

AGE DIFFERENCES IN MAKING CREDIBILITY JUDGMENTS OF ONLINE
HEALTH INFORMATION

BY

QINGZI LIAO

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Human Factors
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Advisor:

Assistant Professor Wai-Tat Fu

ABSTRACT

Older adults constitute a notable group among the exponentially growing population of online health information consumers. To better support their consumption of high quality health information on the Internet, it is important to understand how older adults make credibility judgments of online health information. For this purpose, I conducted two laboratory studies to explore how the credibility cues in message contents, website features and user reviews, which are widely used for online medication information, could differentially impact younger (19 to 26 years of age) and older adults' (58 to 80 years of age) credibility judgments. Results from the first experiment showed that older adults, compared to younger ones, were less sensitive to the credibility cues in message contents (as to differentiate between credible ones and non-credible ones), as well as those in the website features. In the second experiment, I tested whether the addition of user reviews could moderate the age differences in credibility judgments. Results showed that user reviews that were consistent with the credibility cues in message contents could more significantly reinforce older adults' credibility judgments than that of younger adults'. However, when user reviews were inconsistent with the credibility of message contents, older adults seemed to be less swayed than younger adults. In addition, I found that decline in cognitive abilities and lack of Internet experience were two important factors that limit older adults' ability to correctly judge the credibility of online medication information, while individual's better health domain knowledge could possibly compensate for older adults' ability of making correct credibility judgment. Their inexperience with the Internet, and perhaps especially with Web 2.0 applications, was found to be a major factor that influenced their susceptibility to user reviews conveying inconsistent information. These results provided implications for designing health information websites that better support older adults' credibility judgments.

ACKNOWLEDGEMENTS

This thesis would not have been possible without the support of many people. First of all, I would like to express profound gratitude to my advisor, Assistant Professor Wai-Tat Fu, for the guidance he has given me along the way, and the inspiration I have learned from him for being a good researcher. Also I would like to thank to Professor Dan Morrow, for reviewing this thesis, as well as offering me the great opportunities to work with him during my master study. Also thanks to my colleagues, Jessie Chin, Page Roth, Wei Dong, etc, for their assistance and helpful suggestions. In addition, I appreciate the help of the staff of Human Factors Division, Peter Vlach and Karen Ayers. Lastly, I would like to thank to my parents, Ke Liao and Qinghua Luo, for supporting me through the process of earning my master degree.

Table of Contents

CHAPTER 1 INTRODUCTION.....	1
CHAPTER 2 RELATED WORK	5
CHAPTER 3 HYPOTHESES AND RESEARCH QUESTION.....	16
CHAPTER 4 EXPERIMENT 1.....	20
CHAPTER 5 EXPERIMENT 2.....	36
CHAPTER 6 CONCLUSIONS	51
CHAPTER 7 GENERAL DISCUSSION.....	53
REFERENCES	57

CHAPTER 1 INTRODUCTION

According to the Pew Internet Report (Pew Research Center, 2010), 70% of population aged 56-64 and 38% of population aged 65 and older are Internet users. The Pew report also indicates that four out of five Internet users have searched the Web for healthcare information, such as checking on specific diseases and treatments of certain conditions. Given older adults' higher needs for health information in support of their medical decisions, they are often found to be actively engaging in the consumption of online health information. The Pew survey on the popularity of different Internet activities shows that, for the older generation, visiting healthcare information websites was the second most popular one just behind using search engines.

A systematic review showed that 70% of the studies that evaluated health information websites found that the quality (including content quality and presentation quality) of online health information to be unsatisfactory, or even inferior (Eysenbach et al., 2002). Given this situation, older adults' increasing exposure to the World Wide Web (WWW) may increase their vulnerability to low quality, even misleading information for their health. Given that it is difficult to control what information will be available on the WWW, the ability to correctly judge the credibility of information is an important "self-defense" skill against misleading information. Credibility is particularly important for health information because misinformation for making medical decision can be hazardous, or even life threatening. This problem becomes more pressing considering studies have found that, unlike other routine online activities such as reading news or browsing entertainment sites, health information consumers often do not have a trusted brand name or Web portal in mind (Eysenbach, 2007). Instead, they often rely on search

engines to reach for information concerning a particular health problem, and therefore can possibly bypass the home page, which may contain more source information (Hargittai, 2010), and encounter information in a random style. Moreover, since users' health related information seeking is often initiated by demand, they tend to apply or act on the information immediately (Eysenbach, 2007). This further aggravates the potential harm brought by unreliable online health information.

Although important, it is often difficult to implement an effective, top-down quality control mechanism for the largely anonymous, massive amount of health related information on the Web. To cope with this problem, researchers suggested two major approaches to mitigate the potential harm of low quality online health information to consumers: First, researchers have studied how to improve users' ability to make correct judgment of information credibility (e.g., Sillence, 2006; Eysenbach, 2007). Credibility judgment is regarded as the key early stage of information quality evaluation because it would determine users' acceptance or rejection of message from the very beginning (Rieh, 2007). By identifying the factors that impact users' credibility judgment, these studies aimed at helping users to overcome the barriers of distinguishing between trustworthy information and dubious one.

The second approach advocates quality control at the users' end by taking advantage of Web 2.0 technologies such as systems that allow posting of user reviews (O'Grady, 2006; Eysenbach, 2007). Similar to product reviews (e.g., Amazon), user reviews provide health information seekers with additional credibility cues and supplemental information, as well as a "group opinion" for evaluating the product or service in question. This approach is expected to guide users towards high quality information through a collaborative filtering process. However, such peer review system may add another layer of complexity to the credibility judgment processes,

since the largely anonymous and unfiltered user reviews themselves demand credibility assessment, which may in turn interact with the original credibility assessment of the message contents. This problem becomes more pressing as merchants and even fraudsters start to utilize the power of electronic Word of Mouth (eWOM). Their manipulation of user reviews could further obscure the boundaries between unbiased and commercial contents on the Internet.

The volume of research on age differences in using the Internet showed that older adults tend to exhibit distinctive behavior in terms of searching, comprehending, and evaluating online information when compared to younger adults (Czaja, 2009; Hanson, 2009; Chin & Fu, 2010). Many aspects of age-related physical and cognitive declines may increase older adults' difficulties in effectively consuming online health information. Moreover, the older generation is in general less involved in the Web environment. Very often, their inadequate knowledge with information technology and Web elements make it even harder for them to thrive in the Internet environment. Therefore, applications that are developed primarily for younger adults may be suboptimal, or even unusable for older adults. A study aiming at unpacking how factors related to aging may influence older adults' credibility judgments may help increase our understanding on how older adults could benefit more from the freely available massive amount of online health information.

The current paper has two main goals. First, I am interested in investigating the differences in the credibility judgment processes of online health information by younger and older adults. To this end, I conducted a laboratory study to compare older and younger adults' credibility judgment of online health information, and collected detailed verbal protocols during the process to reveal differences in the strategies that they used. To further investigate the differences in credibility judgment processes, I also measured participants' cognitive abilities, Internet experiences and

health related domain knowledge, and tested whether these factors are related to the differences I found in the processes. I chose these factors because they are often considered the well-studied age-related factors in terms of technology use by the older generation. My second goal is to explore the effects of user reviews on credibility judgment, in order to understand whether these user reviews could induce further differences in the credibility processes, and the extent to which they could help younger and older adults to make credibility judgment. Specifically, I am interested in whether older and younger adults would react differently to user reviews that are consistent or inconsistent with the credibility cues associated with the website contents. Before describing the details of the studies, I will first review relevant research that shaped the design of the experiments.

CHAPTER 2 RELATED WORK

Web Credibility

Credibility is often studied as a multi-dimensional concept. Source credibility has been studied in the field of psychology for a long time. It was generally defined along two dimensions: trustworthiness and expertise. Trustworthiness refers to the degree to which a receiver perceives the assertions made by a speaker are valid. Expertise refers to the extent to which a speaker is perceived to be capable of making correct assertions. Fogg (2003b) identified four types of credibility based on different types of evaluation: presumed, reputed, surface and experienced. Presumed credibility refers to the general assumptions about someone or something is credible or not in the perceiver's mind. Reputed credibility describes the perceiver's belief of someone or something is credible or not because of third party's assessment or report. Surface credibility refers to the perceiver's belief based on simple inspection of the object. Experienced credibility refers to the believability based on first-hand experience as people interact with the object over time.

Recently, Web credibility has drawn increasing attention from researchers in the field of Human Computer Interaction. One major genre of the literature aims at identifying factors that influence users' credibility judgments with websites. Rieh (2000) found that source characteristics, which include the institutional authority and individual (author, creator, etc) authority, were the primary criteria people used when making judgments of Web information credibility. Liu's analyses conceptualized credibility assessment as a cognitive process by which information is filtered and selected. Also his study revealed that information's resonance with one's belief, novelty, trustworthiness and high quality have positive effects on perceived credibility. Flanagin and Metzger's (2006) study conceptualized users' perception of credibility of a Web site along three

dimensions: message credibility, sponsor credibility and site credibility. Slightly different, Sundar (2008) and his colleague focused on factors afforded by the Internet media that are capable of cueing cognitive heuristics pertinent to credibility assessment. Their MAIN model identified four such aspects of Website: the modality in terms of presenting textual, audio or video based information, the agency in terms of source authoritativeness, the interactivity of website, and its navigability. One important study for Web credibility features is the Stanford Web Credibility Research Project (Fogg et al., 2001). Involving more than 2600 participants, this study identified a number of features that contribute to users' credibility judgments including design look, information structure, information focus, company motive, usefulness of information, accuracy of information, name recognition and reputation, advertising, bias of information, and tone of writing.

From the information processing perspective, a number of researchers worked on establishing theories to uncover the general patterns of credibility assessment. Among them, stage models proposing sequential activities for credibility assessment are especially helpful for us to understand the process of users' credibility judgments. For example, based on their large scale survey results, Fogg's Prominence-Interpretation Theory (Fogg, 2003) considers credibility assessment as a two-stage process: user first notices something prominent then makes judgment by interpreting it. Wathen and Burkell (2002) proposed a three stage model for credibility assessment: surface credibility, message credibility and content evaluation. First the users will rate the Website credibility based on its surface characteristics such as appearance, usability and information organization. Second, the user will rate the source credibility according to the message quality. Third, the content assessment involves integrating message content and source information with their expertise, domain knowledge, information need and situation. Similarly,

in the context of searching for health advice on the Internet, Sillence's(2006) stage model of trust indicated that users start evaluating health website by first impression and heuristic analysis, then further involve with the site and perform more careful evaluation of the site content, and eventually build subsequent relationship with the site by integration of information across sites and sources, and long term consultation with the particular website.

While these models provided good insights into the credibility judgment processes, most of these theoretical models have not been directly validated by empirical data on the detailed cognitive processes. For those who have empirical evidence, most of them were built upon qualitative results or collected in a reflective situation. I believe that a quantitative study focusing on the credibility evaluation process will provide a good test for these Web credibility models.

Moreover, using verbal protocol analysis to collect first-hand data may be a more valid method than surveys or interviews to understand how people perform credibility judgments (Ericsson & Simon, 1984).

Elaboration Likelihood Model

Many credibility studies were built on some forms of the persuasive communication theories (Jones et al., 2003; Metzger, 2007). This is because credibility assessment is often considered a key early stage in the persuasion process. It will determine people's acceptance or denial of information for further processing, and thus influence the impact of a message on the receiver's attitude formation. One widely cited model among these is the Elaboration Likelihood Model (ELM)(Petty & Cacioppo, 1986). The ELM explains attitudinal changes in individuals as they encounter two distinctive types of cues: central cues are related argument/message quality that requires systematic, deliberative processing. Persuasion through central cues is attained by users' careful consideration of the merits in the arguments in support of advocacy. Peripheral cues are

related to surface features of the information, which can be processed in a heuristic way by relying on practical rules or experience. Peripheral cues can pertain to either source (e.g. reputation of the source) or message (e.g. length of message). Persuasion through peripheral cues can be relatively easily induced. Therefore people who lack either ability or motivation for analyzing the information tend to resort to the peripheral cues to form their attitude. With respect to Web credibility, central cues can be associated with the content of Web information, while peripheral cues may be associated with the contextual features of the websites (e.g., interface design, usability, source information, etc). In this thesis, I will use content cues and contextual cues to refer to the two types of cues particularly for Websites.

There were only a limited number of studies that looked into how cues from these two different routes interact and impact users' credibility judgment. For example, Hong (2006) examined the influence of both message and structural features on perceptions of Website credibility. However, the interaction between content cues and contextual cues appears to be an interesting issue since more often than not, users may receive cues that contradict each other. This is especially important for online health information, as users tend to have high expectation for both information quality and source trustworthiness. In fact, one study showed that for healthcare websites, features of website credibility have only small, or at best moderate, correlation with the accuracy of information it provided (Kunst, 2002). This implies that apparently credible website may not necessarily provide reliable health related information.

Since the processing of content cues tends to be more cognitively demanding, studies proposing the stage model of credibility assessment often found that users tended to initiate credibility assessments by peripheral or heuristic processing regarding contextual features of the Websites(Sillence et al., 2007; Hilligoss & Rieh, 2008). However, there existed mixed results

regarding the effects of specific contextual cues on users' credibility assessment outcomes. There were substantial evidences that notably better interface design and aesthetical features contributed to higher perceived credibility (Alsudani & Casey, 2009). However, when users were not asked to judge the credibility based on these specific features, they were often found to disregard certain specific website features such as third party endorsement, advertisement, etc, when making credibility judgment (Shon et al., 2000; Riegelsberger, 2002; Walther, 2004; Hong, 2006), even though these features were considered by users as important indicators of credibility in studies that were based on interviews or surveys.

When it comes to health related Website, Eysenbach(2002) conducted a study using focus group, observational study and interview to examine how consumers appraise the quality of online health information. Interestingly, during observational study, few participants actually looked at the features they placed emphasis on in the focus group part of the study, and the post-search interviews revealed that few participants took notice or remember about the features of the websites where they retrieved information. These results implied that the application of Elaboration Likelihood Model to Internet can be a complicated issue. This is probably because Web media, as compared to traditional paper and television media, provides more abundant and diverse peripheral cues, and may induce heuristics that have varied availabilities for different individuals.

The Elaboration Likelihood Model lays a good theoretical framework for my studies. It suggests that people are sensitive to the credibility cues derived from both the message content and the contextual features on the Website. While previous studies were conducted primarily on younger participants, there is still a lack of systematic investigation on how older adults respond to the two types of credibility cues differently. These potential age differences could lead to a better

understanding of older adults' capabilities and preferences in evaluating health information on the Internet.

User Reviews

Nowadays, the emergence of concepts like Medicine 2.0 and Health 2.0 indicates increasing adoption of Web 2.0 technologies on the arena of online health information. The employment of user review system to facilitate users' credibility judgment exemplified Meola's(2004) contextual approach to information quality control by using external social resources to establish information quality indicators. It may outperform the traditional approaches that aim at developing checklist to guide users' credibility assessment, as it may relive users from exerting much effort to actively verify the credibility of online information.

As an important phenomenon of social influence, the impact of collective opinions of others on perceiver's attitude and decision has been studied for a long time. People have a drive to compare their own opinions with others and often adopt the collective opinion of others. This tendency to conform to the "group opinion" is motivated by the goals of achieving more accurate perception or decision of their own, creating and maintaining meaningful relationships with others by conforming to the group opinion, and enhancing their self-esteems by behaving consistently with others(Cialdini & Goldstein, 2004).

Although the impact of user reviews on credibility judgment is a relatively new topic in the area of health information, they have been extensively studied in the field of e-commerce (e.g., Gilbert & Karahalios, 2010). User recommendation system has become an essential part of many online retailers' economic models and developed into considerably sophisticated technology. Complex algorithms are often used to provide recommendations that are most informative and

relevant to the particular user based on their stated preferences, online shopping history and user profile. This technology is proved to help driving online sales. For example in the booming music industry, download of popular songs were seen multiplied after the utilization of user recommendation system. Previous studies on consumer trust indicated that social presence, the perceived social context, and the sense of the coexistent of other communicators, could increase consumers' trust in the e-Vendor (Gefen & Straub 2004).

Online user reviews could exert both informational and normative social influence upon individual's judgment in terms of product credibility and purchase intention (Cheung et al, 2009). When reading reviews, users may take the judgments of others as evidence and supplementary information for their decision making; on the other hand, they may also be susceptible to other social processes such as the norm of complying with other's opinion and conforming to "group opinion". The informative influence of user reviews could be strengthened by informative determinants, including argument strength in the review content and the recommendation framing in terms of how negative or positive the reviewers express their opinions. The normative influence of user reviews could be enforced by normative determinants such as recommendation consistency among reviews and recommendation ratings given with the review.

Online user reviews may not change people's mental representations as much as one may expect but rather only act as anchors for people's judgment making. Previous studies showed that users were able to focus on the actual content of item even when the user reviews were presented, and their attitude and decision may be less influenced if the user review encountered later does not match the format of the initial encoding of information, or their impression of the product (Sakamoto, 2010).

Another factor that may influence the effects of online user reviews on peoples' attitude change is the message discrepancy, which refers to the level of difference between the user's original judgment and the claimed position in user reviews. In fact, attitude change, message discrepancy and source credibility were often studied together in literatures regarding attitude change and persuasion. A common finding was that, in high credibility condition, where the recipient perceived the message came from a highly reliable source, the level of attitude change was positively correlated with the discrepancy. However, when the source has low credibility, while their attitude change still positively correlated with discrepancy when the discrepancy is minor or medium, it starts to drop when the discrepancy reached to a certain high level(Chung, 2008). This line of theory carries two implications for my study: first, the level of users' credibility judgment of online medication information may vary depending on the discrepancy between the original content about the item and the argument strength in the user reviews. Second, it is important to consider how the user reviews are presented on the Website. If the perception of the sources who give comments varies in credibility, they may be predetermined to have different influence.

While lately researchers are actively studying social networking service, there is a general lack of studies analyzing how the user generated contents exert influence on older adults differently from younger ones. The pervasive adoption of Web 2.0 component on Internet makes it inevitable for health information seekers to encounter user reviews when looking for medication or other health related product. Whether the employment of user review system and other Web 2.0 features will benefit or hinder older adults' online health information consumption, while still remains largely unknown, is greatly worth analyzing before their further and wider intervention into healthcare websites.

Aging and Credibility Judgment

Research on cognitive aging provided both “bad news” and “good news”. The bad news is that aging is often associated with a variety of cognitive declines, which are often referred to as fluid intelligence. The good news lies in the fact that older adults often have vast stores of knowledge and experience for the situation, which is often referred to as crystallized intelligence (Hayslip & Sterns, 1979; Beier & Ackerman, 2005). There is considerable evidence that when older adults perform a complex task that they are familiar with, their knowledge and experience may moderate their declines in cognitive abilities and enable them to perform at a high level (e.g., Fairweather, 2008). One of the major challenges of cognitive aging research that has not been well addressed is to understand the joint effect of a decline in fluid intelligence and a possible growth in crystallized intelligence on older adults’ daily activities.

When it comes to computer-mediated environment, the issue of aging is even less well studied. On the one hand, older adults’ generally declined cognitive functions, especially lower processing speed and limited working memory, as well as some physical limitations, such as generally worse manual motor skills (which may influence their use of keyboard, mouse, etc) may impede their acceptance and further engagement with new technology. On the other hand, from the “crystallized intelligence” aspect, while older adults may be familiar with the target information or task, they are, however, generally less experienced with the Internet media. It is hard to predict whether older adults could successfully transit their knowledge and experience accumulated in the offline activities into the online environment. Therefore the question of how all the aspects of aging work together to influence older adults’ behavior and performance in online environment remains unclear. A number of such studies reported often worse information search performance of older adults, which was attributed to their lower fluid cognitive ability and

poorer knowledge about Web features (Sharit, 2008). However, there were also studies indicating that older adults can outperform younger adults in cases where only general, unspecific information needs were given for a search task. While younger adults tended to use more bottom-up, interface driven strategies (e.g., click on each link without pre-comprehending of its meaning), older adults tended to use more top-down, knowledge driven strategies (e.g., only click on links that they perceive to make sense to the task), which illustrated that older adults were able to develop effective strategies to adapt to the Internet environment (Chin & Fu, 2009)

The primary focus of the study was to explore age differences in credibility judgment of online health information. While there were a number of studies looking into older adults' information search behavior with health related information, there were, however, few studies explored how older adults evaluate information credibility or quality on health information websites. Results from a survey-based study (Czaja, 2009) indicated that older adults' trust with Internet health information was influenced by website identifiers (e.g. government agencies) and design features, which underscore the importance of contextual cues, i.e. website features that are not integral part of the message, to older adults' credibility judgments.

Considering that credibility judgment is often regarded as attitudinal change induced by content and contextual cues, previous studies regarding age difference in dual processing model may provide some insights into this particular issue (Peters et al., 2007). Studies have provided robust evidence for age-related declines in the efficiency of systematic/ deliberative information processing of central cues, which could be associated with older adults' less efficient information processing, deficits in memory, decrease in the ability to inhibit false and irrelevant information, and less aware of factors that influence their judgments. However, there were less clear results regarding age difference in the processing of peripheral cues. While a number of studies showed

this difference is only minimal, there are evidences that older adults have higher tendency to engage in processing information when there are peripheral cues that could induce emotional changes or relate to their personal experiences. Specifically for Elaboration Likelihood Model, closely related to this study were Yoon and her colleague's (2009) analyses on persuasion and older consumers' behavior. By examining older consumers' attitude change within the Elaboration Likelihood Model framework, they found that the ELM effects, i.e. both central cue effects and peripheral effects on persuasion, were consistent with prior results obtained on younger adults. Interestingly, they found that when presented with information that was high in personal relevance, although older consumer's motivation levels were elevated, the persuasion remained consistently lower. This result suggested that older adults are generally more skeptical about persuasive messages related to domain that carry greater risk or personal relevance. In this sense, it is possible that older adults are less prone to persuasion when processing health-related information.

Another stream of research that may be relevant to this study concerning the effects of user review is the age differences in subjectivity to external persuasive influence on attitude change. Generally speaking, research on aging and attitude change reported that resistance to external influence increases with age. Especially in cases investigating persuasion paradigm which entails presenting detailed arguments to convince a message recipient to adopt an advocated opinion, older adults were less deviated from their original attitudes. The age difference is likely attributed to skills developed with age to defend oneself against persuasion. Also it was shown that older adults have higher internal locus of control and tend to disregard external influence in the presence of internal reinforcement of their own opinion (Lynn, 1977).

CHAPTER 3 HYPOTHESES AND RESEARCH QUESTION

According to the related works reviewed in last section, this thesis was based on previous studies regarding Web credibility and studies on the effects of aging in using Internet. The Elaboration Likelihood Model (ELM) indicates that credibility judgments with online information are influenced by two types of cues: cues in the message content and cues in the contextual features of the Website. In the first experiment, I will manipulate the strength of credibility cues in message content and website features for a number of Web pages, and ask a group of older adults and a group of younger adults to perform the credibility judging task. By comparing their credibility ratings as well as their credibility judging process using verbal protocols, I will investigate the age difference in the abilities of making correct judgments with cues from the two types of sources (i.e., content cues and contextual cues). According to ELM theory, cognitive ability is an important factor that determines the depth of central processing. Considering older adults' generally declined cognitive abilities, they may have more trouble processing content cues as compared to younger adults. On the other hand, since a large proportion of contextual cues is Web specific, e.g., site ownership information, third party endorsement stamp, older adults who are often unfamiliar with Internet environment, may not be able to make sense with all of them. Hence I made the following hypotheses which I expect to see in the first experiment.

Hypothesis 1: Older adults are less sensitive to credibility cues in message content compared to younger adults, and give closer ratings to Web pages with content of high credibility and those of low credibility.

Hypothesis 2: Older adults are less sensitive to credibility cues in Website contextual features compared to younger adults, and give closer ratings to Web pages with contextual features of high credibility and those of low credibility.

Moreover, age difference may not only lie in the credibility rating, but also the strategies that older and younger adults use to make the judgment. This thesis also attempted to answer the following research question:

RQ1: What are the differences in older and younger adults' credibility judging process, which may lead to the potential age differences in their credibility ratings?

To further understand the causes of age differences in credibility judgment of online health information, I will analyze the effects of individual differences on older and younger adults' credibility rating. First, according to ELM theory, cognitive ability may determine the depth of information processing through central route, which is concerned with the ability of making credibility judgment with content cues. Therefore I expect the declines in cognitive ability may be associated with older adults' insensitivity to the credibility cues in Web page content. Second, for credibility judgment with Web information, a large portion of the contextual cues are unique to Internet environment, such as site ownership information, third party endorsement, etc.

Therefore I infer that Internet experience may play a role in users' sensitivity to contextual cues for credibility judgment. Third, research based on ELM suggested domain knowledge could facilitate individual's deliberative processing of content. Also previous studies have shown that topical knowledge influence users' perceived credibility of Web information (Ferebee, 2008).

Hence I made the following hypotheses:

Hypothesis 3: Older adults who are more cognitively declined have more trouble differentiating between message contents of high credibility and those of low credibility.

Hypothesis 4: Older adults who have poorer Internet experience have more trouble differentiating between website contextual features of high credibility and those of low credibility.

Hypothesis 5: Older adults with better health related domain knowledge are more capable of performing correct credibility judgment of online health information.

The influence of user review on older and younger adults' credibility judgments is a particularly interesting question. While previous studies suggested user review may exert great influence on people's credibility judgment about the original contents, I would like to explore the age difference in such influence. For this purpose, I will conduct another experiment to investigate the age differences in making credibility judgments with online health information. I will add user reviews to the original Web pages used in the first experiment, and manipulate their consistency to the credibility cues in the message contents. Then younger and older participants will be randomly assigned to tasks with or without user reviews. By comparing their credibility ratings, I will explore the age difference in susceptibility to the influence of user reviews on credibility judgments.

User reviews that are consistent and inconsistent to the credibility of original content may exert different influence on older adults' credibility judgment. By providing additional credibility cues to further confirm people's credibility judgment formed by reading the content, consistent user

review may have positive effects on users' ability of making credibility judgment. On the other hand, previous aging research in attitude change suggested that older adults were less subject to the influence of external persuasion, which intended to change recipients' initial attitude. Especially when the topic is highly critical for older adults' daily life, e.g. health and finance related topic, older adults were even more skeptical about the persuasive messages (Yoon, 2009). Therefore I expected older adults to be less subject to the influence of inconsistent user reviews. I made the following hypotheses for the second experiment:

Hypothesis 6: User reviews that are consistent with the credibility of original message content could reinforce older adults' credibility judgment.

Hypothesis 7: Older adults are less subject to the influence of user reviews that convey information inconsistent with the credibility of original content than younger adults.

Moreover, I am interested in the effects of Internet experience on younger and older adults' susceptibility to the influence of user reviews. Previous research suggested that the general trust on Internet information will increase as user become more experienced with Internet environment. Since older adults are in general less experienced with the Internet, especially Web 2.0 applications, it is possible that Internet experience will contribute to the potential age difference in subjectivity to influence of user reviews. Hence I made the following hypothesis:

Hypothesis 8: The difference in Internet experience contributes to the potential age differences in the susceptibility to the influence of user reviews.

CHAPTER 4 EXPERIMENT 1

In this study, I conducted a laboratory experiment to explore age differences in making credibility judgments with medication information found on the Internet. Specifically, I am interested in how content argument strength (content cues) and presence or absence of credibility-related Web page features (contextual cues) would differentially influence younger and older adults' credibility judgment, as well as the causes of such age differences. I am also interested in the effects of aging related individual differences on the age differences in making credibility judgment. Specifically, I focused on the effects of decline in cognitive abilities, lack of Internet experience and health related domain knowledge on credibility judgment in this study.

Participants

Sixteen older adults (between 62 and 80 years old, Mean=69.38, SD=5.81; 66.7% were female) and sixteen younger adults (between 19 and 26 years old, Mean=21.56, SD=2.10, 50% were female) participated in the study. All participants were recruited from the Champagin-Urbana area in Illinois, USA. There was no significant difference in the education level between two age groups and most participants (90.6%) have finished at least some year of college. Also there was no significant age difference in the self-reported experiences in seeking health-related information online.

Experimental Design and Materials

A 2×2×2 mixed factor design was used to study older and younger adults' credibility judgment of health information Web pages. There were two within subject variables: content cue strength (strong/weak) and contextual cue strength (strong/weak), and one between subject variable: age (young/old). All participants performed eight credibility judgment tasks, with each task

composed of four Web pages that corresponded to the four possible combinations of strong/weak contextual and content cues.

Materials

In the experiment, articles of alternative medicines were presented on Web pages I created. I used alternative medicines for two major reasons: 1) as argued by many researchers (e.g., Eysenbach,1999), medication information on Internet is expected to take important role in medical informatics and public health, and are especially applied to the purpose of disease prevention and health promotion, and 2) in this experiment, I intended to understand credibility judgment by participants' attitude towards the content (i.e., the medication information) as presented on the Web page not by directly asking them to explicitly judge the credibility of the content, but by providing them with a scenario in which credibility assessment is a natural component in the evaluation process (Metzger, 2007). I believe this is closer to the credibility judgment people perform in real context. Because alternative medication tends to involve a relatively higher level of uncertainties and potential risk, coupled with the rampant “fake medicine” over the Internet, I expect that it would encourage users to naturally invoke the credibility dimension of information evaluation.

Content Cue Manipulation

For content cue manipulation, I adopted the empirical method used by Petty and Cacioppo (1986) to verify the argument strength of the contents shown on the Web pages. First, I selected materials from a popular healthcare website (www.revolutionhealth.com). It has articles of alternative medicine for different diseases, with ratings and reviews provided by users and professionals. I chose diseases with the highest number of reviews to be included in the experiment, which also tended to be common diseases (e.g., diabetes, hypertension, obesity),

which were likely to be familiar to both age groups. Based on the review ratings I selected articles to be materials with “strong” and “weak” content cues. I further modified their use of evidence, argument rigor, information quality and bias, all of which were identified to be information credibility indicators (Hamilton, 1998; Fogg, 2003; Rieh, 2007) (Table 1). I also balanced their length and amount of information to create a set of articles. The medicine names were modified such that they could not be recognized. To further verify the manipulation, I asked a group of 7 pilot participants naïve to the experiment to rate the credibility by reading the article, which means only processing the content cues. I filtered out articles that had the lowest consistencies among the pilot participants and ended up with 8 sets of documents.

Contextual Cue Manipulation

To manipulate the contextual cue, i.e. website features, I focused on features from two categories: design look and source features. Fogg et al.(2003) identified that design look, including layout, typography, white space, images, etc, to have the largest impact on web credibility evaluation. Source features are features that indicate the source authority and expertise, which include references, contact information, privacy policy statement, third-party endorsements, site ownership, displayed awards, commercial features, etc (Hong, 2006). First I selected web page templates from highly recognized and professionally designed healthcare websites (e.g., medlineplus, livestrong, mayo clinic), based on their public reputation, website traffic, and endorsement by authoritative third party such as the Health on the Net Network (HON). I adopted the design look and source features of these Web pages to act as “strong” contextual cues, and randomly chose to deliberately remove 3-5 features from these two categories to create Web pages that had “weak” contextual cues (Table 1). This approach allowed us to avoid focusing on specific Website features which repeatedly happen to every task I used.

Measures of cognitive ability and internet experience and domain knowledge

According to research on dual processing models, a variety of individual variables related to ability and motivation related to information processing can affect outcomes of judgment and decision making, i.e., credibility judgment results in this case. Based on previous studies on age differences, I chose to focus on cognitive ability, Internet experience and domain (health) knowledge, as I expected that they would likely influence participants' ability and motivation to process online health information

The general decline in cognitive function is one major focus of aging research, and also the primary causes of age difference in using technology (Sharit , 2008; Hanson, 2009; Chin & Fu, 2009). For cognitive abilities, I focused on fluid mental abilities (e.g., working memory and processing speed), which are found to be most vulnerable to effects of aging. Previous studies have also identified these abilities to be some of the major causes for older adults' tendency to perform worse in deliberative processing of content cues as compared to younger adults (Peters,2007). Following previous studies, working memory was measured by the Letter Number Sequencing Task (Salthouse, 1991), while processing speed was measured by an automated version of the Operation Span Task (Unsworth,2005).

Research on age differences in Internet use often found that older adults tended to have less experience with Internet. It was often used to explain why their online behavior was different from younger adults. Also previous study suggested that Web use experience could affect individual's credibility evaluation with health related Web information (Dutta-Bergman,2004). To measure Internet experience, I randomly selected twelve questions from the Knowledge-related Internet Information Seeking Semi-structured Interview (KRIISS) (Sharit, 2008), which was developed by Center on Aging at University of Miami to comprehensively capture the

difference between people with different levels of Internet experience. The interview asks wide range of questions regarding how the Internet works, how to use Web browser tools and how to perform information search task.

According to the theory of ELM (Petty & Cacioppo,1986) domain knowledge could facilitate individual's central route processing. Those who have rich domain knowledge tend to have more distinct ratings in face of strong and weak central cues, suggesting that they are capable of making more accurate judgment of the strength of central cues. Also previous studies have shown that topical knowledge will influence a user's perceived credibility of Web information. By comparing the credibility results of average users and experts in health fields, Fogg(2003a) found that experts were considerably less misled by surface features or visual appeal of a Website and expressed more concern about the quality of a health related site's information. In the study, task related domain knowledge was measured by a fluency task, in which participants were asked to generate as many relevant keywords as possible for each of the eight diseases I used in experiment. The average number of keywords for each disease, as an indication of their retrieval of related concepts from memory (Griffiths,2007), was used as an index of individual's task related domain knowledge.

Task and stimuli

Before the experiment, all participants were given the set of standardized pretests to measure their cognitive ability and Internet experience as described above. Participants then read the instruction and the scenario description of each task. Participants were instructed to imagine that they were helping a friend to evaluate some alternative medicines randomly collected from the Internet. The concern for potentially ineffective or even fake medicine was mentioned to implicitly induce their motivation for judging the credibility of each page. Participants were then

presented with the task interface, which presented the 4 web pages under each of the 8 diseases on a regular Web browser. Participants could then click one of the disease names and browse any of the four web pages, each of which described an alternative medicine. Each of the articles had four parts: introduction, side effects, interaction, and dosage information of the medicine. These parts were typical medicine information on most medication websites. Participants could then click on “Rate” button on the interface and submit their ratings of the medicine based on a scale from 1 to 7 (1 =not recommend; 7=highly recommend).

To understand the evaluation process, verbal protocols were collected from participants by asking them to “think aloud” during two of the eight tasks. All the protocols were recorded as digital files by the computer and later transcribed and analyzed.

RESULTS

I divided the analysis into two parts: First, I tested whether there were age differences in the credibility rating given to the medicines (Hypothesis 1 and Hypothesis 2), and looked into the causes by performing verbal protocol analysis (RQ1). Second, I explored the effects of individual differences in cognitive ability (Hypothesis 3), Internet experience (Hypothesis 4) and health related domain knowledge (Hypothesis 5) on younger and older adults’ ability of making correct credibility judgment.

Age Difference in Credibility Judgments

Credibility Rating Analysis

To understand the effects of age difference on credibility judgment of online health information, I performed a three-way ANOVA on their credibility ratings, with content cue strength and contextual cue strength as within-subject variables, and age as between- subject variable. The

results showed that the main effects of content cue ($F(1,30)=22.04, p < 0.01$) and contextual cue ($F(1,30)=41.81, p < 0.01$) were significant. Also the interaction between content cue and age ($F(1,30)=4.18, p=0.05$), and interaction between contextual cue and age ($F(1,30)=5.60, p=0.03$) were significant. No other effect was significant.

The main effects of content cue and contextual cue validated the manipulation of content and contextual cues, as higher credibility ratings were given to Web pages with strong content cue and strong contextual cue rather than those with weak content cue and weak contextual cue. Interestingly, the two-way interactions were caused by the fact that older adults gave closer credibility ratings between strong and weak content, and between strong and weak contextual cues (see Figure 1). The results were consistent with our hypotheses that older adults were less sensitive to credibility cues in both Web page contents and contextual features. Planned comparisons showed that while younger adults could successfully differentiate between strong and weak contextual cues ($F(1,15)=19.10, p < 0.01$), old adults were not able to do so ($F(1,15)=3.90, p = 0.07$). The result indicated that older adults had difficulties in differentiating between strong and weak contextual cues.

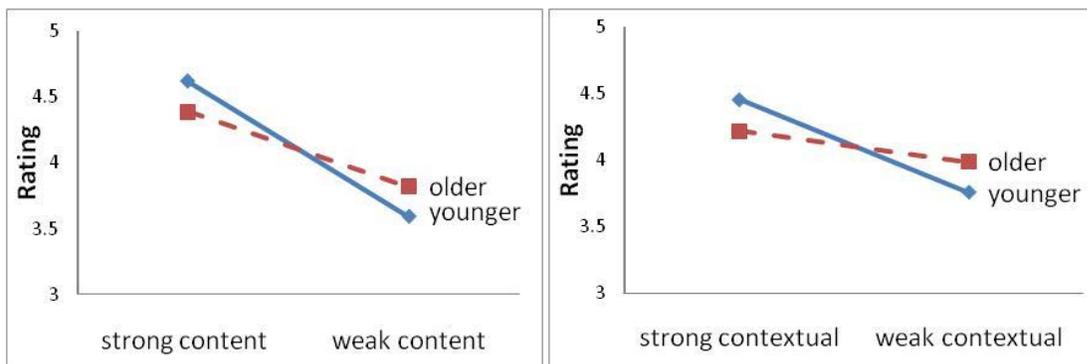


Figure 1. Credibility rating for pages with strong or weak content/contextual cues

Protocol Analysis

To further investigate the causes of older adult' lower sensitivity to credibility cues, I collected verbal protocols from each participant for two of the eight tasks (altogether 8 pages). I classified the transcribed protocols based on whether they were about the content or contextual cues. For content cue processing, I further differentiated protocols showing that participants were passively reading facts or statements on a page from those showing that they were actively deliberating on information they read. This differentiation has important practical implications because more passive reading than active deliberation of information could imply that participants were less sensitive to the information quality or logical deficiencies of the page content, and could potentially be misled or misinformed.

I further broke down the three major categories of protocols (facts reading, deliberation, and contextual cues) into subcategories (See Table 1). For each protocol, I also coded whether the information had a positive, negative, or neutral impact to their credibility judgment. I did this based on whether the words they used to describe the information was positive, negative (e.g., "that's too bad" or "it seems great"), or if they were simply describing the information without any expression of how it impacted their judgment (i.e., neutral). I calculated the correlations between younger and older adults' final rating and the difference in the number of positive and negative cues processed in each page. The reason for doing this was that I wanted to verify whether their final ratings were the integrative outcomes of the processing of the positive and negative cues on each page. The results showed that final ratings were indeed strongly correlated with the difference between the number of positive and negative cues processed (number of positive cues minus number negative cues mentioned for each rating) for both younger ($r=0.81$, $p<0.01$) and older adults($r=0.81$, $p<0.01$).

Categories	Subcategories	Criteria	Examples
Deliberation	1.1 Checking evidence	Checking studies, data, etc	<i>“There is high quality scientific evidence”, “The research looks only preliminary”</i>
	1.2 Evaluating information quality	Commenting on the completeness, accuracy, writing tone, or bias of information,etc	<i>“There is way too much information devoted to healthy lifestyle, not the medicine itself”,</i>
	1.3 Reasoning	Commenting on logical problems, contradictory facts, unclear explanations, etc; doubting claims, motives, etc;	<i>“ It said few adverse effects in the first part, but listed numerous ones in side effects part”</i>
	1.4 Relating to personal experience	Talking about personal experience and preference	<i>“I took similar fiber product before and it helps” “I would not recommend OTC products”</i>
	1.5 Comparson	Comparing with other medicines read	<i>“It works the same way with last one”</i>
Facts reading	2.1 Introduction (effectiveness)	Reading claims in the introduction part, including treaging efficacy, ease of use, history,background, etc	<i>“It lowers cholesterol”, “It has been used in Asia for 1000 years”</i>
	2.2 Side effects	Reading claims in the side effects part	<i>“Side effects included dizziness”</i>
	2.3 Interactions	Reading claims in the interactions part	<i>“Caution advised in people who take drugs lowering blood pressure”</i>
	2.4 Dosage	Reading claims in the dosage part	<i>“The dosage is 25 mg”</i>
Contextual processing	3.1 Design feeling	Aesthetical quality, layout, color, structure,etc	<i>“layout is pretty simple, easy to read”</i>
	3.2 Reference features	Referrence literature, resource links, suggestions for relevant information, etc	<i>“It lists some decent referecne”, “Have links to read abstract of research ”</i>
	3.3 Website source features	Features indicating website reliability, e.g. sponsor information, contact, endorsement	<i>“The website is a non-profit organization”, “American Heart Association recommended”</i>
	3.4 Commertial features	Advertising, promotion, donate button, etc	<i>“The site is covered by advertisements”</i>

Table 1. Coding scheme of protocol analysis for experiment 1

I investigated which major categories of the protocols showed significant age differences. I calculated the percentages of cues mentioned in each category for each participant, to understand whether there were age differences in the relative weightings of different types of cue to make credibility judgment. First I conducted a two-way ANOVA by using age and type of cues as independent variables. It showed that the main effect of type of cues was significant ($F(1,90)=46.33, p < 0.01$), and the interaction between age and type of cues was significant ($F(2,90)=14.55, p < 0.01$). As Figure 2 showed, older and younger adults processed the types of cues differently, which led to the significant interaction. Post-hoc analysis with Bonferroni correction was performed to test the age difference in each type of cue. Results showed that younger adults had higher percentages under the category and contextual cue processing ($p < 0.017$), while older adults had a higher percentage in the category of facts reading ($p < 0.017$).

The results suggested that contextual cue processing contributed more to younger adults' final ratings than that of older adults, while older adults' tended to rely more on accepting the facts they read to make their final ratings.

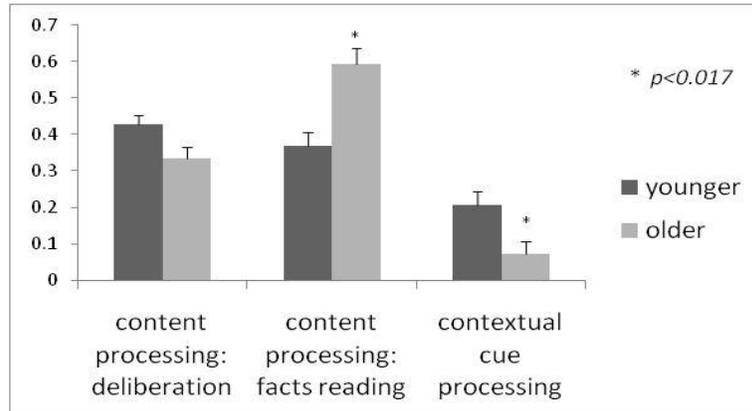


Figure 2. Percentage of cues in each category mentioned by each participant

I was also interested in what cues older and younger adults processed first when they evaluated a Web page. A two-way ANOVA (with type of cues and age as independent variables) on the proportions of the type of cue in the first cue mentioned by each participant was conducted. It showed there were significant main effect of type of cue($F(1,90)= 10.52, p<0.01$) and significant interaction effect between age and type of cues($F(2,90)= 11.99, p<0.01$). The results indicated that older and younger adults processed different types of cues when they first evaluated a Web page. Post-hoc analysis with Bonferroni correction was performed to compare the results of younger and older adults for each type of cue. As Figure 3 showed, younger adults had a higher tendency to process contextual cues than older adults($p<0.017$), while older adults had a higher tendency to read facts of the medicine than younger adults($p<0.017$).

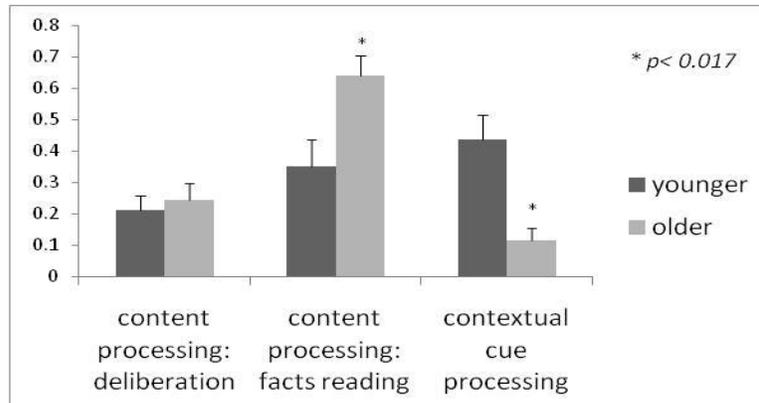


Figure 3. Percentage of different category mentioned for the first cue

Previous studies showed that users tended to first engage in preliminary assessment with the site by processing contextual cues before performing a more systematical, in-depth evaluation of the information on that site (Sillence, 2006; Sillence, 2007; Hilligoss & Rieh, 2008). This was interpreted under the stage model of Web credibility assessment. It is worth noting that these results were not consistent with this model. Older adults had a higher tendency to start by reading facts than deliberation ($p < 0.01$) and processing contextual ($p < 0.01$). This may imply that older adults tended to deviate from “common” behavioral patterns as observed in “regular” Internet users (who were likely younger adults in previous studies). Intuitively, this could be explained by the possibility that older adults were less adapted to processing information on Web pages.

Effects of Individual Differences

In this section, I focused on how differences in cognitive ability, Internet experience, and health domain knowledge could explain differences in older and younger adults’ credibility judgment.

Effects of cognitive ability on content credibility judgment

Consistent with previous studies, the pretest showed that older adults in general had lower cognitive ability than younger adults ($p < 0.01$). To understand how cognitive ability influenced

credibility judgment, I performed median splits in each age group based on the cognitive ability index and compared how the high and low cognitive ability groups differed in their credibility judgment. Within each of these two cognitive ability groups I performed the same three-way ANOVA with content cue, contextual cue and age. By comparing the results I found that the interaction between content cue and age was only significant among users with low cognitive ability ($F(1,14)=10.92, p<0.01$), but not among users with high cognitive ability ($F(1,14)=0.33, p=0.57$). Figure 4 illustrated the differences: for content cue processing, older adults with high cognitive ability could perform almost as well as younger adults. However, older adults with low cognitive ability were less able to differentiate between credible contents and less credible ones. It suggested that the generally declined cognitive ability contributed to older adults' lower sensitivity to credibility of message content.

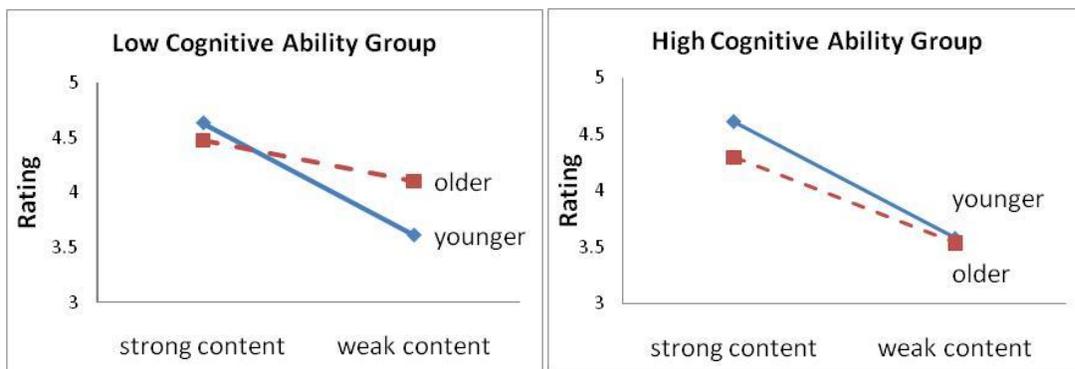


Figure 4. Average credibility rating given to pages with strong or weak content cue by low/high cognitive ability group

Effects of Internet experience on contextual credibility judgment

Consistent with previous studies, our measure showed that older adults were generally less experienced with Internet than younger adults ($p<0.01$). To study the role of Internet experience in credibility judgment, I performed a median split based on the Internet experience index to generate the high and low Internet experience groups, and performed the same three-way

ANOVA in each group. By comparing the results I found that the interaction between contextual cue strength and age was significant in the low Internet experience group ($F(1,14)=5.96$, $p=0.03$), but not in the high Internet experience group ($F(1,14)=0.50$, $p=0.49$). It indicated that older adults with more Internet experience could perform just as well as younger adults in contextual cue processing, while those who lack Internet experience were not able to do so (Figure 5). It suggested that Internet experience was critical for older adults' contextual cue processing when making credibility judgment.

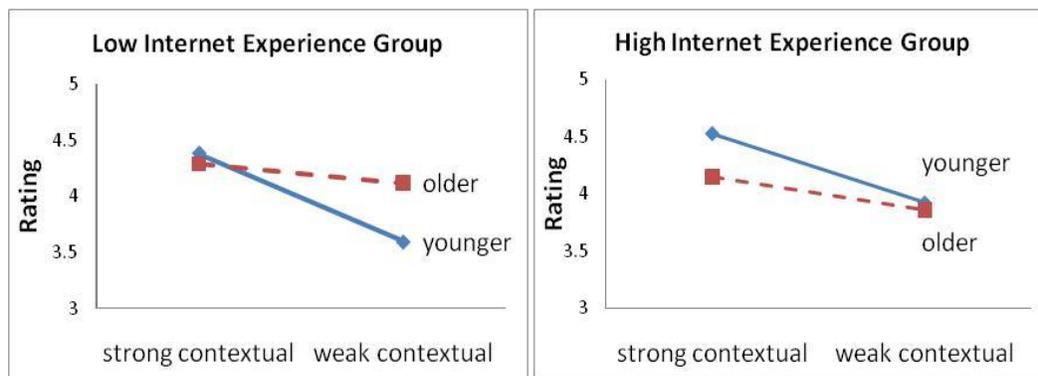


Figure 5. Average credibility rating given to pages with strong or weak contextual cue by low/high Internet experience group

Effects of health related domain knowledge on content and contextual credibility judgment

In this study, there was no significant age difference in domain knowledge between the younger and older groups ($p=0.52$). I divided all participants into groups of low and high domain knowledge by performing median split based on the domain knowledge scores. As shown in Figure 6, in the low domain knowledge group, the two-way interaction between content cue and age ($F(1,14)=4.35$, $p=0.05$), and two-way interaction between contextual cue and age ($F(1,14)=6.09$, $p=0.03$), were still significant. Interestingly, I observed the two-way interaction between age and content cue ($F(1,14)=0.68$, $p=0.42$) and the two-way interaction between age and contextual cue ($F(1,14)=0.50$, $p=0.49$) became not significant in the group of

high domain knowledge. Three-way ANOVA with domain knowledge (high/low), content cue strength and contextual cue strength performed among older adults showed that there was a marginally significant two-way interaction between domain knowledge and contextual cue ($F(1,14)=4.08, p=0.06$), while no similar interaction was observed among younger adults. The results suggested that higher domain knowledge could compensate for older adults' lower abilities in differentiating between strong and weak content, as well as contextual cues. And older adults who had better health knowledge seemed more likely to perform just as well as younger adults in credibility judgments (Figure 6).

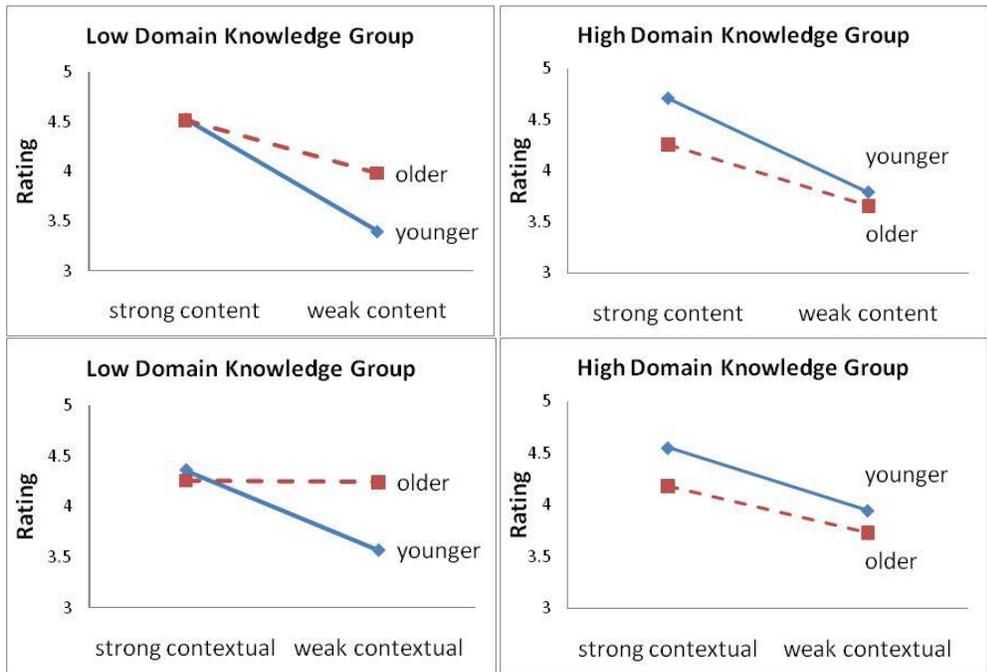


Figure 6. Average credibility rating given to pages with strong or weak contextual/content cue by low/high domain knowledge group

SUMMARY OF RESULTS FROM EXPERIMENT 1

To summarize, I found that older adults were in general less able to make correct credibility of online health information. Specifically, I found that they were less sensitive to the credibility

cues in message content(Hypothesis 1) as well as the website surface features(Hypothesis 2), both of which are considered primary sources of cues for credibility judgment. To answer Research Question 1, verbal protocol analysis was performed to understand participants' credibility judging process. It was found that these age differences could be explained by the differences in the strategies that older and younger adults adopted when processing online health information: 1) For content cue processing, older adults tended to rely on the facts they read in the message content, but less frequently engaged in deliberating on the information quality; 2) For contextual cue processing, older adults would less likely pay attention to features or attributes of the website during credibility judgment. Moreover, while younger adults tended to start with processing contextual cues on the website, older adults would more likely start by directly reading the text on the Web page. The set of results seem to confirm the intuition that older adults' were less adapted to the Web environment than younger adults, and they appeared to be browsing Web pages as if they were processing traditional forms of text such as books or newspapers.

To further understand how aging related individual differences influenced older adults' credibility judgment, I explored the effects of individual differences in cognitive ability, Internet experience and health related domain knowledge on younger and older adults' ability of making correct credibility judgment. It was found that: 1) the generally lower cognitive ability largely contributed to older adults' lower ability to differentiate between strong and weak central cues (Hypothesis 3). Those who were less cognitively declined performed more like younger adults, in which they were more likely to actively engage in deliberating on the credibility of the messages they read; 2) the generally lower Internet experience of older adults could explain their lower ability to differentiate between strong and weak peripheral cues (Hypothesis 4), as older

adults with high Internet experience could perform as well as younger adults in differentiating between strong and weak peripheral cues; and 3) health domain knowledge could compensate for older adults' lower abilities in making credibility judgments (Hypothesis 5), as those who had better health knowledge tended to spend less cognitive resource on reading the information and were less likely to be misled by what was claimed in the message. Also they pay more attention to peripheral features that were specific to health information, which could compensate for older adults' general lower abilities to recognize credibility related Website features. I will discuss the implications from this experiment in the discussion part.

CHAPTER 5 EXPERIMENT 2

In the second study, I focused on how the presence of user reviews as an additional component to the health website could influence younger and older adults' credibility judgments, and how the consistency of credibility cues in user reviews with the credibility cues in content message induced changes to their credibility judgment.

Participants

Twenty-two older adults (age between 58 and 80, Mean=68.45, SD=6.36, 59.1% are female) and twenty-two younger adults (age between 19 and 26, Mean=21.50, SD=1.95, 63.6% are female) participated in the study. All participants were recruited from the Champaign-Urbana area in Illinois, USA. Most participants (93.2%) have completed some years of college. There was no significant age difference in education level or self reported frequency of health information seeking activities on Internet. Also, there was no significant age difference in health domain knowledge ($p=0.64$), as reflected by the same fluency task used in experiment 1.

Experiment Design and Material

A $2 \times 2 \times 2 \times 2$ mixed factor design was used in this study. There were two within-subjects variables: content cues (strong/weak) and contextual cues (strong/weak), and two between-subject variables: age (young/old), and user reviews (with/without user reviews). All participants were asked to finish the same task as in experiment 1. The 32 Web pages, which were combinations of weak/strong content and contextual cues, were the same as those used in experiment 1 (see Table 2). Then I added user reviews to each of the page to create the condition with user reviews. User reviews were randomly assigned to be consistent or inconsistent with the content cues, and evenly distributed across all content cue and contextual cue combinations.

User Reviews

I selected material of user reviews from the same website where I selected content articles (revolutionhealth.com) and modified them to fit the particular medicine. User reviews on the site were given anonymously together with ratings on a one star (not recommended) to five star (highly recommended) scale. The ratings were generally consistent with how negative or positive user reviews were used (see Table 2). These user reviews were primarily about users' experience with the medication, such as its efficacy and side effects, if any. Each entry of user review has less than 100 words. 4-6 entries of user reviews were given to each medication. I manipulate the consistency of user review with content cue strength by selecting positive (three to five stars) or negative (one to three stars) reviews.

	Strong	Weak
Content cue	chosen from “high ranked” medicine, with research evidence, explanation of treating mechanism, comprehensive and accurate information, professional writing	chosen from “low ranked” medicine, lack of evidence, biased information, ill logic, commercial or unprofessional writing tone
Contextual cue	nice layout/color/information structure, with reference/contact information/third party endorsement	bad design, lack of reference/source, typo, advertising, commercial features (e.g. donate button)
User review	<p>Five-star review: <i>“Really effective product!. Combined with the right diet and exercise this meds is capable of producing rigid control of blood sugars in a very tight range.”</i></p> <p>Four-star review: <i>“It works for me. I have been on it for about 5 weeks now. My blood pressure is now high/normal. For the first 4 weeks, this medication caused moderate fatigue but this side effect has now disappeared.”</i></p>	<p>One-star review: <i>“This was a complete waste of time and money for me. I tried different brands one after the other and never even lost half a kg!”</i></p> <p>Two-star review: <i>“Didn't see much difference. I used this remedy for a few months; it did help a little with sugar cravings but I probably loss only a couple of pounds. It made my mouth dry, and I was irritable”</i></p>

Table 2. Examples of content cue, contextual cue and user review manipulation

Procedure

Before the task, participants were asked to take the standardized test for cognitive abilities and Internet experience similar to those used in experiment 1. Participants were then randomly assigned to conditions with or without user reviews. Similar as experiment 1, participants started

by reading the task scenario and then finished the 8 disease tasks. For condition with user reviews, participants could click on a “Read Users’ Review” link to read user reviews, which were presented on the same web page. After that participants clicked on the “Rate” button on the aggregator interface to submit their ratings for the medication.

RESULTS

The analysis was divided into three parts. First I investigated the effects of presence of user reviews on younger and older adults’ credibility with all pages. Second I focused on only the half of pages with user reviews that were consistent with the credibility of message contents to understand the effects of consistent user reviews (Hypothesis 6). Then I investigated the results with the other half of pages to understand the effects of inconsistent user reviews (Hypothesis 7). A series of follow up analysis was also conducted to understand the effects of Internet experience on older and younger adults’ susceptibility to inconsistent user reviews (Hypothesis 8).

Effects of User Review on Credibility Judgment

To understand how the presence of user reviews, which had an equal chance of being consistent or inconsistent to the credibility of content in the experiment, influenced younger and older adults’ credibility judgment, I performed a four-way ANOVA with age and presence of user reviews as between subjects variables, and content cue and contextual cue strength as between subjects variables. The results showed the main effects of content cue ($F(1,40)=56.66, p < .01$) and contextual cue ($F(1,40)=23.05, p < 0.01$) were significant. The interactions between content cue and age ($F(1, 40)=3.42, p=0.07$), interaction between contextual and age ($F(1, 40)=3.42, p=0.07$) and interaction between content cue and presence of user reviews ($F(1,40)=3.79,$

$p=0.06$) were marginally significant. Interestingly, there was a significant three-way interaction between content cue, age and presence of user reviews ($F(1,40)=4.45, p=0.04$).

Figure 7 explained the three-way interaction between content cue, age and presence of user reviews: while the presence of user reviews significantly reduced younger adults' ability to differentiate between strong and weak content cues, this effect was not observed in older adults.

Three-way ANOVA with content cue, contextual cue and presence of user review was performed within each age group to further verify this conclusion. Results showed that while there was significant interaction between content cue and presence of user reviews among younger adults ($F(1, 20)=5.903, p=0.02$), this two-way interaction was not observed among older adults ($F(1,20)=0.022, p=0.88$). It indicated that while the presence of mixed user review significantly affected younger adults' credibility ratings after reading them, it did not show such effect on older adults' credibility ratings.

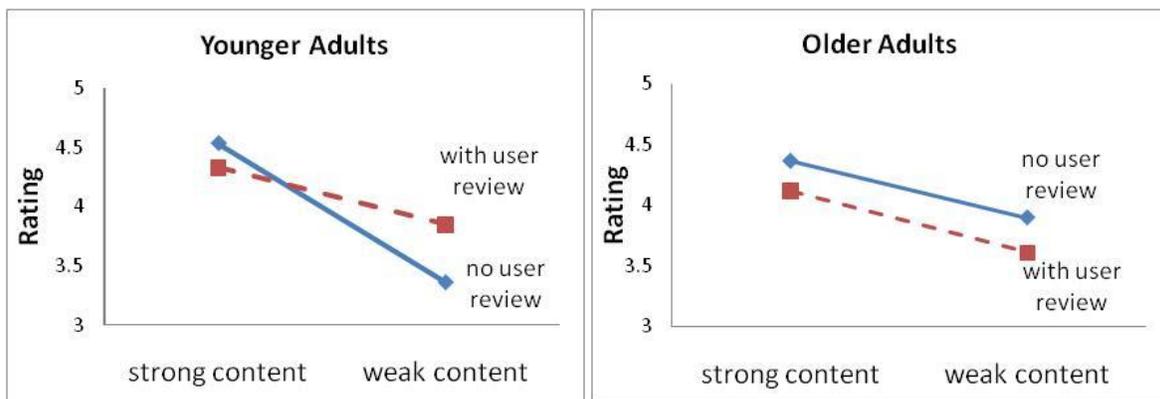


Figure 7. Credibility rating for pages of strong/weak content cues with or without user reviews

I then investigated only the younger group and older group with user reviews by introducing another independent variable: the consistency with content cue strength (consistent/ inconsistent), as a within-subject variable. Four-way ANOVA with content cue, contextual cue, age and consistency of user reviews showed that the main effects of content cues ($F(1,20)=13.27, p<0.01$) and contextual cues ($F(1,20)=8.03, p=0.01$) were still significant. Two way interaction between

content cue strength and consistency of user reviews was significant ($F(1,20)=43.86, p<0.01$), and the two-way interaction between consistency and age was significant ($F(1,20)=4.30, p=0.05$). However, the interaction between content cue and age became non-significant when there were user reviews ($F(1,20)=0.03, p=0.87$).

These results indicated that while older adults were less sensitive to the credibility cues in Web page content, as shown in experiment 1, this age difference disappeared when user reviews were presented. This may happen primarily because mixed user reviews tended to moderate younger adults' credibility judgment with content cues, but less so for older adults. The interaction between consistency and content cue, and interaction between consistency and age suggested that user reviews that were consistent and inconsistent with content cue strength had different impact on older and younger users' credibility judgment. To further understand this effect, I divided all the 32 Web pages into two categories: pages with user reviews consistent with content cue credibility, and those inconsistent with content cue credibility, and analyzed them separately.

Effects of Consistent User Reviews

I analyzed the 16 Web pages with user reviews that are consistent with the content cue credibility. Four-way ANOVA with content cue, contextual cue, age and presence of consistent user reviews showed that the main effects of content cue ($F(1,40)=102.00, p<.01$) and contextual cue ($F(1,40)=14.27, p<.01$) were significant. The interactions between content cue and age ($F(1,40)=4.87, p<.01$), and interaction between content cue and presence of consistent user reviews ($F(1,40)=7.34, p=0.01$) were also significant.

The results indicated that the presence of consistent user reviews had positive effects on users' ability to differentiate between Web content of high credibility and those of low credibility, possibly because they provided additional consistent cues to reinforce participants' initial attitude formed by reading the article. I then performed a three-way ANOVA within each age group. While there was a significant two-way interaction between content cue and consistent user reviews among older adults ($F(1,20)=6.80, p=0.02$), this two-way interaction was not significant among younger adults ($F(1, 20)=1.66, p=0.21$). Figure 8 illustrated this difference: while consistent user reviews did not significantly benefit younger adults' credibility judgment with content cues, it enhanced older adults' ability to differentiate messages with strong and those with weak content cues.

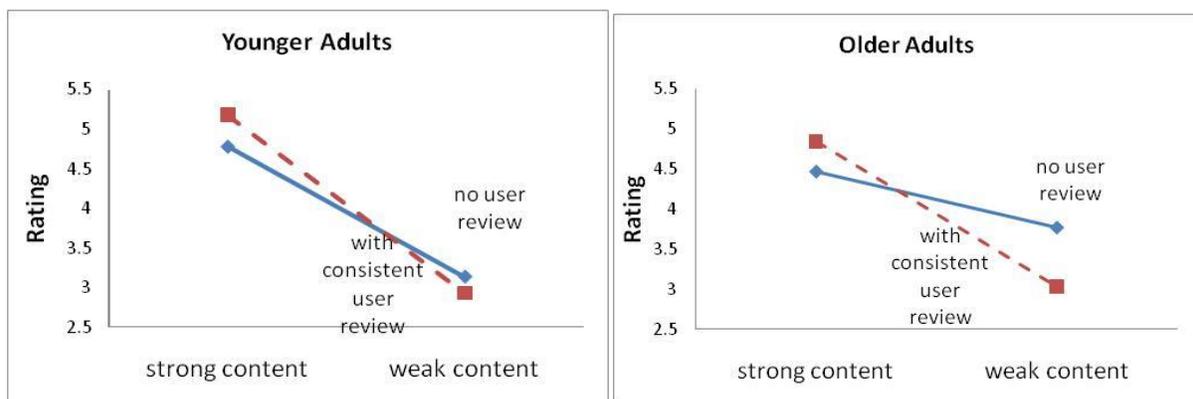


Figure 8. Credibility rating for pages of strong/weak content cues with or without consistent user reviews

In summary, I found that consistent user reviews i.e., favorable user reviews given to credible medication information, and unfavorable user reviews given to non-credible one, enhanced older adults' differential reaction to credible medication information and non-credible one more significantly than younger adults. It confirmed the hypothesis 6 that older adults could benefit from user reviews that may confirm their initial credibility judgment with the original content.

Effects of Inconsistent User Reviews

I then analyzed the 16 Web pages with user reviews that are inconsistent with the content cue strength. Four-way ANOVA showed that the main effect of contextual cue ($F(1,40) = 6.62, p = 0.01$) was still significant, the main effect of content cue is only marginally significant ($F(1,40) = 3.30, p = 0.08$). The interaction between age and contextual cue was marginally significant ($F(1,40) = 3.14, p = 0.08$), and interaction between content cue and presence of inconsistent user reviews was significant ($F(1,40) = 24.72, p < .01$). I found that the interaction between content cue and age was not significant ($F(1,40) = 0.03, p = 0.87$) when inconsistent user reviews were presented. Interestingly, there was a marginally significant three-way interaction between content cue, age and presence of inconsistent user reviews ($F(1,40) = 2.43, p = 0.10$).

From Figure 9, one could see that the marginally significant three-way interaction among age, content cue strength, and presence of inconsistent user reviews was caused by younger adults' higher ratings when inconsistent user reviews were presented than those by older adults. It also led to the disappearance of the age difference in making credibility judgment based on content cues, as shown by the non-significance of two-way interaction between age and content cue. The two-way interaction between inconsistent user review and central cue strength implied that inconsistent user reviews in general had a negative influence on users' ability to differentiate between high credibility and low credibility cues in the Web page content. Alternative three-way ANOVA in each age group showed significant interaction between inconsistent user review and content cues for both younger adults ($F(1,20) = 16.17, p < .01$) and older adults ($F(1,20) = 8.55, p < .01$), which further confirmed this conclusion.

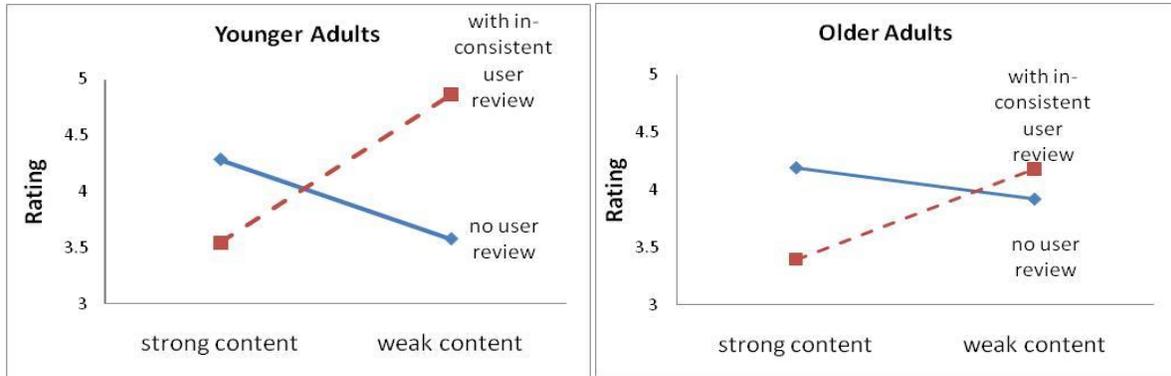


Figure 9. Credibility rating for pages of strong/weak content cues with or without inconsistent user reviews

The effect of inconsistent user review is particularly interesting since negative reviews on well argued articles and positive reviews on poorly argued articles may have distinctive meanings for what we suggested as the social means to health information credibility control. While negative reviews on a well written (strong content cue) article could potentially help prevent users from mistrusting questionable sources, positive reviews on a poorly written (weak content cue) article, on the contrary, could imply deceptive manipulations or spamming related activities.

To understand the effects of negatively inconsistent and positively inconsistent user reviews on older and younger adults' credibility judgment, I analyzed the main effects of the presence of negatively inconsistent user reviews on ratings for high credibility content, and the main effects of positively inconsistent user review on ratings for low credibility content in each age group. It showed that, for older adults, the effect of negatively inconsistent user reviews on strong content cues was marginally significant ($F(1,20)=3.17, p=0.09$), but the effect of positively inconsistent user reviews on weak content cues was non-significant ($F(1,20)=0.32, p=0.58$). For younger adults the effect of positive user reviews on low credibility content was significant ($F(1, 20)=10.96, p<0.01$), and negative user reviews on high credibility content was marginally significant ($F(1,20)=3.45, p=0.08$). Consistent with hypothesis 7, these results suggested that older adults were less susceptible to the influence of inconsistent user reviews. Furthermore,

older adults were especially less swayed by the combination of ill-made arguments and highly appraising user reviews. In that case, they tended to retain their initial negative attitude rather than change their ratings.

Effects of the discrepancy of inconsistent user reviews on change of ratings

In persuasion studies, two key variables were often investigated: message discrepancy and attitude change. Message discrepancy refers to the difference between the belief position advocated by a message (P_M) and the recipient's initial belief position (P_0 .) Post-treatment measure of the recipients' belief position was taken as the new belief position (P_{Eq} .) which will remain until new external forces cause it to move. The amount of belief change y_{Eq} can be calculated as $P_{Eq}-P_0$. The Distance-Proportional Model (Chung, 2008) proposed that the amount of belief change is a linear function of message discrepancy: $y_{Eq} = \beta(P_M - P_0)$, where β is called the proportional factor.

In the experiment, participants were first presented with the article page, and they had to click on the "Read Users' Reviews" link to read the user reviews. By comparing the ratings without reading the user reviews to those after they read the user reviews, I could estimate the extent to which user reviews induced changes to the initial attitude to the credibility of the article. In addition to the absolute consistency of user reviews with central argument, I was interested in how users' credibility judgments were influenced by the varied argument strength in user reviews. I explored the relation between message discrepancy and attitude change by referring to the Distance-Proportional Model as stated above. This model has been shown to work well when there were no negative cues about the reliability of communicator who gave this message [Error!

Reference source not found.] Since in this experiment, all user reviews were shown in the same style anonymously, I assumed there was no difference of source credibility among user reviews.

As I randomly picked up user reviews from a real website, the user rating shown beside each entry of review (one star to five star) was consistent with how strong the advocated belief was.

We calculated the average value of user ratings as the argument strength of user reviews for that particular medicine (P_M). The average rating of each web page given by the control group without user reviews was then used to estimate users' initial attitude (P_0). Younger and older adults' initial attitude was calculated separately. Then the average ratings given by the group with user reviews were used to estimate the final attitude towards each medicine (P_{Eq}). The index of message discrepancy was calculated by the difference between normalized user ratings and normalized initial ratings, while the index of attitude change was calculated by the difference between the final attitude and initial attitude (y_{Eq}). Then I performed regression analyses with these two indexes for each of the 32 medicines among older and younger adults.

Figure 10 showed the regression results of older adults and younger adults. The results indicated that in general both younger and older adults' attitude change fits the Distance Proportional Model well. It showed that the discrepancy of belief position between user reviews and participant's initial belief after reading the article induced attitude change to participant's final credibility judgments. It showed a higher r-square value for younger adults ($R^2=0.67$) than older adults ($R^2=0.54$), and a larger regression coefficient for younger adult ($\beta_1=0.57$) than older adults ($\beta_1=0.39$). Comparison of the regression coefficients of younger and older adults showed that the difference was marginally significant ($t=1.82, p=0.07$). Results provided support to the notion that younger adults' attitude change were more closely related to message discrepancy

than older adults, suggesting that younger adults might have a tendency to be more influenced by user reviews than older adults (Hypothesis 7).

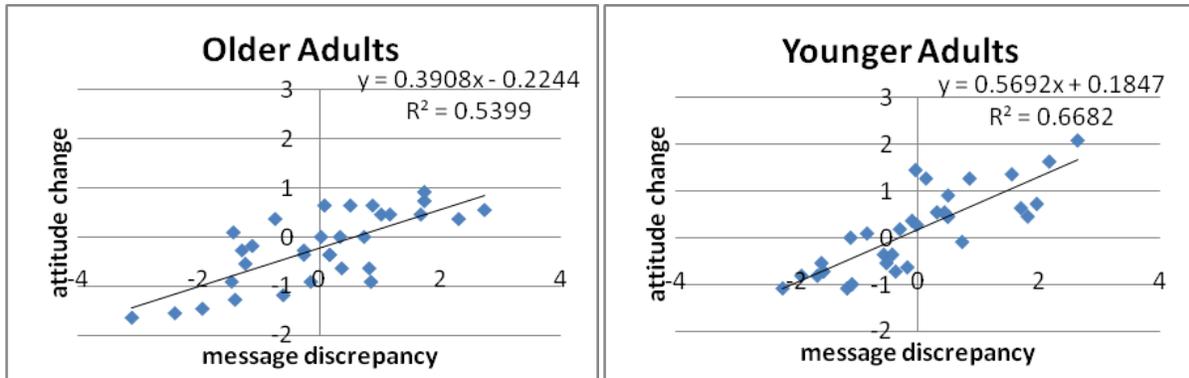


Figure 10. Regression between attitude change and message discrepancy

Effects of Internet experience on subjectivity to inconsistent user reviews

The Internet Experience index I measured showed the older participants in the study had marginally less experience with Internet ($p=0.10$). A number of previous studies provided evidence that people who are actively involved in online environment tend to have higher general trust for information on the Internet (e.g., Dutton, 2006). I hypothesized that people with more Internet experience, who are also more likely to be familiar with Web 2.0 technology, may react differently towards user reviews. Therefore I conducted a regression analysis to investigate the effects of Internet experience on the subjectivity to inconsistent user review.

First I created a variable called content credibility difference (CCD) by averaging the difference between credibility rating given to Web page with high credibility contents and those with low ones. Then I created a dichotomous variable for user review (UR). It gave a value of -1 to the participant who performed task without user reviews, and 1 to those who performed task with user reviews. Internet Experience index (IE) was used as another independent variable for the regression. In addition I tested their interaction by creating the term by multiplying IE and UR.

The linear regression analysis on CCD with IE, UR, IE*UR showed the effect of IE*UR is significant ($p=0.04$). This significant variable indicated that the effect of user reviews on the differentiation of content credibility was moderated by Internet experience. I plotted the content credibility difference value by Internet experience for group with user review and group without user review in Figure 11. It illustrated that people who have higher Internet experience were in general more influenced by inconsistent user review, as the difference of CCD between pages with user review and pages without user review was larger for participants who had higher Internet experience.

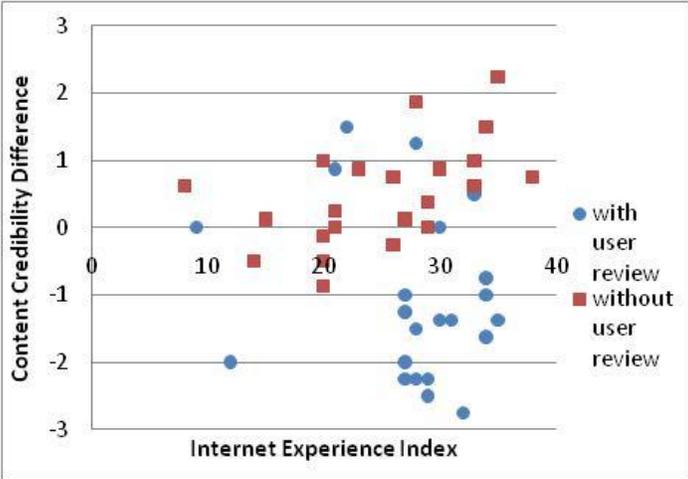


Figure 11. Plot of content credibility difference (CCD) with Internet experience (IE) index for group with and without user reviews

This result confirmed the hypothesis that Internet experience is related to the change in ratings induced by inconsistent user reviews. People who have higher Internet experience seemed to be more susceptible to the influence of user reviews. The reason may be that they are more likely to have experience with user review system or other user generated content, and therefore tend to give more credits to information or cues from user reviews. Considering older generation is in general less involved in Web environment and often only participate in limited activities (Pew

Internet Report, 2010), it is, therefore, understandable that I observed that older adults were less affected by user reviews that convey inconsistent cues with the original content.

To analyze this effect of Internet experience for younger and older adults, I divided each age group into high IE group and low IE group by performing median split with IE index. I then performed three-way ANOVA with content cue, contextual cue and presence of inconsistent user reviews in each group. It was found that, for older adults, in the high IE group, I did not observe significant interaction between content cue credibility and presence of user experience ($F(1,10)=1.84, p=0.21$), while in the low IE group, this interaction was significant ($F(1,10)=9.56, p=0.01$). Similarly, for younger adults, this interaction was non-significant for low IE group ($F(1,10)=2.01, p=0.19$), but significant for high IE group ($F(1,10)=13.41, p<0.01$). This result indicated that for both younger and older adults, those with higher Internet experience were more subject to the influence of inconsistent user review.

To further detect whether older and younger adults with low or high Internet experience react to inconsistent user review in the same pattern, I analyzed the main effects of presence of negatively inconsistent user reviews on ratings of high credibility content cues, and the presence of positively inconsistent user reviews on ratings of weak content cues for older and younger adults with low or high Internet experience. I found that older adults with low Internet experience were not affected by either positive reviews on page with weak content cues ($F(1, 10)=0.04, p=0.85$) or negative reviews on page with strong content cues ($F(1, 10)=1.18, p=0.30$). Older adults with high Internet experience, as well, were not affected by either positive reviews on page with weak content cues ($F(1, 10)=1.35, p=0.27$) or negative reviews on page with strong content cues ($F(1, 10)=0.68, p=0.43$). Younger adults who had low Internet

experience were only subject to the influence of positive user reviews on weak content cues ($F(1, 10)=8.37, p=0.02$) but not negative review on strong content cues ($F(1, 10)=0.18, p=0.68$).

Younger adults with high Internet experience, however, were significantly influenced by both negative user review on strong content cues ($F(1, 10)=8.78, p=0.01$) and positive user review on weak content cues ($F(1, 10)=5.05, p=0.05$).

The pattern of results implied that, in general, users with poorer Internet experience were less inclined to integrate cues and information from user reviews when making credibility judgment (Hypothesis 8). However, the generally poorer Internet experience of older adults could only partially explain the age difference on the effects of inconsistent user reviews. Additionally, there appeared to be age difference on the tendency to discount positive or negative user reviews. Specifically, I found that older adults would be more likely to discount positive reviews in articles with weak content cues than younger adults.

SUMMARY OF RESULTS FROM EXPERIMENT 2

In general, user reviews were found to have strong impact on users' credibility judgments. I found significant effects of central cues, peripheral cues, and presence of user reviews on users' credibility judgment rating. Also, I found that user reviews could influence older and younger adults' credibility judgment in different ways. When credibility cues in user reviews were consistent with the credibility cues in website content, it could more significantly facilitate older adults to make correct judgment about content cues (Hypothesis 6), and thus help to overcome the age difference in the ability of making correct credibility judgment by supplementing the generally lower systematic/deliberative processing abilities of older adults.

The most interesting finding in this study was the effect of user reviews that convey information inconsistent to the credibility of original content. It showed that older adults were less subject to the influence of inconsistent user review as compared to younger adults (Hypothesis 7).

Especially when appraising user reviews appeared on a flawed website, older adults tended to discount information from user reviews while younger adults were easily swayed by the user review. Also I found that participants' attitude change was related to the discrepancy between the position reflected in user reviews and their initial attitude towards the website content.

To understand the cause of such age difference, I explored the effect of Internet experience on subjectivity to inconsistent user review and found that people who were inexperienced with Web environment tend to be less influenced by user reviews (Hypothesis 8). Therefore, older adults' generally lower Internet experience seem to at least partially explain their lower subjectivity to the influence of inconsistent user reviews, although there might be additional age differences in terms of the pattern of credibility judgment change by user review. E.g., younger adults with low Internet experience seemed to be still subject to the influence of positive user reviews on non-credible content, but not that of negative user reviews on highly credible content, while older adults with low Internet experience were influence by neither of them.

CHAPTER 6 CONCLUSIONS

The results of the two experiments showed that, when making credibility judgments of online health information, people integrate credibility cues from different sources including message content, contextual website features and user reviews. However, there exists age difference in all these three aspects: 1) for content cues, older adults were less able to differentiate content with high credibility from those with low credibility. Protocol analysis revealed that older adults tended to accept what was claimed by the article without adequate deliberation on the quality of information. This can be problematic as they may be easily misled by unreliable information. On the other hand, I found that, particularly for older adults, the ability of making correct credibility judgment with content cues were associated with the person's cognitive ability. Those older adults who were more cognitively declined had more trouble judging the credibility of message content. However, by gaining more health related domain knowledge, older adults could possibly compensate for their cognitive declines to maintain the ability of making correct credibility judgment with the message content. 2) for contextual cues, older adults were less sensitive to the credibility indicated by Website features. Protocol analysis seemed to imply that older adults were more inclined to skip evaluating the contextual Website features, but simply focus on reading the content. This tendency could be moderated by increasing Internet experience. Those older adults who were less familiar with Internet were especially vulnerable to the risk of ignoring credibility cues in Website contextual features. On the other hand, those older adults who have better health related domain knowledge seemed to be able to perform better in credibility judgment with contextual cues. 3) For user reviews, older adults exhibited different reaction to user reviews that were consistent to and inconsistent to the credibility of original message content (content cues). The consistent user reviews seemed to better support older

adults' credibility judgment of the medication by providing additional consistent cues to reinforce their initial judgment with the content. Thus it helps to moderate the age difference in the ability of making correct credibility judgment. The inconsistent user reviews, however, had less impact on older adults' credibility judgment as compared to younger adults. Especially when positive user reviews were given to a low credibility medication, older adults' credibility judgment was less swayed. One important factor contributing to this age difference is older adults' generally inexperience with Internet, as I found that people with low Internet experience were less likely to be influenced by inconsistent user reviews.

CHAPTER 7 GENERAL DISCUSSION

The results confirmed that there are age differences in credibility judgment of online health information. Based on these results, implications for improving older adults' credibility judgment of online health information could be drawn. From a training perspective, I identified two credibility judgment strategies that older adults should be encouraged to adopt: 1) actively deliberate on the credibility of the message rather than passively read facts stated on the Web page 2) initiate credibility judging process by examining contextual cues first, as it takes less cognitive effort and could be more easily adopted by older adults. The results seemed to suggest that if we could provide instructions or training to older adults with these strategies, their credibility judgment outcomes could be improved.

From a design perspective, one could consider methods to compensate for older adults' declined cognitive ability and generally poorer Internet knowledge to facilitate their credibility judgment. As older adults are less likely to spontaneously deliberate on the credibility dimension of information, it would be helpful to provide explicit, prominent guidance to encourage older adults to deliberate on the information. Similarly, designers should take into account older adults' generally limited Internet experience, as reflected by their tendency to pay limited attention to web-specific contextual cues in their credibility judgments. For example, while collaborative implementations to provide authority/ credibility indicators (e.g. HON stamp and other displayed awards) have been considered promising solutions for regulating health information website, the study (as well as others) seemed to suggest that users often failed to actively check these features. One way to ameliorate the problem is to utilize external tools such as the use of pop-up checklist that "reminds" users to ask critical questions relevant to credibility of information and to pay attention to Web-specific features, which may help older adults to naturally adopt more effective

credibility judgment when consuming online health information. To successfully guide users' attention towards credibility related features, and relieve users from excessively deliberating on the quality of information, visualization techniques seem to provide useful tools to enforce users' credibility judgment in a systematic and user-friendly way (Yamamoto, 2011; Schwarz, 2011).

Another focus of the study is the impact of user reviews users' credibility judgment. As stories about how people benefit from other patients' experience shared online abound, the impact of "user reviews" on health information website is likely to be more than just another tool for people to socialize. The study showed that user reviews could have complex influence on older and younger adults' credibility judgments. It suggested that knowing exactly how individuals with different background profiles would be influenced by user reviews is important in understanding the long-term effects of these policies.

As I found that older adults' lower sensitivity to content credibility could be moderated by providing user reviews that are consistent with the credibility of original content, it implied that a well designed user review system, which encourage users to give informative and accurate reviews, may benefit older adults' credibility judgment. For example, customizable sorting function to provide user reviews that users concern most could efficiently help users to make informed credibility judgment. Also, to highlight the linkage between content and user reviews in some way (e.g., providing hyperlinks between a claim and user reviews that confirm or disconfirm it) may be helpful for users who look for evidence or counterevidence for their initial credibility judgment with the original content.

One most interesting finding in this study was older adults' lower sensitivity to inconsistent user reviews, especially when positive user reviews appeared on a flawed website. This phenomenon

could be interpreted from two aspects: First, older adults' lower susceptibility to attitude change may be a possible explanation, as reflected by the smaller regression coefficients in attitude change versus message discrepancy in the analysis. As suggested by previous research, older adults are more resistant to attitude change when facing persuasive influence, in which detailed argument is provided to convince message recipient to adopt the position advocated. Also, research on age difference in dual processing with external influence cues indicated that motivational and emotional variables could vary older adults' deliberative processing level and affect the outcome of influence (Peters, 2007). I infer that these age-related differences may cause older adults to be less sensitive to the influence of user reviews, especially in the situation of negative initial attitude. It was possible that older adults' initial negative attitude towards low-credibility content made them to selectively stop further deliberative processing when reading and comprehending user reviews.

The second possible reason for the lower sensitivity to inconsistent user reviews could be attributed to the generally lower internet experience of older adults. The non-significant age difference among younger and older users who had low Internet experience provided some support to this possibility. Indeed, previous studies showed that people's general trust with Internet is positively related to users' Internet experience (Dutton, 2006). Frequent users of Internet tend to have more certainty and more confidence in online information. We could further extend this view to social networking applications. For example, there is research showing that younger adults often use Internet for entertainment and social networking, while older adults tend to use Internet as a tool for research, shopping and banking (Sydney, 2009). This may imply that, compared to general Internet experience, older adults may have an even lower experience with user reviews and other social networking features (both in terms of actively

contributing and passively reading). This may lead to a higher tendency for older adults to distrust and discount cues associated with user reviews.

We should state that older adults' relatively insensitive reaction to positive user reviews on a low credibility site, and to negative user reviews on a seemingly credible site could be interpreted as a positive reaction. In fact, one could argue that younger adults might be more susceptible to biased information than older adults. While in the study I did not deliberately test this notion, future research could be directed to how easily younger and older adults can detect biased or even wrong information in user reviews.

On the other hand, to make user review system a helpful tool for quality control of online health information, designers should give more thoughts on how to motivate older adults to utilize such system. Good user experience, both in terms of the interaction with the system and the quality of user reviews it provides, is key to engage older adults, who may be new to Web 2.0 technologies, in further involvement and higher trust with user reviews. Further research is needed to understand how different policies adopted by user review systems, such as those that limit the minimum length of review and the number of entries that one user could give daily, and collaborative evaluation of reviews themselves, such as the "like" button on a review, will provide better solutions to harness the potential of user reviews for information quality control, and to benefit older adults.

REFERENCES

- Alsudani, F. and M. Casey., The effect of aesthetics on web credibility. In *Proc. 23rd British HCI Group Annual Conference on People and Computers* (2009), 512-519.
- Beier, M.E., Ackerman, P., Age, Ability and the Role of Prior Knowledge on the Acquisition of New Domain Knowledge: Promising Results in a Real-World Learning Environment. *Psychology and Aging*, 20, 2 (2005), 341-355.
- Cheung, M., et al., Credibility of Electronic Word-of-Mouth: Informational and Normative Determinants of On-line Consumer Recommendations. *Int. J. Electron. Commerce* 13, 4 (2009), 9-38.
- Chung, S., E.L. Fink, and S.A. Kaplowitz, The Comparative Statics and Dynamics of Beliefs: The Effect of Message Discrepancy and Source Credibility. *Communication Monographs* 2, (2008), 158 - 189
- Chin, J. and W.-T. Fu., Interactive effects of age and interface differences on search strategies and performance. In *Pro. CHI2010*. ACM(2010), 403-412.
- Cialdini R.B. & Goldstein N. Social Influence: Compliance and Conformity. *Annual Review of Psychology* 55(2004), 591-621
- Czaja, S.J., et al., Older Adults and Internet Health Information Seeking. In *Proc. HFES2009*. 126-130.
- Chung, S., E.L. Fink, and S.A. Kaplowitz, The Comparative Statics and Dynamics of Beliefs: The Effect of Message Discrepancy and Source Credibility. *Communication Monographs* 2, (2008), 158 - 189
- Dutta-Bergman, M.J., The Impact of Completeness and Web Use Motivation on the Credibility of e-Health Information. *Journal of Communication* 54. 2 (2004), 253-269.
- Dutton, W.H. and A. Shepherd, Trust in the Internet as an experience technology. *Information, Communication & Society* 9, 4 (2006), 433 - 451.
- Ericsson, K. A., Simon, H., *Protocol Analysis*(1984), Cambridge, MA: MIT Press
- Eysenbach, G., Credibility of Health Information and Digital Media: New Perspectives and Implications for Youth. The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning (2007), 123-154.
- Eysenbach, G., et al., Empirical Studies Assessing the Quality of Health Information for Consumers on the World Wide Web: A Systematic Review. *JAMA* 287. 20 (2002), 2691-2700.
- Eysenbach, G., E.R. Sa., and T.L. Diepgen., Shopping around the internet today and tomorrow: towards the millennium of cybermedicine. *British Medical Journal* 319. 1294 (1999), 1-5.
- Fairweather, P.G. How older and younger adults differ in their approach to problem solving on a complex website. In *Proc. SIGACCESS2008*. ACM Press(2008), 67-72

- Ferebee, S., An examination of the influence of involvement level of web site users on the perceived credibility of web sites. *In Proc. 2nd International Conference on Persuasive Technology*. Springer-Verlag(2007), 176-186
- Fogg, B.J., et al., What makes Web sites credible?: a report on a large quantitative study. *In Proc. CHI2001*. ACM(2001), 61-68
- Fogg, B.J., Prominence-Interpretation Theory: Explaining How People Assess Credibility Online. *In Proc. CHI 2003*. ACM Press(2003a),
- Fogg, B. J. (2003b). *Persuasive technology: Using computers to change what we think and do*. San Francisco: Morgan Kaufmann.
- Gefen, D. and D.W. Straub, Consumer trust in B2C e-Commerce and the importance of social presence. *Experiments in e-Products and e-Services* 32, 6 (2004), 407-424.
- Griffiths, L.T., M. Steyvers, and A. Firl., Google and the Mind: Predicting Fluency with PageRank. *Psychological Science* 18. 12 (2007), 1069-1076.
- Hamilton, M.A., Message Variables That Mediate and Moderate the Effect of Equivocal Language on Source Credibility. *Journal of Language and Social Psychology* 17,1 (1998), 109-143.
- Hanson, V.L., Cognition, Age, and Web Browsing. *In Proc. UAHCI2009*. Springer-Verlag(2009), 245-250
- Hargittai, E., Fullerton, L., Menchen-Trevino, E., Thmoas, K., Trust Online: Young Adults' Evaluation of Web Content. *International Journal of Communication* 4 (2010), 468-494.
- Hayslip B & Sterns, Age Differences in Relationships Between Crystallized and Fluid Intelligences and Problem Solving. *Journal of Gerontology* 34, 3(1979). 404-414
- Heinke Kunst, D.G., Pallavi M Latthe, Manish Latthe, Khalid S Khan, Accuracy of information on apparently credible websites: survey of five common health topics. *British Medical Journal* 324 (2002), 581-582.
- Hilligoss, B. and S.Y. Rieh, Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. *Information Processing & Management* 44. 4 (2008), 1467-1484.
- Hong, T., The influence of structural and message features on Web site credibility. *Journal of the American Society for Information Science and Technology* 57. 1 (2006), 114-127.
- Jones, L.W., R.C. Sinclair, and K.S. Courneya, The Effects of Source Credibility and Message Framing on Exercise Intentions, Behaviors, and Attitudes: An Integration of the Elaboration Likelihood Model and Prospect Theory1. *Journal of Applied Social Psychology* 33. 1 (2003), 179-196.
- Liao, Q.V. Effects of cognitive aging on credibility assessment of online health information. *Ext. Abstract CHI2010*. ACM Press(2010), 4321-4326

- Lynn W. Phillips, B.S., Age Differences in Information Processing: A Perspective on the Aged Consumer. *Journal of Marketing Research* 14, (1977), 444-457.
- Man Yee Cheung, C.L., Choon, Ling Sia, and Huaping Chen, Credibility of Electronic Word-of-Mouth: Informational and Normative Determinants of On-line Consumer Recommendations. *International Journal of Electronic Commerce* 13, 4 (2009), 9-38.
- Meola, M., Chucking the checklist: A contextual approach to teaching undergraduates Web-site evaluation. *Libraries and the Academy*, 4,3(2004) 331-344.
- Metzger, M.J., 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 (2007), 2078-2091.
- O'Grady, L., Future directions for depicting credibility in health care web sites. *International Journal of Medical Informatics* (2006), 58-65.
- Peters, E., et al., Adult Age Differences in Dual Information Processes: Implications for the Role of Affective and Deliberative Processes in Older Adults' Decision Making. *Perspectives on Psychological Science* 2. 1 (2007), 1-23.
- Petty, R.E., Cacioppo, J.T, The Elaboration Likelihood Model. *Advances in Experimental Social Psychology* 19. 123-205 (1986).
- Riegelsberger, J. and M.A. Sasse. Face it - photos don't make a web site trustworthy. *In CHI '02 extended abstracts*, ACM(2002), 742-743
- Rieh, S.Y., Credibility: A Multidisciplinary Framework. *Annual Review of Information Science and Technology* 41 (2007), 307-364.
- Sakamoto, Y. (2010). The impact of collective opinion on online judgment. *In Proc. 32nd Annual Conference of the Cognitive Science Society*.
- Salganik, M.J., P.S. Dodds, and D.J. Watts, Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science* 311. 5762 (2006), 854-856.
- Salthouse, T.A., Mediation of Adult Age Differences in Cognition by Reductions in Working Memory and Speed of Processing. *Psychological Science*, (1991), 2, 179- 183.
- Schwarz, J. & Morris, M., Augmenting Web Pages and Search Results to Support Credibility Assessment. *In Proc. CHI2011*. ACM(2011), 1245-1254.
- Shon, J., J. Marshall, and M.A. Musen. The Impact of Displayed Awards on the Credibility and Retention of Web Site Information. *In Proc. AMIA Symp.* 2000.794-798
- Sharit, J., et al., Investigating the Roles of Knowledge and Cognitive Abilities in Older Adult Information Seeking on the Web. *ACM Trans. Comput.-Hum. Interact* 15. 1 (2008), 1-25.
- Sillence, E., et al., A framework for understanding trust factors in web-based health advice. *International Journal of Human-Computer Studies* 64. 8 (2006), 697-713.

- Sillence, E., et al., How do patients evaluate and make use of online health information? *Social Science & Medicine* 64. 9 (2007), 1853-1862.
- Sundar, S. Shyam. "The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility." *Digital Media, Youth, and Credibility*. Cambridge, MA: The MIT Press, 2008. 73–100.
- Sydney, S.F., Generations Online in 2009, Pew Internet & American Life Project. 2009.
- Toma, C.L. Perceptions of trustworthiness online: the role of visual and textual information. *In Proc. CSCW2010*. ACM(2010), 13-22
- Tseng, S. and B.J. Fogg, Credibility and computing technology. *Commun* 42, 5 (1999), 39-44.
- Unsworth, N., R. Heitz, J. Schrock, and R. Engle, An Automated Version of the Operation Span Task. *Behavior Research Methods*, (2005), 37, 498-505.
- Walther, J.B., Z. Wang, and T. Loh, The Effect of Top-Level Domains and Advertisements on Health Web Site Credibility. *Journal of Medical Internet Research* 6. 3 (2004), e24.
- Wathen, C.N. and J. Burkell, Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology* 53. 2 (2002), 134-144.
- Yamamoto, Y. & Tanaka, K., Enhancing Credibility Judgment of Web Search Results. *In Proc. CHI2011*. ACM(2011), 1235-1244.
- Yoon, C., Cole, C. & Lee. M., Consumer Decision Making and Aging: Current Knowledge and Future Directions. *Journal of Consumer Psychology* 19(2009) 2-16