Open Access

Open Journal of Information Systems (OJIS)
Volume 5, Issue 1, 2018

www.ronpub.com/ojis
ISSN 2198-9281

Halo Effect Contamination in Assessments of Web Interface Design

Daniel S. Soper, Farnaz Piepkorn

Department of Information Systems & Decision Sciences, California State University,
Fullerton 800 N. State College Blvd., Fullerton, California, USA, {dsoper, fpiepkorn}@fullerton.edu

ABSTRACT

This paper relies on findings and theory from both the human-computer interaction and cognitive psychology literatures in order to inquire into the extent to which the halo effect contaminates web interface design assessments. As a human cognitive bias, the halo effect manifests itself when a judge's evaluations of an entity's individual characteristics are negatively or positively distorted by the judge's overall affect toward the entity being judged. These distortions and halo-induced delusions have substantial negative implications for rational decision-making and the ability to objectively evaluate businesses, technologies, or other humans, and should hence be a critical consideration for both managers and organizations alike. Here we inquire into the halo effect using a controlled, randomized experiment involving more than 1,200 research subjects. Subjects' preexisting affective states were activated using polarizing issues including abortion rights, immigration policy, and gun control laws. Subjects were then asked to evaluate specific interface characteristics of six different types of websites, the textual content of which either supported or contradicted their preexisting affective beliefs. Comparing subject responses to objective control evaluations revealed strong evidence of halo effect contamination in assessments of web interface design, particularly among men. In light of the results, a theoretical framework integrating elements from cognitive and evolutionary psychology is proposed to explain the origins and purpose of the halo effect.

TYPE OF PAPER AND KEYWORDS

Regular research paper: *halo effect, cognitive bias, user interface design, web design*

1 INTRODUCTION

Although technical research remains an important part of the contemporary information systems (IS) research agenda, the IS research community has clearly expressed an increased interest in cognitive and behavioral phenomena over the past several decades [2, 4, 18, 26, 49, 69, 81, 89, 95, 99]. In accordance with this general trend, IS researchers have conducted a large number of inquiries into human interactions with technology, including many studies in the areas of web design and usability [1, 25, 28, 33, 35, 52, 59, 68, 70, 88]. These two areas of inquiry are, of course, not mutually exclusive, and in this paper we contribute to their integration by inquiring into the extent to which a particular cognitive bias known as the *halo effect* produces inaccuracies and systematic errors in judgment in the context of web interface design assessments.

Almost everyone has experienced and is familiar with the halo effect. Like all cognitive biases, however, this familiarity exists primarily at an unconscious level. Put simply, the halo effect (or *halo error*) is the tendency to judge an entity's individual traits according to one's general feelings or affect toward the entity [12, 92]. If,

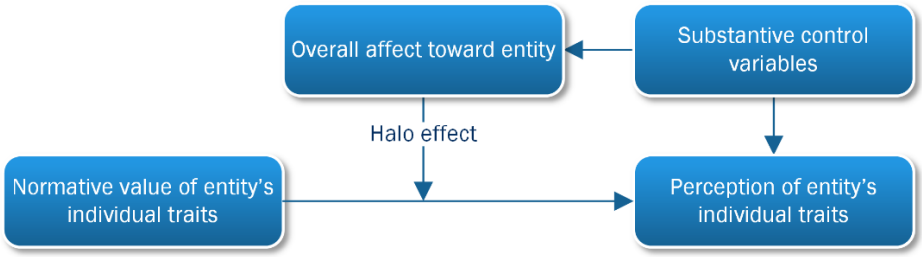

Figure 1. A theoretical representation of the halo effect

for example, you hold a colleague in particularly high regard, then you would be likely to assign high ratings to her individual characteristics (*e.g.*, intellect, trustworthiness, efficiency, cleanliness, manners, reliability, etc.), even if she may not objectively merit such high ratings. Although the halo effect clearly influences one's perceptions of other people, it has also been shown to distort the way in which we perceive businesses, educational institutions, government entities, consumer goods, and even certain technologies. The extent to which the halo effect colors assessments of interface design, however, is not currently well-understood. Given that websites now commonly serve as the most publicly visible face of an organization, this is not an insignificant or trivial oversight.

On the contrary, the ability of managers to accurately judge the quality of their organization's website has critical implications for the way in which both customers and the public at large will perceive the organization itself. In an effort to reconcile this oversight, the current study employs a controlled, randomized experiment involving 42 web interface design variations and 1,230 research subjects in order to assess the extent to which halo error distorts evaluations of web interface design. The study described herein substantially extends and refines our preliminary work on the topic [87], and proposes a new theoretical framework to explain the origins and evolutionary purpose of the halo effect.

Despite more than a century of research in other fields, the halo effect has – with a few notable exceptions – been largely ignored in IS research. The following section therefore seeks to familiarize IS researchers with the halo effect by providing a detailed overview of the history and development of halo effect theory. We then describe our research hypotheses and methodology, followed by a presentation of our analytical results and a discussion thereof. The manuscript concludes with a summary, limitations, and directions for future research.

2 THE HALO EFFECT

The halo effect is a cognitive bias in which one's overall affect toward an entity produces overtly positive or negative distortions in evaluations of the entity's individual characteristics. An *entity* in this characterization may be a person, thing, location, or even a non-material concept such as an idea. As a cognitive bias, the halo effect is one of several known phenomena that lead to systematic errors in human judgment. Together, such errors cause us to behave very differently from what might otherwise be expected under a normative model of rational decision-making. From a theoretical perspective, the halo effect can be readily represented using a structural model, such as that depicted in Figure 1.

As shown in the figure, the theory posits that the way in which an entity's individual traits are perceived depends not only on the objective or normative value of those traits, but also on the person's overall affect toward the entity. The final assessment of a specific individual trait, then, is a function of the normative value of the trait, the distortion introduced by the judge's preexisting attitude toward the entity, and random inter-judge variation, which may be explicable through substantive control variables. A positive (negative) broad impression can thus be expected to produce positively (negatively) biased estimates of individual traits. Inter-judge random variation notwithstanding, the magnitude of the halo effect can be quantified as the extent to which a judge's rating of an individual trait is influenced by his or her overall affect toward the entity, after controlling for the normative value of that trait. Although the theory was originally described in the cognitive psychology literature in the context of one person's judgments of another, research on the topic has since been extended into additional disciplines and realms of inquiry, notably for purposes of the current study including those of both business and technology.

2.1 History of the Halo Effect

It has now been more than a century since the halo effect first appeared in the scientific literature. In his 1907 manuscript on literary merit, psychologist Frederic Wells observed an unjustifiably strong correlation between critics' judgments of an author's general literary merit and the author's individual traits, which included qualities such as charm, imagination, and wholesomeness [98]. Wells could not possibly have

known at the time that this observation would inaugurate a stream of research that would endure for more than a century.

In accord with a common pattern in the history of science, Wells' findings did not immediately capture the attention of the scientific community. In fact, more than a decade would pass before the halo effect would again be mentioned in the literature, this time by psychologist Edward Thorndike, who is generally credited with giving the halo effect its name [92]. Five years earlier, Thorndike had noticed that the ratings assigned to corporate employees for traits such as intelligence, reliability, and technical skill were, to a suspicious extent, very highly and evenly correlated. Quoting Thorndike:

> *It consequently appeared probable that those giving the ratings were unable to analyze out these different aspects of the person's nature and achievement and rate each in independence of the others. Their ratings were apparently affected by a marked tendency to think of the person in general as rather good or rather inferior and to color the judgments of the qualities by this general feeling.* (p. 25)

Thorndike formally tested this supposition by evaluating ratings assigned to military officers by their superiors. As suspected, he discovered strong relationships among traits that should ostensibly have been nearly independent, an example being a correlation of 0.51 among officer intelligence and physique. He concluded that, "…even a very capable foreman, employer, teacher, or department head is unable to treat an individual as a compound of separate qualities and to assign a magnitude to each of these in independence of the others" (pp. 28-29). As his *ultima admonitio*, Thorndike warned that all future scientific studies employing multi-attribute rating scales must account for the impact of the halo effect. It is somewhat disturbing to consider that the vast majority of such research studies published during the past century – including those in the IS field – have not heeded this admonition, hence casting an ominous shadow over the validity and accuracy of their findings.

2.2 The Halo Effect in Human Beings

Some of the most visible and widely disseminated findings in this area of research have emerged from studies focusing on one person's judgments of another person. In their highly cited paper, Dion et al. [29] demonstrated that physically attractive men and women were presumed to have more socially desirable personalities, a higher occupational status, more marital competence, and more social and professional happiness than persons of average or below average attractiveness. The boundaries of this phenomenon were later extended by Landy and Sigall [48], who found that assessments of writing quality were distorted by the attractiveness of the author. When asked to evaluate the quality of an essay along several dimensions, skilled judges who were shown a photo of an attractive woman and told that she was the author consistently rated the essay more favorably than judges who were shown a photo of an unattractive author. The ratings of judges who were not shown a photo fell between these two extremes. A later replication and extension of this work found that the halo effect was largely isolated to the condition in which male judges were evaluating an essay written by an author whom they believed to be female [43], thus revealing the gender of the rater and ratee to be a potentially important factor in halo effect research. This notion was confirmed by Lucker et al. [53], who found that the magnitude of the halo effect varied according to whether the judge was male or female, and whether the person being judged was male or female.

In addition to studies focusing on physical characteristics, the halo effect has also been observed in other domains involving the evaluation of one person's traits by another. Early work by Remmers [75], for example, found strong evidence of the halo effect in students' evaluations of their instructors, with the effect being especially pronounced among university students. In a similar vein, students watching a video of a university instructor rated his accent, mannerisms, and appearance as appealing when he projected a warm and friendly persona, while the same traits were judged to be irritating and irksome when the instructor behaved in a cold and distant manner [65]. It has further been shown that the scores assigned by examiners for a person's performance on the most widely used intelligence tests are strongly contaminated by the halo effect when the examiner is provided *a priori* with an overall assessment of the person's level of intelligence [83].

All of these judgmental errors and departures from rationality carry manifold implications for both science and society at large, and it is for this reason that several authors have investigated whether the halo effect can be mitigated through the use of training procedures. In a cross-sectional study involving nurses, for example, it was found that those nurses who were trained about the nature of the halo effect exhibited less pronounced distortions in judgment than their untrained counterparts [16]. This finding was later replicated in the context of managerial assessments of subordinates [14], and in university students' evaluations of their instructors [9]. Despite these findings, the utility of training in the mitigation of the halo effect appears to be neither universal nor permanent. In a replication of the aforementioned Nisbett and Wilson [65] persona study, for example, it was found that subjects remained highly

susceptible to the halo effect even after receiving training regarding its consequences [101]. Indeed, after evaluating nine different methods aimed at reducing the halo effect, Cooper [23] concluded that each left a residual illusory halo, while work by Bernardin [8] revealed that improvements in halo error subsequent to training disappear rapidly over time.

Together, the studies reviewed above indicate that one's overall affect toward others exerts a powerful influence over assessments of their individual traits. These broad, overall impressions are formed from whatever information is available, be it years of personal experience, or, in the absence of all other information, something as simple as an impression of the person's physical characteristics. The halo effect clearly has a potent negative impact on one's ability to accurately and rationally judge the traits of others. It is a fundamental and deeply entrenched property of human cognition, and does not appear to be subject to permanent excision by any known means.

2.3 The Halo Effect in Business

A realization of the potential scope and scale of the negative consequences of the halo effect for business has engendered a strong and growing interest among business researchers, especially in recent decades. To date, the halo effect has been studied in a wide range of business contexts, including organizational reputation, management, investing, marketing, customer satisfaction, consumer research, and risk assessment. Some of the earliest business-related research into the halo effect was conducted in the area of management. When performing employment interviews, Bingham and Moore [13] noted that managers' assessments of interviewee traits were marred by halo effect contamination. Interestingly, the magnitude of the contamination was found to correlate with other factors, including the number of interviewers involved in the process. These empirical observations suggest that the halo effect is not insulated from the influence of other elements within the decision space, and that considering substantive control variables with a view toward isolating the halo effect is both prudent and judicious.

The halo effect has further been studied in the context of managers' assessments of their superiors, colleagues, and subordinates [31]. Using a moderation model, these researchers found the halo effect to be responsible for more than 10% of the variance in managers' ratings. Further evidence of halo effect contamination in the managerial domain was noted by Grove and Kerr [34] in their study of employee morale. When compared to a control group of employees from a financially sound firm, a group of employees from a firm in receivership reported very low values on a multi-item scale designed to measure job satisfaction. In accordance with halo effect theory, the distress felt by the employees of the bankrupt firm produced low ratings not only for dimensions such as job security, but also for other dimensions such as salary and working conditions – both of which were objectively superior to those of the employees in the financially sound firm. It is important to note that the distorted perceptions of a firm caused by the halo effect can be positive as well. Kauffman and Wang [44], for example, concluded that many of the positive assessments and subsequent investments made into individual e-commerce companies during the DotCom bubble were tainted by halo error arising from the positive overall impression of e-commerce during that era.

In addition to management and investment decisions, halo error has also been troubling in the field of marketing. In a study of viewer attitudes toward television shows, for example, Beckwith and Lehmann [6] found that the multi-attribute measurement models commonly used in marketing are strongly confounded by the halo effect. These authors warn that directly measured beliefs for different attributes of a product or service are often just reflections of an individual's overall affect toward that product or service. A later meta-study echoed this position, and concluded that halo effects contaminate a wide variety of marketing research settings [5]. This study also noted that most people – contrary to halo effect theory and more than a century of empirical evidence – still erroneously believe that their overall impression of a product or service emerges from a rational assessment of its individual traits, rather than *vice versa*. Further work reveals that halo-related problems with multi-attribute measurement models continue to plague marketing research, and that these problems can yield poor-quality decisions with respect to product modifications or strategy [50].

Customer satisfaction research has also been beleaguered by halo-based distortions in judgment. In the context of business-to-consumer (B2C) relationships, the halo effect has been shown to influence different attributes of customer satisfaction in the insurance industry [24], the banking services and airline industries [104], and when making travel reservations [103]. Assessments of satisfaction with business-to-business (B2B) relationships have also been shown to be highly susceptible to halo error [94], as have assessments of satisfaction in consumer-to-consumer (C2C) transactions, especially in the context of online auctions [20]. This seemingly ubiquitous halo error can be particularly troublesome for firms seeking to identify strengths or weaknesses in their relationships with customers, since attribute-level measurements of customer satisfaction will inevitably be biased according to a customer's overall impression of the firm.

Unfortunately, halo effect contamination extends much more deeply into the world of business than has thus far been described. Holbrook [36], for example, identified the halo effect as problematical to evaluative judgments by consumers in situations involving product research, while O'Donnell and Schultz [67] found strong evidence of halo error in strategic risk assessment. Neither is an organization's reputation immune from the scourge of the halo effect. Assessments of different attributes of store image, for example, were observed to be susceptible to halo error [106], with a judge's gender and level of experience being identified as significant predictors. Interestingly, Coombs and Holladay [22] found that the halo effect can actually serve as a shield that protects the reputation of an organization subsequent to an accident in a crisis situation. When an accident caused by human error occurs during the response to a crisis, the responding organization is less likely to be perceived as culpable and more likely to be given the benefit of the doubt if it enjoyed a favorable reputation prior to the crisis.

The notion of leveraging the halo effect for gain ranges into other aspects of corporate and managerial reputation, as well. Brown and Perry [15], for example, noted that a company's past financial performance generates a halo that heavily influences future corporate ratings. It has similarly been noted that the words used to describe the characteristics of corporate leaders depend almost entirely upon corporate performance [56, 57]. As a more general rule, when a company is perceived to be doing well (or poorly), specific attributes of the company such as its management, strategy, personnel, or work environment will be judged as superior (or inferior), regardless of objective reality [77, 78].

2.4 The Halo Effect and Technology

Research conducted in the past few decades has begun to reveal that halo effect contamination is a much more pervasive and farther reaching phenomenon than was originally supposed, with recent studies indicating that assessments of technology are also commonly contaminated by halo error. In the context of university technology transfer, for example, outside evaluators' assessments of the merits of a new technology are artificially inflated by both the prestige of the university where the technology was developed [85] and by the university's willingness to accept an equity position [30]. Moreover, although evaluations of IT capabilities do not appear to be linked to past financial performance among firms that are considered IT leaders [11], it has been noted that halo effects contaminate many studies that examine the relationship between IT capabilities and firm performance [81]. More recent work in the area of IT capabilities has used a data gathering process intentionally designed to minimize halo error [58]. The need to consider and control for halo error has also been highlighted in the context of collaborative filtering algorithms, where the halo effect was found to contaminate movie ratings that involved multiple components [79].

Additional manifestations of the halo effect in the context of technology have lately appeared in a wide variety of academic outlets. Kim [46], for example, concluded that positive past experiences with a technology are often projected onto similar new technologies by way of a halo effect, with Massey and Montoya-Weiss [55] reporting a specific example of this phenomenon in perceptions of knowledge conversion technologies. Perceptions also form the basis of Fred Davis' [26] Technology Acceptance Model (TAM), and since all of the major constructs in TAM (*i.e.*, ease of use, usefulness, and usage intentions) are self-reported, Szajna [90] found that those constructs are subject to contamination by the halo effect. Lee and Chen [49] expressed similar concerns with respect to halo error in participants' self-assessments of past usage of and future intentions to use virtual worlds. It appears that not all self-assessments are contaminated by halo error, however, given that no halo effects were detected in the context of software developers' assessments of their own performance [74].

In addition to influencing assessments of specific technologies and systems, halo effects have also been suspected by several authors to produce distortions in evaluations of web design. Hartmann et al. [35], for example, speculated that halo error may influence judgments of website usability and aesthetics, while Deng and Poole [28] speculated that the initial emotional response evoked by the visual characteristics of a website may be carried by way of a halo effect to evaluations of other aspects of the website. In a vein similar to that of the halo effect research conducted in the context of human attractiveness, it has been noted that a user's overall assessment of a website's attractiveness may lead to inflated judgments of the quality of other attributes of the website such as perceived usefulness [69]. Inversely, assessments of a website's visual appeal have been observed to be higher when other characteristics of the website have also been judged to be of a high quality, suggesting the possibility of halo effect contamination among such assessments [99]. While the studies above speculate that halo effect contamination may be present in assessments of web design, it is important to note that in each case such speculations were only secondary conjecture – heretofore the role of the halo effect in assessments of web design has never been explicitly or directly tested.

3 RESEARCH HYPOTHESES

Having described and reviewed the foundational literature on halo effect theory, and having illustrated the ways in which the halo effect has been found to exert an undesirable influence on people's assessments of organizations, technologies, and other human beings, we now turn our attention to the research hypotheses that form the core of the current study. Specifically, in this section we rely on halo effect theory and the findings of past research in this area to develop a series of theory-driven hypotheses about the role of the halo effect in assessments of web interface design.

To begin, one of the foundational concepts of halo effect theory is that perceptions of an entity's individual characteristics or traits are derived largely – but not entirely – from objective reality; *i.e.*, the normative values of an entity's individual traits serve as the perceptual foundation upon which evaluations or judgments of those traits are based. The theory simply predicts that one's final perceptions of individual traits will deviate from reality because of the heuristic nature of human cognitive processing. In the context of assessments of web interface design, we therefore expect perceptions of web interface design characteristics to be linked to objective reality. Namely:

H1: Perceptions of web interface design traits are positively related to the normative values of those traits.

The relationship between the normative values of an entity's individual traits and perceptions of those traits notwithstanding, the primary prediction of halo effect theory holds that we are cognitively unable to separate our overall affect toward an entity from evaluations of the entity's individual characteristics. That is, perceptions of an entity's individual traits are influenced by general feelings toward the entity, and we are hence unable to independently assess those traits without our ratings being colored by our general feelings toward the entity. Importantly, this overall affect emerges from an aggregation of whatever information we have available about the entity, be it years of personal experience, or something as simple as a photograph. When assessing a never before seen web interface, we therefore expect perceptions of the interface's design characteristics to be distorted according to the judge's overall affect toward the textual content of the website being judged. Namely:

H2: Perceptions of web interface design traits are positively related to the judge's general feelings toward the textual content of the website whose traits are being judged.

Inasmuch as halo effect theory predicts that overall affect will produce distortions in judgments of individual characteristics, it is in turn important to consider conditions that can influence a person's affect. A core phenomenon that has been shown to influence affective states is the priming effect [47]. Like the halo effect, the priming effect is a cognitive bias that can cause people to behave differently than what would otherwise be expected under a model of rational decision-making. The priming effect is linked to implicit memory, and is characterized by exposure to a stimulus that unconsciously influences reactions to or feelings toward future stimuli [47, 93]. Inquiring into a person's affective states loads those feelings into the person's working memory. This act unconsciously alters the person's cognitive processes, and hence influences her future responses. In the context of the current experiment, isolating and controlling for the influence of priming was therefore critical to obtaining an accurate measurement of any halo effects. Since the priming effect predicts that a person's self-reported affective feelings toward a web interface will vary according to whether she is asked about those feelings before or after judging the interface's individual traits, we hypothesize the presence of psychological priming so as to control for its anticipated effects in our measurement model. Thus:

H3: A subject's self-reported overall affect toward a website varies according to whether she is asked about her affective feelings before or after judging the website's design traits.

Beyond psychological priming, we also sought to isolate and control for substantive inter-judge differences that could significantly influence a judge's affect or decision-making when evaluating a web interface. First, the age of a subject has been clearly identified in the psychology and managerial literatures as a factor that influences both affect [19, 62, 63] and judgment [41, 54, 91]. Insofar as affect and judgment are central constructs in the current study, we deemed it judicious to control for the effects of subject age on these constructs:

H4a: A subject's age influences her overall affect toward a website.

H4b: A subject's age influences her judgments of website interface design characteristics.

There is also well-established evidence that suggests that a subject's gender influences both her affective states [32, 51] and her decision-making [73, 97]. Recent work has also found strong gender effects in the context of interactions with information technologies [95, 96], with fMRI brain-scanning research revealing gender-based differences in both human brain function and neural information processing [76]. With specific

respect to gender differences in the context of the halo effect, the literature suggests that the magnitude of the halo effect depends upon gender [43], and that halo error is more pronounced in men than among women [53, 106]. Thus:

H5a: A subject's gender influences her overall affect toward a website.

H5b: Halo effect-based distortions in judgments of web interface design are more pronounced among men than among women.

Finally, there appears to be at least some evidence that suggests that a subject's affective states and judgments within a domain are related to the amount of experience that the subject has accumulated within that domain. Experienced physicians, for example, have been found to show little empathy toward suffering patients [27, 60], while experienced soldiers have a well-known proclivity to dehumanize enemy soldiers and civilians to a greater extent than do their less experienced counterparts [37, 45]. Although experience is known to be a substantially weaker predictor of decision performance than other factors [91], research in the managerial and marketing literatures suggests that experience can yield more accurate decisions and less distortion in the context of halo effect-related judgments [80, 106]. Thus:

H6a: A subject's level of web experience influences her overall affect toward a website.

H6b: Distortions in judgments of web interface design characteristics are weaker among more experienced web users.

4 METHODOLOGICAL APPROACH

Inquiry into the hypotheses developed above was carried out in accordance with early foundational research by Johnson and Vidulich [40], who established controlled experimentation as the soundest approach to assessing the predictions of halo effect theory. Quoting these authors, "The correct method of investigation is experimental variation of the conditions of judgment and comparison of halo effects manifested under the different conditions" (p. 134). To that end, our experiment was broadly structured around two large groups of subjects; namely, a baseline (control) group and a halo effect (treatment) group. With a view toward conducting a rigorous investigation into the nature of halo effect contamination in assessments of web interface design, a rather broad scope was adopted for the experiment. *In toto*, the experiment incorporated 3 polarizing issues, 6 different types of websites, 42 web interface variations, and 1,230 research subjects. In the following subsections we provide details regarding the design and execution of the experiment, as well as a description of the means by which the resulting data were analyzed.

4.1 Website Types and Interface Design Characteristics

In the aggregate, subjects in the study were variously exposed to and asked to evaluate the interface design characteristics of six different types of websites, including a search website, an e-commerce website, a news website, a blog website, a social networking website, and a video sharing website. This was done to ensure a wide range of coverage with respect to the variety of websites that a contemporary web user might encounter, and to assess the extent to which any halo effects were consistent and generalizable across different types of websites. This approach also helped mitigate the possibility of any observed halo effects being attributable solely to random chance.

Following a detailed review of the academic and trade literatures, Aladwania and Palvia [2] concluded that the appearance of a web interface is a unique construct, and should be treated separately from other constructs such as content or technical adequacy. In deference to this conceptualization, a web interface template was created for each of the six different types of websites described above. This approach was taken so as to allow the textual content of the web page (*e.g.*, the title, text, etc.) to be modified while preserving the layout and appearance of the web interface itself. In this way, subjects could be asked to evaluate the design characteristics of a web interface while being subjected to experimentally manipulated variations in textual content.

Given that subjects had never before seen the interface or its textual content, their preexisting, overall affect toward the website would emerge solely from these informational cues. Holding the design of the interface constant thus ensured that differences in subjects' affect toward the website would be linked specifically to the experimentally manipulated variations in textual content. Further details about the role of these interface templates in the experiment will be provided shortly in the subsections describing the baseline and halo effect groups. An example of an interface template being subjected to the content substitution process is illustrated in Figure 2.

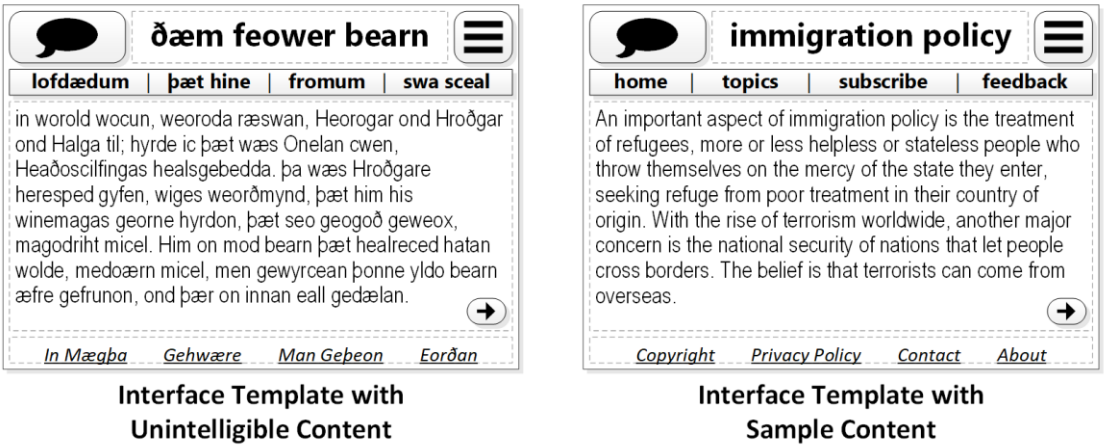

Figure 2. Sample web interface template and content substitution process

Table 1. Original subscale items and modified items as used in the current experiment

Original Statement [2]	Modified Statement Used in Current Experiment
____'s website looks attractive.	This website looks attractive.
____'s website looks organized.	This website looks organized.
____'s website uses fonts properly.	This website uses fonts properly.
____'s website uses colors properly.	This website uses colors properly.
____'s website uses multimedia features properly.	This website uses multimedia features properly.

Inasmuch as the halo effect is concerned with distorted evaluations of an entity's individual characteristics, it was necessary to identify a set of specific interface design characteristics that subjects could be asked to evaluate in order for the experiment to be properly aligned with halo effect theory. To this end, we adopted a five-item subscale designed to measure the attractiveness of a web interface design from within a broader instrument oriented toward measuring user perceptions of overall website quality [2].

In accordance with the original instrument, subjects in the experiment were asked to respond to the evaluative statements in the subscale using a seven-point, Likert-type scale anchored at *1 = strongly disagree* and *7 = strongly agree*. Minor modifications were made to the wording of the items in the subscale in order to adapt those items to the context of the current experiment, as shown in Table 1.

4.2 Baseline Group

From a theoretical perspective, a set of normative or objective ratings of a web interface's individual traits is required as a basis of comparison if one hopes to determine whether and to what extent the halo effect may be contaminating assessments of those traits (*vide supra*, Figure 1). A baseline (control) group was therefore used to establish normative ratings for the design characteristics of the web interfaces used in the study. For guidance as to how normative ratings could be obtained, we again turned to the foundational literature on halo effect theory. First, since the halo effect manifests itself when a judge is asked to rate multiple individual traits for the same entity, Bingham [12] concluded that the halo effect could be mitigated by directing subjects to focus on just one specific trait of the entity being judged – a position later echoed by Johnson and Vidulich [40]. Quoting their work:

> *One group rated five well-known individuals on five traits under conditions designed to maximize the halo effect. They rated one individual each experimental day on all traits. The other group rated the same individuals on the same traits under conditions designed to minimize the halo effect. They rated all individuals on one trait each experimental day.* (p. 134)

Second, halo effect theory posits that the strength and direction of halo error is influenced by the nature of one's overall affect toward the entity being judged. In the absence of a preexisting, broad impression, past research suggests that any such distortions in judgment

will be minimized. Recall, for example, the work of Landy and Sigall [48], in which judges who were shown *ex ante* a photo of an attractive author consistently rated the quality of an essay more favorably than when the author was supposed to be unattractive. Meanwhile, judges who were not shown a photo of the author provided more objective ratings that consistently fell between those of the judges who had seen a photo. Sattler et al. [83] observed a similar phenomenon in their work examining the validity of intelligence testing. Evaluators who had not been provided with any information regarding subjects' intelligence prior to administering intelligence tests consistently arrived at scores that fell between those produced by evaluators who had been told *ex ante* that subjects were of either above or below average intelligence. From a theoretical perspective, the experimentally manipulated material served as informational cues that allowed judges to formulate affective associations toward a subject prior to the evaluation process. The subsequent effects of these associations manifested themselves as distortions in judgment (*i.e.*, as halo effects).

In light of these considerations, the current experiment incorporated the mechanisms described above to obtain normative assessments of the design characteristics of each web interface used in the study. To wit, subjects in the baseline group evaluated each of the six web interfaces, but were asked to evaluate just one design trait for each interface, with the subject evaluating the same trait for all six interfaces. Upon completion of data gathering for the baseline group, normative values could then be computed as the mean of subjects' ratings for each trait. The specific design characteristic that each subject in the baseline group was asked to evaluate was determined using iterative assignment, and the order in which the six web interfaces were presented to each subject was randomized to mitigate any ordering or self-generated validity effects [18, 82].

Finally, since any intelligible textual content displayed on a web interface might serve as an informational cue that could trigger an affective association toward that content – and by extension, distortions in judgment – it was necessary for each web interface evaluated by the baseline group subjects to be populated with neutral, unintelligible content. For purposes of the current study, the original, untranslated text from the ancient epic poem *Beowulf* [3] was thus used as the content basis for web interfaces evaluated by the baseline group. Since this epic poem was written in Old English (Ænglisc), which has now been a dead language for more than 800 years, its original text is essentially unintelligible to modern readers (highly specialized linguists notwithstanding). An example of such unintelligible content appears in Figure 2. As a final contrivance aimed at obtaining normative ratings, baseline group subjects were specifically admonished to disregard the textual content displayed on the web interfaces, and instead focus their attention solely on their design characteristics.

4.3 Halo Effect Group

The experiment also incorporated a halo effect (treatment) group, the purpose of which was to generate responses that could be compared to those of the baseline group for evidence of halo error. Whereas the experimental conditions to which the baseline group was subjected were designed to elicit objective, normative ratings, members of the halo effect group were subjected to conditions designed to allow any potential halo effects to emerge and be measured. Whereas a subject in the baseline group was asked to rate all six web interfaces along a single, randomly assigned dimension, subjects in the halo effect group were asked to evaluate a single interface along all five dimensions.

Following past research [40], this deviation from the baseline group required subjects to rate multiple individual characteristics of a single entity – in this case, a web interface. Since halo effect theory posits that subjects are cognitively incapable of providing objective ratings of an entity's individual traits when tasked with evaluating the entity along several dimensions simultaneously [92, 98], this design was adopted so as to allow any such departures from objectivity to be identified. The specific web interface that each halo effect subject was asked to evaluate was determined using iterative assignment, and the order in which the five evaluative statements were presented to subjects was randomized with a view toward nullifying any potential ordering or self-generated validity effects [18, 82].

The other major difference in the halo effect group related to the content that was displayed on the web interface templates whose design characteristics subjects were asked to evaluate. Whereas subjects in the baseline group were exposed to unintelligible content, the content appearing on interfaces evaluated by the halo effect group was intentionally polarizing in nature. Inasmuch as halo effect theory postulates that subject ratings will be colored by the subject's broad affective impressions, this approach was adopted in order to indirectly activate or trigger subjects' preexisting cognitive associations. It is important at this point to emphasize and reiterate that subjects were not asked to evaluate the textual content appearing on their assigned web interface, but rather only to rate the interface's design characteristics. As with the previously described studies on author attractiveness [43, 48], this experimental manipulation was intended to provide. a

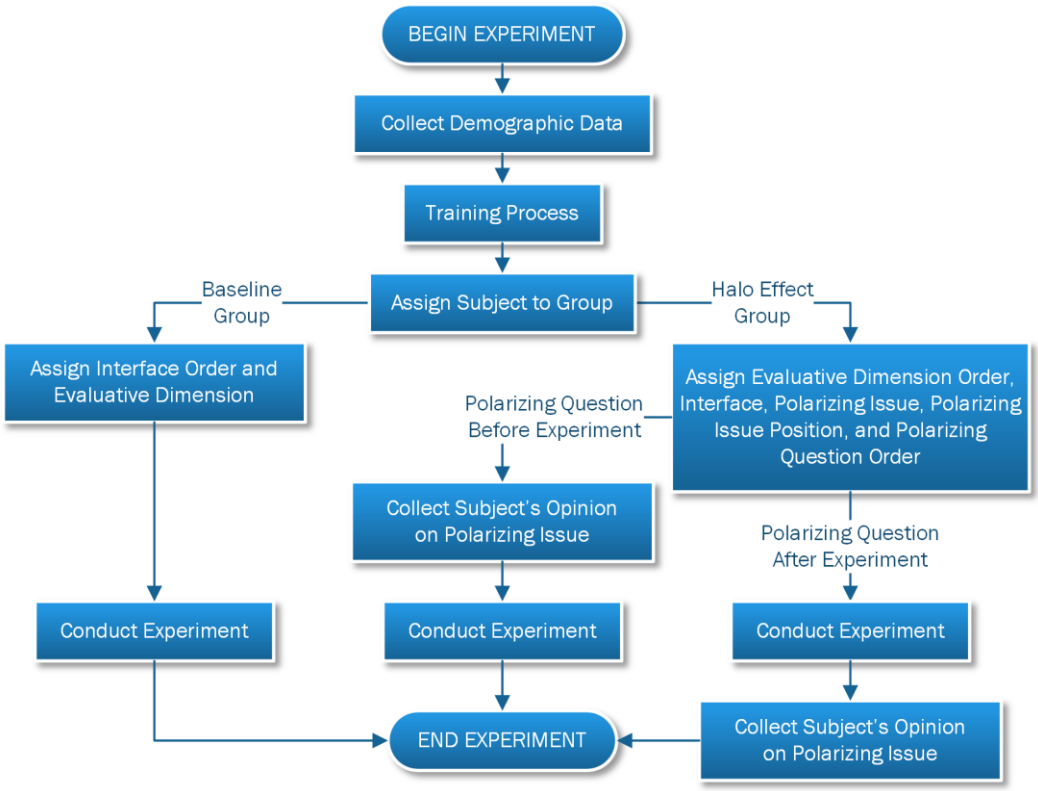

Figure 3. Experimental process

strong psychoemotional anchor – unrelated to the task at hand – that could induce preexisting affective associations and subsequent distortions in judgment. Three issues acknowledged by scholars to be socially polarizing in the United States were therefore included in the study; namely, abortion rights [61], immigration policy [64], and gun control laws [102].

For each of the three polarizing issues, two variations in content were developed for use with the web interfaces evaluated by the halo effect group. These two variations in content reflected the diametrically opposing points of view for each polarizing issue. Using the abortion rights issue by way of example, one set of textual content shown to halo effect subjects would be supportive of abortion rights, while the other set of content would be opposed to abortion rights. The specific polarizing issue to which a halo effect subject was exposed was assigned iteratively, as was the position – either supportive or oppositional – of the textual content seen by that subject. With six web interface templates, three polarizing issues, and two content positions for each issue, a total of 36 web interface variations were created for use with the halo effect group. In order to determine whether and to what extent any halo effects were related to subjects' preexisting affective impressions, it was necessary to inquire into each subject's position regarding the polarizing issue with which he or she was presented. Subjects were thus asked to specify their feelings toward the polarizing issue to which they had been assigned using a seven-point, Likert-type scale. Anchors for the scale were established at *1 = I am strongly opposed to* ____ and *7 = I am strongly supportive of* ____, where "____" may have been either *abortion rights*, *immigration*, or *gun control*, depending upon the polarizing content to which the subject was exposed.

With a view toward determining the extent to which priming effects may influence halo effect contamination [47], half of the subjects in the halo effect group were asked to specify their feelings toward the polarizing issue prior to rating the web interface, while the remaining half were asked to specify their feelings after already having rated the interface. Whether a given halo effect subject was asked to identify their position before or after having rated the web interface was determined by means of iterative assignment.

4.4 Experimental Process and Execution

All of the abovementioned elements were aggregated into an interactive, web-based software system that was designed not only to carry out the experiment, but also to gather and store all of the data generated therefrom. After collecting demographic data (*i.e.*, age, gender, and

level of web experience), the system guided subjects through a training process that pointed out the salient features of the software program and instructed them as to how they could respond to the questions posed during the experiment. Upon completing the training process, the system next randomly assigned 14.3% of subjects to the baseline group and 85.7% of subjects to the halo effect group, per the sample size calculations described shortly hereafter. Depending upon the group into which a subject was allocated, the system then iteratively assigned the balance of the parameter settings required for that group, after which the experiment itself was executed. The overall flow of the experimental process is shown in Figure 3.

As noted previously, the halo effect group involved six different types of websites, three polarizing issues, and two possible positions for each issue (either supportive or oppositional). Halo effect subjects would thus be exposed to one of 36 possible interface configurations. For purposes of statistical validity a minimum of 30 subjects was required for each possible configuration, thus bringing the preliminary minimum sample size for the halo effect group to 1,080 subjects, notwithstanding any sample size requirements imposed by the structure of the measurement model. The baseline group by contrast involved six interfaces containing unintelligible content (one interface for each of the six different types of websites) and five possible evaluative dimensions for a given interface, thus bringing the total number of possible configurations to 30. For purposes of statistical validity each possible configuration required a minimum of 30 responses, thus bringing the total minimum number of responses to 900. Since each subject in the baseline group evaluated all six unintelligible interfaces along the same dimension, six responses were recorded per subject, bringing the minimum sample size for the baseline group to 150 subjects, and the minimum sample size for the overall experiment to 1,230 subjects.

The sample size calculations above are, of course, based solely on the design characteristics of the experiment, and do not directly consider the sample size required to detect a significant effect in light of the structural properties of the measurement model. We therefore conducted an *a priori* sample size analysis using the most recently developed computational methods for performing such analyses for structural equation models [86, 100]. The results of these calculations revealed first that a minimum sample size of 579 subjects would be required in order to detect a small but statistically significant effect size of 0.1 for the halo effect group [21] while maintaining an adequate statistical power level of 0.8, and second, that a minimum sample size of 700 subjects would be required based upon the structural properties of the research model. It was therefore concluded that the preliminary minimum sample size of 1,080 subjects for the halo effect group would be sufficient to accommodate the competing demands imposed by statistical power and anticipated effect size in light of the design of the structural model.

Motivated by the experiment's large sample size requirements, and considering that the study's target population was adult web users in the United States, we engaged the leading global online advertising firm to craft a targeted campaign for the purpose of soliciting volunteers for the experiment. The firm's advanced technology allowed subject recruitment to be constrained to web users in the United States who were at least 18 years old, thus ensuring that only those subjects who were a part of the target population would be included. IP address restrictions were also enforced to help prevent the same person from participating in the study more than once. In total, our campaign attracted 1,230 subjects, of whom 48.0% were male and 52.0% were female. Subjects reported an average level of web experience of 3.85 on a five-point, Likert-type scale anchored at *1 = very little experience* and *5 = extensive experience*. Subjects ranged in age from 18 to 78 years. The average age was 33.65 years with an interquartile range of 25 to 40 years, indicating an age distribution skewed in the direction of youth. These values were consistent with Internet usage demographics among adults in the United States [71].

4.5 Data Gathering and Preparation

Data were gathered continuously until the necessary sample size had been achieved. Since each of the 150 subjects in the baseline group rated each of the six different types of websites along a single evaluative dimension, a total of 900 ratings were generated by the baseline group. By contrast, each of the 1,080 subjects in the halo effect group rated one of the six different types of websites along all five evaluative dimensions, thus yielding a total of 5,400 ratings for the halo effect group, and a grand total of 6,300 ratings for the entire experiment. Data from both groups were assembled into a single repository for analysis, with binary dummy-coded values of 0 and 1 being used to identify subject gender and group membership. A similar series of binary dummy variables was used to encode the type of website and evaluative dimension to which each rating

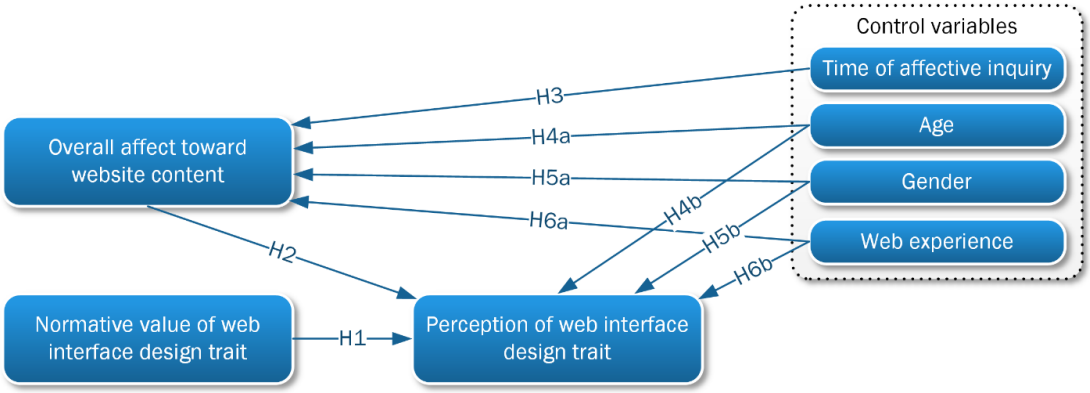

Figure 4. Research model

applied, as well as to identify the polarizing issue, content position (either supportive or oppositional), and priming effect condition to which halo effect subjects were exposed.

Finally, it was necessary to compute halo effect subjects' overall levels of affect toward the textual content with which they were presented. Affect levels were coded using a seven-item, ordinal integer scale ranging from -3 (strongly negative affect) to +3 (strongly positive affect), and were computed as a function of both a subject's feelings toward the polarizing issue to which she had been assigned and the randomly assigned content position. By way of example, consider a subject who had been assigned the abortion rights issue, and was asked about her feelings using a seven-point, Likert type scale (anchored at *1 = I am strongly opposed to abortion rights* and *7 = I am strongly supportive of abortion rights*). If the subject responded with a value of *7* and had been randomly assigned to see content that was strongly supportive of abortion rights, then the "distance" between her feelings and the content would be minimal, thus leading to an assignation of +3 on the affect scale (*i.e.*, strongly positive affect toward the content).

If, however, the subject responded with a value of *7* and had been randomly assigned to see content that was strongly opposed to abortion rights, then the "distance" between her feelings and the content would be maximal, thus leading to an assignation of -3 on the affect scale (*i.e.*, strongly negative affect toward the content). More specifically, affect levels were assigned using the following conditional formula:

$$a = s\ ?\ r - 4 : 4 - r \quad (1)$$

Where a is the subject's level of affect, s is a Boolean value indicating whether the content seen by the subject was supportive of the polarizing issue to which she was assigned, and r is the subject's numeric Likert-scale response regarding her feelings toward that polarizing issue (an integer ranging from 1 to 7).

4.6 Research Model and Analytical Approach

In light of the structural model of halo effect theory shown in Figure 1, and in light of the hypotheses developed previously, the research model depicted in Figure 4 formed the foundation of inquiry for the study. As shown in Figure 4, after accounting for the influence of several literature-derived control variables (H3 through H6), a subject's rating of each web interface design trait was predicted to be a function of the normative value of that design trait (H1) and the distortions in judgment (*i.e.*, the halo effects) introduced by the subject's overall affect toward the textual content displayed on the webpage (H2). In deference to the structural nature of the research model, all of the hypothesized relationships were estimated simultaneously using structural equation modeling (SEM) techniques with maximum likelihood estimation.

Prior to estimating the structural model, however, it was first necessary to determine whether doing so would be statistically defensible. A series of preliminary analyses was hence conducted with a view toward (1) establishing the validity of the measurement model, and (2) acquiring a deeper understanding of the nature of the experimental data. With respect to the former of these two objectives, discriminant validity and common method bias were assessed by means of correlation analyses [70] and Harman's single factor test [72, 84]. With respect to the latter objective, one-way analyses of variance, Levene tests for homogeneity of variance [66], and tests of between-subjects effects were conducted in order to evaluate the extent to which differences were present among the responses of subjects in the baseline and halo effect groups [39].

Table 2. Pearson correlations and significances among study variables

	Age	Gender	Web experience	Priming	Affect	Normative value	Subject response
Age	1.000						
Gender	0.063**	1.000					
Web experience	-0.082**	0.201**	1.000				
Priming	0.027*	0.036**	-0.047**	1.000			
Affect	-0.040**	-0.041**	0.034*	0.035*	1.000		
Normative value	-0.027*	0.018	-0.017	-0.036**	0.005	1.000	
Subject response	-0.064**	-0.043**	-0.015	-0.021	0.032*	0.189**	1.000

*$p < 0.05$ **$p < 0.01$

5 Analytical Results

5.1 Preliminary Analyses

Excepting for age and gender, all of the subject response values for the study were obtained using Likert-type scales, and to this end it was necessary to establish discriminant validity and consider the potential impacts of common method variance (CMV). An examination of the correlation matrix for the study (shown in Table 2) revealed no particularly large correlations among the variables used in the study, thus providing preliminary evidence of discriminant validity and supporting the notion that no issues were present with respect to CMV [70]. Further support for the validity of the measurement model was provided by Harman's single factor test, the results of which revealed seven principal components, with the largest accounting for just 17.73% of the total variance. In light of the comparatively small correlations among the study variables, the results of this test indicate that the study had no significant problems with CMV [72, 84].

Next, a series of preliminary one-way ANOVAs and Levene tests for homogeneity of variance was conducted to obtain deeper insights into the nature of the study's data [39, 66]. The first of these preliminary tests considered differences in the distributions of web interface design ratings between the two groups. The mean rating for interface traits in the baseline group was 4.70 (on a scale ranging from 1 to 7) with a standard deviation (*sd*) of 1.63, while the mean rating for the same characteristics in the halo effect group was 4.21 (*sd* = 1.76). Without controlling for any other factors, the ANOVA thus revealed that the distribution of response values in the halo effect group was significantly different from the distribution of response values in the baseline group ($p < 0.001$). The results of the Levene test also revealed a significantly larger degree of variation in the responses of subjects in the halo effect group as compared to their counterparts in the baseline group ($p < 0.05$).

Considering next the nature of the baseline group in greater detail, tests of between-subjects effects for the baseline group revealed no statistically significant main effects of gender, age, or web experience on subject ratings regarding the various web interface design traits considered in the study. Recalling that differences in age, gender, and experience have been shown by past research to be significant factors in situations wherein halo error was present, these results were encouraging since they implied that the theory-driven manipulations used in the study created an experimental environment in which the normative measures of interface quality obtained from the baseline group were indeed statistically free of halo effect-induced distortions in judgment. Further, statistically significant main effects in subject responses were detected with respect to both the web interface design traits being judged and the type of website being judged. Together, these observations were also encouraging insofar as they revealed that subjects in the baseline group were able to discriminate not only among the various web interface design traits, but also among the various types or categories of web interfaces to which they were exposed.

Finally, considering the halo effect group in greater detail, tests of between-subjects differences among the members of the halo effect group revealed no main effects of web experience or priming on subject ratings of the various web interface design traits. By contrast, the normative value of each interface characteristic (as obtained from the baseline group), along with age, gender, subject affect toward the polarizing issue, the polarizing issue itself, the type or category of website being considered, and the specific interface characteristic being evaluated were all observed to have significant main effects on the judgment of subjects in

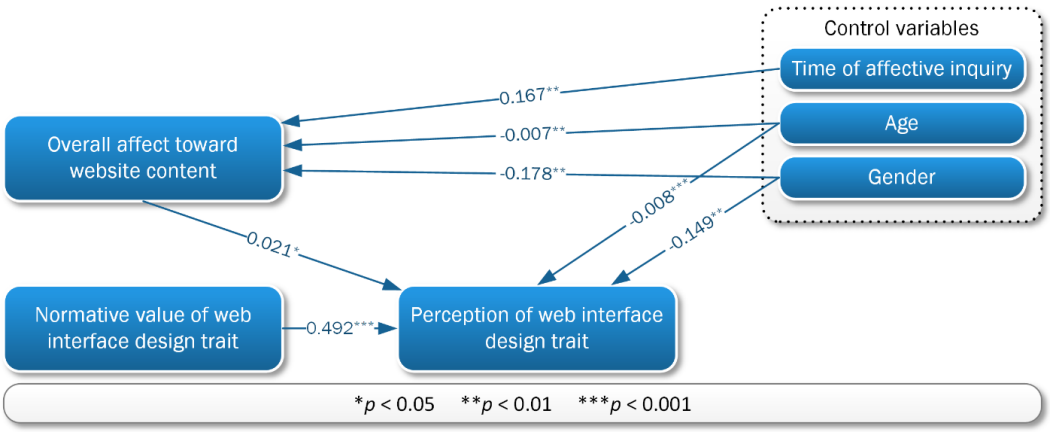

Figure 5. Research model parameter estimates and significances

Table 3. Measures of fit for the research model

Measure of Model Fit	Value
Chi-square (*p*):	14.107 (0.015)
Standardized root mean square residual (SRMR):	0.012
Comparative fit index (CFI):	0.968
Tucker-Lewis index (TLI):	0.904
Normed fit index (NFI):	0.953
Adjusted goodness of fit index (AGFI):	0.996
Root mean square error of approximation (RMSEA):	0.018
Probability of close fit (PCLOSE):	0.999

the halo effect group. Together, these findings provided reasonable preliminary evidence that one or more of the study's primary hypotheses may have been tenable, and justified proceeding with the estimation of the primary structural research model.

5.2 Structural Model Estimation

Initial estimation of the structural model revealed significant relationships among all of the hypothesized inter-construct relationships, except for those associated with web experience (*i.e.*, Hypotheses 6a and 6b were not supported). Since web experience was not observed to be related to overall affect toward website content or perceptions of the design of a web interface, those paths were removed from the structural model, after which the model was re-estimated. The final parameter estimates for the structural model are shown in Figure 5.

In light of the structural equation modeling methodology that was used to obtain the parameter estimates shown in Figure 5, it was also necessary to evaluate the extent to which the research model fit the observed data. To this end, Table 3 below provides the values of several of the commonest measures of fit as they apply to the research model described above.

Excepting for the value of the chi-square metric, each of the measures of fit shown in the table implies a very good to excellent fit between the research model and the observed data. Although the chi-square metric was statistically significant, chi-square values are known to be inflated for models that use large samples [7], and under such circumstances are almost always statistically significant [38]. In light of the unusually large sample size used in our study, we do not believe that the significant chi-square value indicates poor model fit; the preponderance of the evidence reported in Table 3 implies a very good fit between the model and the observed data, and we therefore conclude that the model was properly specified.

5.3 Tests of Hypotheses

Our first hypothesis addressed one of the foundational concepts of halo effect theory; namely, that perceptions of an entity's individual characteristics or traits are derived largely – but not entirely – from objective

reality. Since perceptions of web interface design traits were hypothesized to be positively related to their normative values, those values were hypothesized to serve as the perceptual foundation upon which evaluations of such traits would be based. Results obtained from the estimation of the structural model revealed a strong positive relationship between the normative values of web interface design traits and subjects' perceptions of those traits ($p < 0.001$). Hypothesis 1 was therefore supported.

Our second hypothesis addressed the primary prediction of halo effect theory, which holds that human beings are cognitively unable to separate our overall affect toward an entity from our evaluations of the entity's individual traits, with our affect being derived from an aggregation of whatever information we have available. In the context of the current study, perceptions of web interface design traits were therefore hypothesized to be distorted by a subject's overall affect toward the textual content of the website whose traits were being judged. The results obtained from estimating the structural model revealed a significant positive relationship between a subject's overall affect toward the content of a website and her perceptions of the website's interface design traits ($p < 0.05$). Hypothesis 2 was therefore supported.

Our third hypothesis sought to assess and control for the role of psychological priming, which has been shown to exert a strong influence on a person's affective state. The simple act of inquiring into a person's affective or emotional states can unconsciously alter her cognitive processes, and hence influence her future responses. In an effort to isolate and accurately quantify halo effect-based distortions in judgment, it was therefore deemed prudent to control for the effects of psychological priming. A person's self-reported affective feelings toward a website's textual content were therefore hypothesized to vary according to whether she was asked about those feelings before or after rating the characteristics of the website's interface. Results obtained from the estimation of the structural model revealed a strong positive relationship between the time of affective inquiry and a subject's self-reported overall affect toward the textual content of a website. Hypothesis 3 was therefore supported.

Given that age has been identified in the psychology and managerial literatures as a significant factor that can influence affect and judgment, the study included a set of hypotheses designed to isolate and control for these effects. Hypothesis 4a sought to assess and control for the influence of a subject's age on her overall affect toward a website's content, while Hypothesis 4b sought to assess and control for the influence of the subject's age on her perceptions of a website's interface design traits. The results obtained from estimating the structural model revealed a significant inverse relationship between a subject's age and her overall affect toward the content of a website ($p < 0.01$), thus providing support for Hypothesis 4a. A significant inverse relationship was also observed between a subject's age and her perceptions of the website's interface design traits ($p < 0.001$), thus providing support for Hypothesis 4b.

In addition to age, the literature also provides well-established evidence that indicates that a subject's gender has a significant impact on both affective states and decision-making. Gender has also been identified by past research examining the halo effect as a highly significant factor with respect to the degree to which a subject's judgments are contaminated by halo error. The study therefore included a set of hypotheses designed to isolate and assess these gender effects. Hypothesis 5a sought to assess the extent to which a subject's gender influences her overall affect toward website content, and Hypothesis 5b proposed that the degree of halo-based error in assessments of web interface design traits would be more pronounced among men than among women. Evidence from the primary structural model indicated that a subject's gender does indeed influence her overall affect toward the content of a website ($p < 0.01$), thus providing support for Hypothesis 5a.

The primary structural model also provided evidence of a significant gender-based difference in perceptions of web interface design traits ($p < 0.01$). In light of this significant difference, it was deemed prudent and justifiable to proceed with a full test of Hypothesis 5b. For this purpose, the primary data set was split into two subsets, the first of which contained the data gathered from the 551 female members of the halo effect group, and the second of which contained the data for the 529 male members of the halo effect group. The structural model was then re-estimated using these two independent data sets.

Among men, a highly significant positive relationship was observed between overall affect toward the content of a website and perceptions of the website's interface design traits (parameter estimate = 0.042, $p < 0.01$). Among women, however, no significant relationship was observed among these constructs (parameter estimate = 0.007, $p = 0.780$). These findings indicate that halo error in assessments of web interface design is, from a statistical perspective, entirely isolated to men, thus providing support for Hypothesis 5b. Since these observations have important implications for the growing body of literature that documents gender-based differences in human brain structure, function, and neural information processing, this finding will be discussed in greater detail in the following section.

Finally, some evidence exists in the literature that suggests that a subject's affective states and judgments within a domain may be related to the amount of

experience that the subject has accumulated within that domain. To this end, a set of hypotheses was included in the study with the intent of isolating and controlling for the effects of web experience: Hypothesis 6a proposed that a subject's level of web experience can influence her overall affect toward the content of a website, while Hypothesis 6b proposed that distortions in judgments of web interface design characteristics are weaker among more experienced web users. The results obtained from estimating the structural model revealed no significant relationship between subjects' levels of web experience and their overall affect toward the content of a website, or between their levels of web experience and their perceptions of a website's interface design characteristics. Hypotheses 6a and 6b were therefore not supported.

6 DISCUSSION AND IMPLICATIONS

The halo effect is a cognitive bias that causes human beings to behave very differently than what would be expected under a Bernoullian, utility-based model of decision-making [10]. Specifically, halo effect theory proposes that a person's overall affect toward an entity will produce distortions in evaluations of the entity's individual characteristics, thus leading to systematic errors in judgment. Like reflections in a funhouse mirror, perceptions of an entity's individual traits have a basis in reality, but are distorted by one's general feelings toward the entity. The halo effect has been studied in many domains of inquiry, including cognitive psychology, business, and technology, and has been found to be both a pervasive and pernicious phenomenon.

After controlling for other cognitive biases, including both priming and ordering (or self-generated validity) effects, and after also controlling for many other factors known to influence human judgment and decision-making, this study found direct evidence of halo effect contamination in assessments of web interface design. Insofar as it is the first known direct empirical test of the halo effect in the context of interface design, the current study extends the boundaries of halo effect theory into the realm of human-computer interaction (HCI), and this extension of the scope of the theory stands as a major contribution of this work. The findings of past HCI research – especially those studies that employed multi-item measurement scales – may therefore need to be revisited in light of the substantial halo-induced distortions in judgment that were detected here.

Beyond establishing halo error as a legitimate factor for HCI studies concerned with judgment and decision-making, this study also provides one of the first direct tests of and evidence for gender effects at the intersection of halo effect theory and information technology. Specifically, the data revealed strong and highly significant halo effects among men, while no statistically significant halo effects were observed among women. Since these findings were derived from the responses of more than 1,000 men and women, we believe that the likelihood of the observed results being attributable to a random statistical anomaly is quite remote. The findings reported here thus add to the growing body of scientific evidence that suggests that both affective states and decision-making are directly influenced by gender [32, 51, 73, 97], and that halo effect-based distortions in judgment are, at least in certain circumstances, largely isolated to men [43, 53, 106].

Although such conclusions have traditionally been considered controversial, past IS research has identified gender as a significant factor in people's interactions with technology [95, 96], and with the results of recent fMRI brain-scanning research, the scientific community is rapidly awakening to the fact that measurable and quantifiable differences do indeed exist between men and women with respect to brain structure, function, and neural information processing [76]. As noted by Cahill [17], "...investigators have documented an astonishing array of structural, chemical, and functional variations in the brains of males and females." The presence of strong halo effects in men may very well be an evolved manifestation of such variations; this idea will be expounded upon shortly in the general discussion.

6.1 Implications and Recommendations

The presence of halo error in evaluations of interface design – and potentially many additional facets of HCI – has manifold implications for both researchers and managers alike. With respect to the former, researchers must not expect evaluations of web interfaces to be accurate if those evaluations are made using traditional, multi-item measurement scales. More broadly, and much more profoundly, a century of research clearly warns that the results of *all* studies that employ multi-item measurement scales but do not control for halo error should be looked upon with a healthy dose of skepticism. Although several disciplines have begun to heed this clarion call, due to the lack of familiarity with the halo effect and its attendant implications, the IS field has not yet responded to this problem. The halo effect is particularly troubling for IS research since so many of our most hallowed and well-established theoretical constructs (such as *perceived usefulness* and *perceived ease-of-use*) are almost always measured using multi-item scales. Even if such scales exhibit high degrees of statistical validity and reliability, the bitter reality is that those measures of validity and reliability are likely based

on data that are contaminated with halo error. Adopting methods such as those used to establish normative ratings in the current study should therefore be considered in the context of multi-item measurement.

The results of this study also have important implications for managers. In the modern era, a website now commonly serves as the public face of an organization. To this end, the design of an organization's website has critical implications for the way in which customers and other interested parties will perceive the organization, with such perceptions influencing the organization's prospects for success. Since the design, layout, and appearance of a website can directly influence the way in which people think about the underlying organization, managers should be highly interested in ensuring that the design of their organization's website is well-aligned with the expectations of its users. In many cases, however, proposed designs are evaluated by parties that have a vested interest in the website or in the underlying organization itself. The results reported herein suggest that the overall affect of these parties may unintentionally contaminate their assessments of the website's design, thereby leading to erroneous conclusions. In order to avoid such problems, managers should strongly consider adopting mechanisms that have been shown to minimize halo error when soliciting assessments of proposed website designs. Examples of such mechanisms include those used herein to solicit normative website design ratings; *i.e.*, (1) directing judges to focus on just a single, specific trait of the interfaces being judged [12, 40], and (2) removing or minimizing any informational cues that may trigger an affective response in the judge [48, 83].

6.2 General Discussion

Although the halo effect has been successfully used by scholars in many different disciplines as a mechanism for both predicting and explaining deviations from rational judgment, after more than a century of research, we are just now beginning to gain preliminary insights into the origins and purpose of the halo effect. Put another way, despite knowing about the existence of the halo effect, the question of *why* the halo effect exists has remained a mystery. Why is it that under normal circumstances, human beings seem either unwilling or unable to evaluate an entity's individual traits in isolation without referencing their overall perceptions of the entity? Further, why does this cognitive phenomenon appear, at least in certain circumstances, to be more pronounced in men than in women? The remainder of this section will draw on elements from both cognitive and evolutionary psychology to outline an integrated theoretical framework that may prove fruitful in helping to address these questions.

In his Nobel Prize-winning work, psychologist Daniel Kahneman erected the scaffolding of a great deal of what is now known about human cognition, judgment, and decision-making. Kahneman's work in these areas led to the development of a dual-mode (or two-system) theoretical model of human thought, with Kahneman labeling the two systems as System 1 and System 2 [42]. In this theory, System 1-based thought is very fast, automatic, and occurs primarily at an unconscious level. System 1 is believed to account for the vast majority of human cognition. By contrast, System 2-based thought is slow, deliberative, and effortful. System 2 is believed to account for virtually all of what is commonly referred to as conscious thought. One of the key insights of Kahneman's theory is that consciousness is a highly limited resource. Put another way, in comparison with System 1, the human brain's capacity for conscious, directed, sentient (*i.e.*, System 2) thought is very limited. In an effort to conserve this limited resource, the brain employs a number of heuristic mechanisms that allow System 1 to handle much of the workload that would otherwise be relegated to System 2. Kahneman's work suggests that one of these mechanisms in particular – substitution – may lie at the very heart of the halo effect.

The substitution heuristic is characterized by the human brain's tendency to substitute or replace a complex or difficult problem with one that is easier to solve, or with one for which it has already arrived at a solution. Halo effect theory, for example, posits that a person's overall opinion of an entity will color that person's judgments of the entity's individual traits. Considering this proposition in the milieu of Kahneman's substitution heuristic, it becomes evident that answering a series of questions about an entity's specific traits requires more deliberation and concentrated thought than does referencing one's existing overall opinion about the entity. Rather than expending its limited System 2 resources, the human brain instead prefers to rely on its existing, pre-computed opinion in order to quickly arrive at an answer that it believes is likely to be *approximately* correct. Thus, when a subject is asked what he thinks about a particular characteristic of a web interface, his brain substitutes a simpler question, such as: "What do I think of this website?" Since System 1 can quickly answer the simpler question by accessing the subject's existing attitudes and opinions toward the website's content, the subject's brain is able to effectively take a shortcut to the solution, thereby avoiding the need to engage in slow, cognitively expensive System 2-based thought.

Despite the appealing explanation for the halo effect that can be derived from Kahneman's theory, the

substitution heuristic is silent with respect to the growing body of empirical evidence that suggests that men are more susceptible to halo error than are women. With respect to this situation, theory from the field of evolutionary psychology suggests that there may be a biological basis for halo-based judgments. Specifically, it has been proposed that survival for our ancient male ancestors was linked to their ability to quickly identify potential comrades, threats, or foes [105]. The ability to judge the nature of a situation quickly within the boundaries of a reasonable margin of error would thus confer a distinct evolutionary advantage on our ancient male forebears, thereby improving their chances of survival in a dangerous and highly dynamic environment. This implies the tantalizing, but as yet untested possibility that the halo effect is an evolved psychological mechanism – an obligate adaptation – that was useful to our ancestors in the unforgiving environment of the ancient world, but which causes unwelcome and detrimental errors in judgment for human beings living in the modern era.

7 SUMMARY, LIMITATIONS, AND CONCLUDING REMARKS

The study described herein inquired into the extent to which evaluations of web interface design are distorted by halo effect-based contamination. For this purpose a controlled, randomized experiment was used that incorporated 3 polarizing issues, 6 different types of websites, 42 web interface design variations, and more than 1,200 research subjects. After controlling for other cognitive biases and several additional factors that are known to influence judgment and decision-making, the quantitative results revealed substantial halo error in assessments of web interface design. Additional analyses subsequently confirmed that these halo effect-based distortions in judgment are, statistically speaking, entirely isolated to men. The results therefore suggest that on average, women's assessments of web interface design traits are more closely aligned with objective ratings than are analogous assessments made by men. These findings parallel past research in which halo error has been found to be prevalent primarily among men, and add to the rapidly growing body of scientific literature that documents important structural, functional, and neural information processing differences between the male and female brains.

There are several limitations to this work that merit acknowledgement. First, the findings were derived only from the responses of adult web users. Indeed, virtually all halo effect-related research has focused on adult judgment and decision-making. Whether this phenomenon extends into the judgment and decision-making capabilities of children is an open question that is meritorious of scientific inquiry. Second, the results are similarly limited by the fact that our research subjects were all English-speaking web users in the United States. Although halo error has been identified in many different countries and cultures, variations in web design are evident according to whether a website's primary language is written from left-to-right, right-to-left, top-to-bottom, etc. [88].

These variations naturally imply that judgments regarding what constitutes high-quality web design may vary from language to language. Third, we relied upon the findings of past research to create experimental conditions that were designed to maximize the halo effect. The extent to which the magnitude of the observed distortions in judgment would vary under conditions that were not intentionally designed to maximize the halo effect remains unknown. Finally, despite our efforts to include many different types of websites and web interface designs in the experiment, we acknowledge that these web-based variations represent only a subset of the immense and ever-changing variety of technology interfaces with which human beings interact. Whether and to what extent the findings reported here are generalizable to application software interfaces, mobile interfaces, game interfaces, etc. remains unknown. Together, these limitations represent rich and fertile ground for future research.

The limitations noted above notwithstanding, there are three key contributions of this work that we would like to emphasize and reiterate. First, the results reveal that the halo effect plays a highly significant role in assessments of web interface design. Since this study represents the first known direct test of the halo effect in the context of interface design, the findings reported here extend the boundaries of halo effect theory into the realm of human-computer interaction, and establish halo error as a legitimate factor for HCI research concerned with judgment and decision-making. In light of the substantial halo-based errors in judgment detected here, the results of past HCI-related research studies that employed multi-item measurement scales may need to be revisited.

Next, this study represents one of the first known direct tests of and evidence for gender effects at the intersection of halo effect theory and information technology. The experimental data reveal strong and highly significant halo effect contamination among men, while no statistically significant halo-based distortions were observed among women. Since this outcome is based on the responses of more than 1,000 research subjects interacting with many different types of websites and dozens of interface design variations, and in light of the size of the observed effect, it is highly unlikely that this gender-based difference can be

attributed to a statistical anomaly. These results can therefore be added to the growing list of gender-based differences in brain structure, function, and neural information processing that have been accumulating in the scientific literature over the past few decades.

Finally, this paper developed an integrated framework that draws on established research in both cognitive and evolutionary psychology to offer a theory-based explanation as to the origins and purpose of the halo effect in human males. This theoretical framework proposes that the halo effect is a manifestation of the human brain's well-known tendency to use substitution heuristics, and that reliance on halo-based judgments evolved as an obligate adaptation in the brains of our ancient male ancestors. In essence, we believe that the halo effect can best be understood as a cognitive trade-off in which the male brain is willing to accept a modest amount of error in the accuracy of its judgments in exchange for the capacity to quickly arrive at conclusions. Although this cognitive trade-off may have limited value in the modern world, when viewed from the perspective of the traditional gender roles and the unforgiving climate that characterized the ancient world, such a trade-off may arguably have conferred a distinct evolutionary advantage on our ancient male ancestors. This proposition, of course, remains to be tested, but if ultimately proven tenable would reveal that the true origins of the halo effect are rooted firmly in biological evolution.

REFERENCES

[1] R. Agarwal and V. Venkatesh, "Assessing a firm's web presence: a heuristic evaluation procedure for the measurement of usability," *Information Systems Research,* vol. 13, pp. 168-186, 2002.

[2] A. M. Aladwania and P. C. Palvia, "Developing and validating an instrument for measuring user-perceived web quality," *Information & Management,* vol. 39, pp. 467-476, 2002.

[3] Anonymous, *Beowulf*: Unknown Publisher, 8th-11th century.

[4] N. Au, E. W. T. Ngai, and T. C. E. Cheng, "Extending the Understanding of End User Information Systems Satisfaction Formation: An Equitable Needs Fulfillment Model Approach," *MIS Quarterly,* vol. 32, pp. 43-66, 2008.

[5] N. E. Beckwith, H. H. Kassarjian, and D. R. Lehmann, "Halo effects in marketing research: Review and prognosis," *Advances in Consumer Research,* vol. 5, pp. 465-467, 1978.

[6] N. E. Beckwith and D. R. Lehmann, "The Importance of Halo Effects in Multi-Attribute Attitude Models," *Journal of Marketing Research,* vol. 12, pp. 265-275, 1975.

[7] P. M. Bentler and D. C. Bonnet, "Significance Tests and Goodness of Fit in the Analysis of Covariance Structures," *Psychological Bulletin,* vol. 88, pp. 588-606, 1980.

[8] H. J. Bernardin, "Effects of rater training on leniency and halo errors in student ratings of instructors," *Journal of Applied Psychology,* vol. 63, pp. 301-308, 1978.

[9] H. J. Bernardin and C. S. Walter, "Effects of rater training and diary-keeping on psychometric error in ratings," *Journal of Applied Psychology,* vol. 62, pp. 64-69, 1977.

[10] D. Bernoulli, *Specimen Theoriae Novae de Mensura Sortis*. Upper Saddle River, NJ: Gregg Press, 1738 [1967].

[11] A. S. Bharadwaj, "A Resource-Based Perspective on Information Technology Capability and Firm Performance: An Empirical Investigation," *MIS Quarterly,* vol. 24, pp. 169-196, 2000.

[12] W. V. Bingham, "Halo, invalid and valid," *Journal of Applied Psychology,* vol. 23, pp. 221-228, 1939.

[13] W. V. Bingham and B. V. Moore, *How to interview*. Oxford, England: Harpers, 1931.

[14] W. C. Borman, "Effects of instructions to avoid halo error on reliability and validity of performance evaluation ratings," *Journal of Applied Psychology,* vol. 60, pp. 556-560, 1975.

[15] B. Brown and S. Perry, "Removing the Financial Performance Halo from Fortune's "Most Admired" Companies," *Academy of Management Journal,* vol. 37, pp. 1347-1359, 1994.

[16] E. M. Brown, "Influence of Training, Method, and Relationship of the Halo Effect," *Journal of Applied Psychology,* vol. 52, pp. 195-199, 1968.

[17] L. Cahill, "His Brain, Her Brain," *Scientific American,* vol. 292, pp. 40-47, 2005.

[18] P. Chandon, V. G. Morwitz, and W. J. Reinartz, "Do Intentions Really Predict Behavior? Self-Generated Validity Effects in Survey Research," *Journal of Marketing,* vol. 69, pp. 1-14, 2005.

[19] S. T. Charles, C. A. Reynolds, and M. Gatz, "Age-related differences and change in positive and negative affect over 23 years," *Journal of Personality and Social Psychology,* vol. 80, pp. 136-151, 2001.

[20] B. Chong and M. Wong, "Crafting an effective customer retention strategy: a review of halo effect on customer satisfaction in online auctions," *International Journal of Management and Enterprise Development,* vol. 2, pp. 12-26, 2005.

[21] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences (2nd Edition)*. Hillsdale, NJ: Lawrence Earlbaum Associates, 1988.

[22] W. T. Coombs and S. J. Holladay, "Unpacking the halo effect: reputation and crisis management," *Journal of Communication Management,* vol. 10, pp. 123-137, 2006.

[23] W. H. Cooper, "Ubiquitous halo," *Psychological Bulletin,* vol. 80, pp. 218-244, 1981.

[24] L. A. Crosby and N. Stephens, "Effects of Relationship Marketing on Satisfaction, Retention, and Prices in the Life Insurance Industry," *Journal of Marketing Research,* vol. 24, pp. 404-411, 1987.

[25] D. Cyr, M. Head, H. Larios, and B. Pan, "Exploring human images in website design: a multi-method approach," *MIS Quarterly,* vol. 33, pp. 539-566, 2009.

[26] F. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly,* vol. 13, pp. 319-340, 1989.

[27] J. Decety, C.-Y. Yang, and Y. Cheng, "Physicians down-regulate their pain empathy response: an event-related brain potential study," *Neuroimage,* vol. 50, pp. 1676-1682, 2010.

[28] L. Deng and M. S. Poole, "Affect in Web Interfaces: A Study of the Impacts of Web Page Visual Complexity and Order," *MIS Quarterly,* vol. 34, pp. 711-730, 2010.

[29] K. Dion, E. Berscheid, and E. Walster, "What is beautiful is good," *Journal of Personality and Social Psychology,* vol. 24, pp. 285-290, 1972.

[30] M. Feldman, I. Feller, J. Bercovitz, and R. Burton, "Equity and the Technology Transfer Strategies of American Research Universities," *Management Science,* vol. 48, pp. 105-121, 2002.

[31] S. Fox, A. Bizman, and E. Herrmann, "The halo effect: Is it a unitary concept?," *Journal of Occupational Psychology,* vol. 56, pp. 289-296, 1983.

[32] F. Fujita, E. Diener, and E. Sandvik, "Gender differences in negative affect and well-being: The case for emotional intensity. ," *Journal of Personality and Social Psychology,* vol. 61, pp. 427-434, 1991.

[33] D. Gefen, E. Karahanna, and D. W. Straub, "Trust and TAM in online shopping: an integrated model.," *MIS Quarterly,* vol. 27, pp. 51-90, 2003.

[34] B. A. Grove and W. A. Kerr, "Specific Evidence on Origin of Halo Effect in Measurement of Employee Morale," *The Journal of Social Psychology,* vol. 34, pp. 165-170, 1951.

[35] J. Hartmann, A. Sutcliffe, and A. D. Angeli, "Investigating attractiveness in web user interfaces," in *ACM SIGCHI Conference on Human Factors in Computing Systems*, San Jose, CA, 2007, pp. 387-396.

[36] M. B. Holbrook, "Using a Structural Model of Halo Effect to Assess Perceptual Distortion Due to Affective Overtones," *Journal of Consumer Research,* vol. 10, pp. 247-252, 1983.

[37] G. Hooks and C. Mosher, "Outrages Against Personal Dignity: Rationalizing Abuse and Torture in the War on Terror," *Social Forces,* vol. 83, pp. 1627-1645, 2005.

[38] D. Hooper, J. Coughlan, and M. Mullen, "Structural Equation Modelling: Guidelines for Determining Model Fit," *Electronic Journal of Business Research Methods,* vol. 6, pp. 53-60, 2008.

[39] G. R. Iversen and H. P. Norpoth, *Analysis of Variance*. Thoisand Oaks, CA: Sage Publications, 1987.

[40] D. M. Johnson and R. N. Vidulich, "Experimental manipulation of the halo effect," *Journal of Applied Psychology,* vol. 40, pp. 130-134, 1956.

[41] M. M. S. Johnson, "Age Differences in Decision Making: A Process Methodology for Examining Strategic Information Processing," *Journal of Gerontology,* vol. 45, pp. 75-78, 1990.

[42] D. Kahneman, *Thinking, Fast and Slow*. New York, NY: Farrar, Straus, and Giroux, 2011.

[43] R. M. Kaplan, "Is beauty talent? Sex interaction in the attractiveness halo effect," *Sex Roles,* vol. 4, pp. 195-204, 1978.

[44] R. J. Kauffman and B. Wang, "New Buyers' Arrival Under Dynamic Pricing Market Microstructure: The Case of Group-Buying Discounts on the Internet," *Journal of Management Information Systems,* vol. 18, pp. 157-188, 2001.

[45] S. Keen, *Faces of the Enemy: Reflections of the Hostile Imagination*. Harper & Row: San Francisco, CA, 1991.

[46] S. S. Kim, "The Integrative Framework of Technology Use: An Extension and Test," *MIS Quarterly*, vol. 33, pp. 513-538, 2009.

[47] B. Kolb and I. Q. Whishaw, *Fundamentals of Human Neuropsychology*. New York, NY: Worth Publishers, 2008.

[48] D. Landy and H. Sigall, "Beauty is talent: Task evaluation as a function of the performer's physical attractiveness," *Journal of Personality and Social Psychology*, vol. 29, pp. 299-304, 1974.

[49] Y. Lee and A. N. K. Chen, "Usability Design and Psychological Ownership of a Virtual World," *Journal of Management Information Systems*, vol. 28, pp. 269-307, 2011.

[50] L. Leuthesser, C. S. Kohli, and K. R. Harich, "Brand equity: the halo effect measure," *European Journal of Marketing*, vol. 29, pp. 57-66, 1995.

[51] R. W. Levenson, L. L. Carstensen, and J. M. Gottman, "Influence of age and gender on affect, physiology, and their interrelations," *Journal of Personality and Social Psychology*, vol. 67, pp. 56-68, 1994.

[52] C. Liu and K. P. Arnett, "Exploring the factors associated with Web site success in the context of electronic commerce," *Information & Management*, vol. 38, pp. 23-33, 2000.

[53] G. W. Lucker, W. E. Beane, and R. L. Helmreich, "The Strength of the Halo Effect in Physical Attractiveness Research," *The Journal of Psychology: Interdisciplinary and Applied*, vol. 107, pp. 69-75, 1981.

[54] S. E. MacPherson, L. H. Phillips, and S. Della Sala, "Age, executive function and social decision making: A dorsolateral prefrontal theory of cognitive aging," *Psychology and Aging*, vol. 17, pp. 598-609, 2002.

[55] A. P. Massey and M. M. Montoya-Weiss, "Unraveling the Temporal Fabric of Knowledge Conversion: A Model of Media Selection and Use," *MIS Quarterly*, vol. 30, pp. 99-114, 2006.

[56] J. R. Meindl and S. B. Ehrlich, "The Romance of Leadership and The Evaluation of Organizational Performance," *Academy of Management Journal*, vol. 30, pp. 91-109, 1987.

[57] J. R. Meindl, S. B. Ehrlich, and J. M. Dukerich, "The Romance of Leadership," *Administrative Science Quarterly*, vol. 30, pp. 78-102, 1985.

[58] S. Mithas, N. Ramasubbu, and V. Sambamurthy, "How Information Management Capability Influences Firm Performance," *MIS Quarterly*, vol. 35, pp. 237-256, 2011.

[59] J.-W. Moon and Y.-G. Kim, "Extending the TAM for a World-Wide-Web context," *Information & Management*, vol. 38, pp. 217-230, 2001.

[60] D. S. Morse, E. A. Edwardsen, and H. S. Gordon, "Missed Opportunities for Interval Empathy in Lung Cancer Communication," *Journal of the American Medical Association: Internal Medicine*, vol. 168, pp. 1853-1858, 2008.

[61] T. Mouw and M. E. Sobel, "Culture Wars and Opinion Polarization: The Case of Abortion," *American Journal of Sociology*, vol. 106, pp. 913-943, 2001.

[62] D. K. Mroczek, "Age and Emotion in Adulthood," *Current Directions in Psychological Science*, vol. 10, pp. 87-90, 2001.

[63] D. K. Mroczek and C. M. Kolarz, "The effect of age on positive and negative affect: A developmental perspective on happiness," *Journal of Personality and Social Psychology*, vol. 75, pp. 1333-1349, 1998.

[64] L. Newton, *Illegal, Alien, or Immigrant: The Politics of Immigration Reform*. New York, NY: New York University Press, 2008.

[65] R. E. Nisbett and T. D. Wilson, "The halo effect: Evidence for unconscious alteration of judgments," *Journal of Personality and Social Psychology*, vol. 35, pp. 250-256, 1977.

[66] R. G. O'Brien, "A simple test for variance effects in experimental designs," *Psychological Bulletin*, vol. 89, pp. 570-574, 1981.

[67] E. O'Donnell and J. J. Schultz, "The Halo Effect in Business Risk Audits: Can Strategic Risk Assessment Bias Auditor Judgment about Accounting Details?," *The Accounting Review*, vol. 80, pp. 921-939, 2005.

[68] J. W. Palmer, "Web Site Usability, Design, and Performance Metrics," *Information Systems Research*, vol. 13, pp. 151-167, 2002.

[69] D. V. Parboteeah, J. S. Valacich, and J. D. Wells, "The Influence of Website Characteristics on a Consumer's Urge to Buy Impulsively "

Information Systems Research, vol. 20, pp. 60-78, 2009.

[70] P. A. Pavlou, H. Liang, and Y. Xue, "Understanding and mitigating uncertainty in online exchange relationships: A principal-agent perspective," *MIS Quarterly*, vol. 31, pp. 105-136, 2007.

[71] Pew Research Center., *Who's Online: Internet User Demographics*. Washington, D.C.: Internet & American Life Project, 2012.

[72] P. M. Podsakoff, S. B. MacKenzie, J.-Y. Lee, and N. P. Podsakoff, "Common method biases in behavioral research: A critical review of the literature and recommended remedies," *Journal of Applied Psychology*, vol. 88, pp. 879-903, 2003.

[73] M. Powell and D. Ansic, "Gender differences in risk behaviour in financial decision-making," *Journal of Economic Psychology*, vol. 18, pp. 605-628, 1997.

[74] R. Rasch and H. Tosi, "Factors Affecting Software Developers' Performance: An Integrated Approach," *MIS Quarterly*, vol. 16, pp. 395-413, 1992.

[75] H. H. Remmers, "Reliability and halo effect of high school and college students' judgments of their teachers," *Journal of Applied Psychology*, vol. 18, pp. 619-630, 1934.

[76] R. Riedl, M. Hubert, and P. Kenning, "Are there neural gender differences in online trust? an fMRI study on the perceived trustworthiness of ebay offers," *MIS Quarterly*, vol. 34, pp. 397-428, 2010.

[77] P. Rosenzweig, "Misunderstanding the Nature of Company Performance: The Halo Effect and Other Business Delusions," *California Management Review*, vol. 49, pp. 6-20, 2007.

[78] P. Rosenzweig, *The Halo Effect... and the Eight Other Business Delusions That Deceive Managers*. New York, NY: Free Press, 2009.

[79] N. Sahoo, R. Krishnan, G. Duncan, and J. Callan, "The Halo Effect in Multicomponent Ratings and Its Implications for Recommender Systems: The Case of Yahoo! Movies," *Information Systems Research*, vol. 23, pp. 231-246, 2012.

[80] E. Salas, M. A. Rosen, and D. DiazGranados, "Expertise-Based Intuition and Decision Making in Organizations," *Journal of Management*, vol. 36, pp. 941-973, 2010.

[81] R. Santhanam and E. Hartono, "Issues in Linking Information Technology Capability to Firm Performance," *MIS Quarterly*, vol. 27, pp. 125-153, 2003.

[82] W. E. Saris and I. N. Gallhofer, *Design, Evaluation, and Analysis of Questionnaires for Survey Research*. Hoboken, NJ: John Wiley & Sons, 2007.

[83] J. M. Sattler, W. A. Hillix, and L. A. Neher, "Halo effect in examiner scoring of intelligence test responses," *Journal of Consulting and Clinical Psychology*, vol. 34, pp. 172-176, 1970.

[84] R. Sharma, P. Yetton, and J. Crawford, "Estimating the Effect of Common Method Variance: The Method-Method Pair Technique with an Illustration from TAM Research," *MIS Quarterly*, vol. 33, pp. 473-490, 2009.

[85] W. D. Sine, S. Shane, and D. Di Gregorio, "The Halo Effect and Technology Licensing: The Influence of Institutional Prestige on the Licensing of University Inventions," *Management Science*, vol. 49, pp. 478-496, 2003.

[86] D. S. Soper. (2016, September 16). *A-priori Sample Size Calculator for Structural Equation Models [Software]*. Available: https://www.danielsoper.com/statcalc

[87] D. S. Soper, "User Interface Design and the Halo Effect: Some Preliminary Evidence," presented at the Twentieth Americas Conference on Information Systems (AMCIS), Savannah, GA, 2014.

[88] D. S. Soper and S. Mitra, "An Inquiry into Mental Models of Web Interface Design," in *19th Americas Conference on Information Systems (AMCIS)*, Chicago, Illinois, 2013.

[89] Y. Sun, A. Bhattacherjee, and Q. Ma, "Extending technology usage to work settings: The role of perceived work compatibility in ERP implementation," *Information & Management*, vol. 46, pp. 351-356, 2009.

[90] B. Szajna, "Empirical Evaluation of the Revised Technology Acceptance Model," *Management Science*, vol. 42, pp. 85-92, 1996.

[91] R. N. Taylor, "Age and Experience as Determinants of Managerial Information Processing and Decision Making Performance," *Academy of Management Journal*, vol. 18, pp. 74-81, 1975.

[92] E. L. Thorndike, "A constant error in psychological ratings," *Journal of Applied Psychology*, vol. 4, pp. 25-29, 1920.

[93] E. Tulving, D. L. Schacter, and H. A. Stark, "Priming Effects in Word Fragment Completion are independent of Recognition Memory," *Journal of Experimental Psychology*, vol. 8, pp. 336-342, 1982.

[94] J. Van Doom, "Is There a Halo Effect in Satisfaction Formation in Business-to-Business Services?," *Journal of Service Research*, vol. 11, pp. 124-141, 2008.

[95] V. Venkatesh and M. G. Morris, "Why Don't Men Ever Stop to Ask for Directions? Gender, Social Influence, and Their Role in Technology Acceptance and Usage Behavior," *MIS Quarterly*, vol. 24, pp. 115-139, 2000.

[96] V. Venkatesh, M. G. Morris, and P. L. Ackerman, "A Longitudinal Field Investigation of Gender Differences in Individual Technology Adoption Decision-Making Processes," *Organizational Behavior and Human Decision Processes*, vol. 83, pp. 33-60, 2000.

[97] G. R. Wark and D. L. Krebs, "Gender and dilemma differences in real-life moral judgment," *Developmental Psychology*, vol. 32, pp. 220-230, 1996.

[98] F. L. Wells, *A statistical study of literary merit*. New York, NY: The Science Press, 1907.

[99] J. D. Wells, J. S. Valacich, and T. J. Hess, "What Signals Are You Sending? How Website Quality Influences Perceptions of Product Quality and Purchase Intentions," *MIS Quarterly*, vol. 35, pp. 373-396, 2011.

[100] J. C. Westland, "Lower bounds on sample size in structural equation modeling," *Electronic Commerce Research and Applications*, vol. 9, pp. 476-487, 2010.

[101] C. G. Wetzel, T. D. Wilson, and J. Kort, "The halo effect revisited: Forewarned is not forearmed," *Journal of Experimental Social Psychology*, vol. 17, pp. 427-439, 1981.

[102] A. Winkler, *Gunfight: The Battle over the Right to Bear Arms in America*. New York, NY: W. W. Norton & Company, 2011.

[103] J. Wirtz, "An examination of the presence, magnitude and impact of halo on consumer satisfaction measures," *Journal of Retailing and Consumer Services*, vol. 7, pp. 89-99, 2000.

[104] J. Wirtz and J. E. G. Bateson, "An experimental investigation of halo effects in satisfaction measures of service attributes," *International Journal of Service Industry Management*, vol. 6, pp. 84-102, 1995.

[105] L. Workman and W. Reader, *Evolutionary Psychology: An Introduction*, 2nd ed. Cambridge, UK: Cambridge University Press, 2008.

[106] B. T. W. Wu and S. M. Petroshius, "The halo effect in store image measurement," *Journal of the Academy of Marketing Science*, vol. 15, pp. 44-51, 1987.

AUTHOR BIOGRAPHIES

Dr. Daniel S. Soper is an Associate Professor of Information Systems and Decision Sciences in the Mihaylo College of Business and Economics at California State University, Fullerton. His current research interests lie in the realms of human cognition, interface design, and innovative applications of computational linguistics. His research has appeared in a wide variety of leading information systems and computer science journals and conference proceedings. In his free time he enjoys travelling, painting, building websites and software, and attempting to teach himself to play the piano.

Dr. Farnaz Piepkorn is a Training Coordinator and Technical Trainer at Northrop Grumman Corporation. She has over fifteen years of experience in curriculum development, training coordination, and instruction of computer applications and business courses in a wide variety of academic and business environments. In addition to her professional duties at Northrop Grumman, she enjoys teaching a variety of university courses, including computer-based applications, statistics, management, college survival skills, and business writing.