in susceptibility to online influence: A theoretical review. *Computers in Human Behavior*, 72, 412-421.
- Wright, K. (2002). Social support within an on-line cancer community: An assessment of emotional support, perceptions of advantages and disadvantages, and motives for using the community from a communication perspective. *Journal of Applied Communication Research*, 30(3), 195-209.
- Yang, S. C., Hung, W. C., Sung, K., & Farn, C. K. (2006). Investigating initial trust toward e-tailers from the elaboration likelihood model perspective. *Psychology & Marketing*, 23(5), 429-445.
- Zarcadoolas, C., Blanco, M., Boyer, J. F., & Pleasant, A. (2002). Unweaving the web: An exploratory study of low-literate adults' navigation skills on the world wide web. *Journal of Health Communication*, 7(4), 309-324.
- Zickuhr, K., & Madden, M. (2012). Older adults and internet use: For the first time, half of adults ages 65 and older are online. Pew Research Center's Internet & American Life Project Report (2012). Pewinternet.org/Reports/2012/Older-Adults-and-Internet-use.