Calibration of consumer knowledge of the web

Kishore Gopalakrishna Pillai a,⁎, Charles Hofacker b,⁎

a Department of Marketing, Leeds University Business School, University of Leeds, Leeds, LS2 9JT, United Kingdom
b Department of Marketing, College of Business, Florida State University, Tallahassee, FL 32306, USA

Abstract

Calibration of consumer knowledge of the web refers to the correspondence between accuracy and confidence in knowledge of the web. Being well-calibrated means that a person is realistic in his or her assessment of the level of knowledge that he or she possesses. This study finds that involvement leads to better calibration and that calibration is higher for procedural knowledge and common knowledge, as compared to declarative knowledge and specialized knowledge. Neither usage, nor experience, has any effect on calibration of knowledge of the web. No difference in calibration is observed between genders. But, in agreement with previous findings, this study also finds that males are more confident in their knowledge of the web. The results point out that calibration could be more a function of knowledge-specific factors and less that of individual-specific factors. The study also identifies flow and frustration with the web as consequences of calibration of knowledge of the web and draws the attention of future researchers to examine these aspects.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Calibration; Knowledge; Web; Involvement; Knowledge type

1. Introduction

Consumer knowledge has been a topic of considerable research interest in marketing for the past three decades (Alba & Hutchinson, 1987; Bettman & Park, 1980; Brucks, 1985; Mitchell & Dacin, 1996; Roy & Cornwell, 2004). The importance of consumer knowledge can be located in the fact that knowledge is central to understanding consumer behaviors such as information search and information processing. Researchers have examined the antecedents and consequences of prior knowledge of consumers (Alba, 1983; Chase & Simon, 1973) and the different types of consumer knowledge such as product class knowledge, price knowledge, knowledge of the World Wide Web, etc. (Brucks, 1985; Estelami, Lehmann, & Holden, 2001; Page & Uncles, 2004). Researchers have also attempted to distinguish between objective knowledge, which refers to the absolute knowledge possessed by the consumer, and subjective knowledge, which refers to consumers’ perception of their own knowledge (Raju, Lonial, & Mangold, 1995).

An aspect of knowledge that has been omitted from the discourse in consumer research is knowledge calibration. Knowledge calibration refers to the correspondence between accuracy and confidence in knowledge. Notions of correct and incorrect knowledge tap into only the first dimension, i.e., accuracy. A person can be well calibrated even when possessing inaccurate knowledge. Being well-calibrated means that a person is realistic in his or her assessment of the level of knowledge that he or she possesses. This, in turn, is likely to facilitate appropriate actions such as information search or optimal decision making. Hence, examination of this construct is likely to unravel processes that have not been brought to light by the research on consumer knowledge. The latter construct has attracted considerable research, while research on calibration of consumer knowledge is relatively new. Alba and Hutchinson (2000) highlighted the importance of this construct and its implications in consumer research and called for research in this area. This paper responds to the call by examining the calibration of consumer knowledge of the web, a topic that is of considerable relevance, given the increasing importance of the web for consumers. The paper proposes a model of calibration of knowledge of the web and examines the...
antecedent roles of involvement with the web, experience, usage, knowledge type, and gender on calibration of consumer knowledge of the web.

The paper’s contribution is threefold. One, it seeks to empirically examine knowledge calibration in the consumer domain. Second, in doing so, it adds a new dimension to the research on knowledge and usage of the web. Third, it contributes to the larger stream of research on knowledge calibration through the examination of knowledge type as an antecedent of calibration.

We start by introducing the concept of knowledge calibration, followed by a brief review of the literature on knowledge calibration. The hypotheses for the study are presented afterwards. Subsequently, the method and the results of the study are presented. The results indicate the antecedent effects of involvement and knowledge type in enhancing calibration. Broadly, the results imply that calibration of knowledge of the web in particular and calibration of knowledge in general could be determined by knowledge variables such as types of knowledge. The study discounts the role of individual specific factors, other than involvement, as antecedents of knowledge calibration.

2. Literature review and conceptual framework

2.1. Knowledge calibration

As defined earlier, calibration of knowledge refers to the correspondence between accuracy and confidence in knowledge. High accuracy and high confidence together lead to high calibration. Similarly, low accuracy and low confidence lead to high calibration. Lack of correspondence between the two lead to low calibration; thus, low accuracy along with high confidence (overconfidence), and high accuracy along with low confidence (underconfidence) are instances of poor calibration, termed as miscalibration. The following diagram, adapted from Goldsmith and Pillai (2006), captures the above distinctions in a 2 × 2 matrix.

Several scientists, including statisticians, meteorologists, and psychologists, have sought to address the measurement and explanation of judgments of confidence and their relation to accuracy (e.g., Gigerenzer, Hoffrage, & Kleinbolting, 1991; Harvey, 1997; Yates, 1990). A recurring finding from many of these studies is that people are systematically overconfident about the accuracy of their knowledge and judgment (Pillai & Goldsmith, 2006). Overconfidence can be defined as the overestimation of the probability of an event. Thus, in a knowledge assessment task, if the subject estimates that 70% of his/her answers are correct, but only 50% are actually so, the subject is overconfident.

Reviewing research on calibration, Spence (1996) states that much research indicates that individuals are overconfident, though certain professions have displayed high calibration in judgments. For example, weather forecasters and odds-makers (those who calculate and set betting odds based on the prediction of the result of a contest, such as a horserace or an election) have been shown to be well calibrated (Hoerl & Fallin, 1974; Murphy & Winkler, 1977), whereas doctors’ diagnoses have not (Christen-Szalanski & Busyhead, 1981; Oskamp, 1962). Explanations for this disparate finding include (a) the repetitiousness of forecasting with immediate, unambiguous feedback which enables validation and correction, facilitating calibration (Lichtenstein, Fischoff, & Philips, 1982), (b) different payoff matrices or loss functions for forecasters and doctors, whereby doctors perceive the penalty for false negative as more devastating and hence state their opinions with greater conviction (Keren, 1991; O’Connor, 1989). The several studies across various domains point to the fact that high calibration is rarely achieved and moderate levels that include some degree of systematic bias, usually overconfidence, are the norm (Pillai & Goldsmith, 2006).

Noteworthy among other domains that have examined the issue of the confidence–accuracy relationship is forensic psychology. In criminal investigations, witnesses often attempt to identify a culprit and give a statement about the certainty of the identification, expressed as a confidence judgment. It has been demonstrated that a confident witness is regarded as a credible witness, with a great impact on the outcome of a trial (Olsson, 2000). Yet many experimental studies of witness identification have shown that the confidence–accuracy relationship is not reliably different from zero (Loftus, 1974; Narby, Cutler, & Penrod, 1996). Consequently, the issue of calibration or “realism in confidence assessments” has attracted considerable scholarly attention in this field.

As mentioned earlier, this study is an attempt to apply the construct of knowledge calibration to consumer research. In this study, we examine the calibration of consumers’ knowledge of the web. The latter construct is dealt with in the following section.

2.2. Consumers’ knowledge of the web

As noted earlier, examination of consumer knowledge has a rich tradition in the marketing literature. Specifically, the influence of product knowledge on consumers’ information processing has been of great interest to researchers (e.g., Alba & Hutchinson, 1987; Bettman & Park, 1980; Brucks, 1985). Product knowledge has been found to be influential in consumers’ product evaluations (Blair & Innis, 1996), precision of price estimates (Biswas & Sherrell, 1993), and processing of advertising messages (Maheswaran & Stemthal, 1990). Research on consumers’ knowledge of the web is relatively recent, as the web itself is relatively recent.

During the past decade, pre-existing firms have sought to realize benefits for themselves and for their customers online by providing alternative purchasing channels, inexpensive self-service options, complementary services, longer product lines, recommendation agents and interactive advertisements. Other firms have sprung up to provide new services like online auctions, interactive driving directions, video upload and exchange and e-mail. While we tend to think of virtual channels as being less expensive than physical channels, these new interfaces have their own costs in terms of the consumer knowledge that is required. Not all consumers are ready for this onslaught of
technology (Parasuraman, 2000). The examination of consumer knowledge of the web and the calibration of such knowledge becomes relevant in this context.

Knowledge has been defined as the body of facts and principles (i.e., information and understanding) accumulated by mankind (i.e., stored in memory) about a domain (Delbridge & Bernard, 1998). It follows that consumers’ knowledge of the web refers to the information and understanding of the web stored in the consumers’ memory. Examples of consumer knowledge of the web include information stored in a consumer’s memory about how to use a search engine, what cookies are, or the benefit of “what’s new” links.

Researchers in the past have used proxy measures to infer consumer knowledge of the web, such as number of hours spent online or months and years that has been used (Diaz, Hammond, & McWilliam, 1997). Self-perception judgments of Web skill have also been used to measure consumer skill on the web (Novak, Hoffman, & Yung, 2000). Recently, Page and Uncles (2004) proposed a scale to measure consumers’ web knowledge content. Because the vast array of knowledge that consumers could possess about the web imposes limitations on valid measurement, these authors restricted their focus to knowledge arising from machine interactivity of the web. While person-to-person interaction occurs frequently online, we also restrict our study to machine interactivity. The antecedent role of knowledge type is one of the key contributions of this study, and we chose those distinctions that are of considerable import to the study of web knowledge.

Fig. 2 shows the model of the antecedents of calibration of consumer knowledge of the web. We start with the discussion of individual-specific factors.

2.3. Antecedents of calibration of knowledge of the web

We discuss two different types of antecedent variables: (a) individual-specific factors such as involvement, experience, usage, and gender, and (b) knowledge-specific factors which includes the distinctions between declarative and procedural knowledge on the one hand and common and specialized knowledge on the other. The selection of the antecedents was determined, in part, by the research focus on these constructs. For example, involvement is a variable that has attracted considerable attention as a determinant of consumer behavior. As discussed below, experience, usage, and gender have also been shown to be of relevance to the stream of research on knowledge.

As detailed in the Measures section, we use the scale proposed by Page and Uncles (2004) to measure consumers’ knowledge of the web. In the following paragraphs, we discuss each of the antecedent variables of the study and offer propositions. A model of antecedents of calibration of knowledge is proposed. While the model is not comprehensive, many of the more important variables that affect calibration are represented. This is followed by a test of the model.

It should be noted that the hypotheses discussed in the following sections are derived from the interpretation of knowledge calibration as the realism of perceptions on knowledge. The four quadrants presented earlier (see Fig. 1) are not dealt with separately, for two reasons. The first is that accuracy (at an aggregate, between subjects, level) and confidence are continuous variables, and the dichotomous division was intended for conceptual clarification. More importantly, respondents in the study exhibited high levels of accuracy, rendering the distinction between the two types of calibration redundant.

2.3.1. Involvement

Involvement has been defined in terms of personal relevance of the product or the advertisement to the subject. Zaichkowski (1985) defines involvement as a person’s perceived relevance of the object based on inherent needs, values, and interests. Involvement is generally considered to be a function of (a) individual characteristics such as interests, values, and goals, (b) situational factors such as the purchase occasion or perceived risk, and (c) characteristics of the object or stimulus (Laurent & Kapferer, 1985). Outcomes associated with high involvement include more time and effort spent in search-related activities (Bloch, Sherrel, & Ridgway, 1986), more extensive decision making, greater perceived differences in product attributes, and a greater likelihood of establishing brand preferences (Zaichkowski, 1985).

The effects of involvement on the processes mediating the effects of a persuasive communication about an issue have been extensively investigated (Maheswaran & Sternthal, 1990; Miniard, Bhatia, Lord, Dickson, & Unnava, 1991; Okechuku, 1992; Verlegh, Steenkamp, & Meulenberg, 2005). A robust finding of this literature is that uninvolved message recipients, compared to those who are highly involved, are characterized by (a) less attention to and less cognitive elaboration of attribute information, and (b) more reliance on peripheral cues available in the message (Celsi & Olson, 1988; Batra & Ray, 1985; Petty,
Cacioppo, & Schumann, 1983). Starting with Ray et al. (1973), large numbers of studies have reported that consumers involved in a situation or product are more active processors of cognitive information about it. Celsi and Olson (1988) tested several hypotheses concerning the effects of felt involvement on the amount of attention and comprehension effort, the focus of attention and comprehension processes, and the extent of cognitive elaboration during comprehension. These authors found that involved consumers attend to and comprehend more information about shopping situations and products and produce more elaborate meanings and inferences about them (Celsi & Olson, 1988).

Research in memory-based product judgments has shown that individuals are cognitive misers who attempt to minimize search effort (Wyer & Srull, 1986), and involvement during a memory-based judgment influences this processing objective and affects search intensity (Park & Hastak, 1994). Compared to uninvolved consumers, involved consumers are more motivated to form a relatively accurate judgment and hence they search more intensely for judgment-relevant information.

The above discussion highlights the role of involvement in enhancing effort by consumers. Involvement can also be expected to increase familiarity with the domain, and practice in dealing with specific issues concerning the domain. Effort and practice are cues associated with performance, which refers to knowledge accuracy and correct assessment of confidence in a testing task where accuracy and confidence are measured, as in the case of this study. More involved consumers are likely to collect more information about the product, and will be more motivated to form realistic and correct notions regarding their knowledge of the domain. Hence,

**H1.** Higher the involvement with the world wide web, higher the calibration of knowledge of the web.

### 2.3.2. Cumulative online experience

Consumers’ ability to understand and represent web-based information is structured and constrained according to their existing domain experience (Moreau, Lehmann, & Markman, 2001), which is driven by their tenure online, or duration of time they have used the web. Consumer experience may include both direct experiences such as information searching, evaluation, purchase, and consumption of interactive products, and indirect experiences such as advertising exposure and the observation of others’ consumption (Alba & Hutchinson, 1987).

Considerable evidence suggests that cumulative consumer experience increases knowledge (e.g., Park, Mothersbaugh, & Feick, 1994), and thus some researchers measured experience as a proxy of consumer knowledge (e.g., Bettman & Park, 1980). Alba and Hutchinson (1987) proposed that consumer knowledge consists of two dimensions — familiarity and expertise. The existing literature indicates that domain experience is...
closely related to both familiarity and expertise. Consumer experience is considered as a direct antecedent to familiarity (Alba & Hutchinson, 1987), and research has shown that experience is related to expertise as well (Mitchell & Dacin, 1996). We argue that the effect of experience on familiarity will be stronger than the effect on expertise, leading to the feeling of knowing. Of course expertise should increase as a function of cumulative time logged online, but not nearly as fast as their sense of familiarity. Research on the concept of confidence has examined practice – a variable closely related to experience – supports this argument. This research suggests that practice produces dissociation between confidence and performance because of elevations in confidence that are unaccompanied by corresponding increases in accuracy (Paese & Sniezek, 1991). Similar themes were echoed by Fischoff and Slovic (1980), who argued that even minimal familiarity with a task or tentative rules of performance in the task can produce surprising amounts of conviction, and by Heath and Tversky (1991), who noted that familiarity and knowledge are accompanied by feelings of competence. In the information systems field, research has recorded the role of experience in increasing users’ confidence in their ability to master and use computers supporting their task (DeLone, 1988). Based on these, we reason that experience will be positively related to confidence in knowledge, but will not be related to calibration, since it produces a dissociation between confidence and accuracy.

**H2.** Higher the experience, higher the level of confidence in knowledge.

### 2.3.3. Weekly usage

Here, usage refers to the extent of current weekly use of the World Wide Web. The dominant paradigm that informs us of adoption of the web is the technology acceptance model (TAM). Recent research has employed TAM to explain continued usage of the web (Kim & Malhotra, 2005). In the context of web usage, TAM suggests that there exists a direct and positive effect between attitude towards web usage, usage intention, and actual usage. Perceived usefulness and ease of use determine the attitudes toward using the web. In turn, usage intentions are determined by these attitudes and perceived usefulness. Finally, usage intentions lead to usage. These relationships have found strong empirical support (Davis, 1993; Venkatesh & Davis, 2000). However, a significant body of theoretical evidence underscores the importance of intrinsic motivators in web usage. Researchers have become increasingly aware of the relevance of the non-extrinsic motives of use such as intrinsically enjoyable experiences (i.e., flow) in understanding attitudes and behaviors. Flow is defined as the holistic sensation that people feel when they act with total involvement (Csikszentmihalyi, 1977). Flow is a positive, highly enjoyable state of consciousness that occurs when our perceived skills match the perceived challenges we are undertaking. When this occurs, an individual derives intrinsic enjoyment from the activity and tends to continue with it. If the task is too easy, the person becomes bored. If the work demands skills beyond the capabilities of the individual, anxiety is created. When goals are clear, and the individual’s abilities are up to the challenge, then he or she becomes involved in the activity and is intrinsically motivated.

Thus, it can be seen that TAM posits a goal-directed, extrinsically motivated route to usage of the web, while the flow theory argues the case of an intrinsically motivated path. We argue that the extrinsically motivated route is likely to operate in a manner similar to experience. That is, higher usage will lead to people developing a feeling of knowing, more than what is warranted by an actual increase in knowledge. While the accuracy increases, calibration does not. But for the intrinsically motivated people, higher usage will lead to, as well as flow from, higher involvement, and this is likely to lead to more realistic assessments of their knowledge of the web; i.e., better calibration. Examining the disparate effects of different types of motivation is beyond the scope of this research. Previous research has noted that some people are motivated by extrinsic factors while others by intrinsic factors (Darley, 1999), and often both coexist in individuals (Roberts, Hann, & Slaughter, 2006). Given the null effect from extrinsically motivated individuals and the positive effect from intrinsically motivated individuals, we expect a weak positive effect of usage on calibration of consumer knowledge.

**H3.** Higher the usage, higher the calibration of knowledge of the web.

#### 2.3.4. Gender

Research has recorded the role of gender on confidence. Males have been found to have higher confidence than females in various domains of knowledge (Estes & Hosseini, 1988; Goldsmith & Goldsmith, 1997; Kalaian & Freeman, 1994). Estes and Hosseini (1988) examined confidence in investment decisions in the stock market and found that men are more confident than women. Sutton’s (1987) research on the illusion of control in insurance also corroborates these findings. Sutton (1987) found that women are better prospects than men for auto insurance because of a more realistic perception of their risk of auto accidents. We generalize from these findings and expect males to have higher confidence in their knowledge of the web. Also, previous research has noted a gender gap in computer literacy in which males are more literate than females (Chen, 1986; Sachs & Bellisimo, 1993). It has also been found that, in general, men tend to have more favorable attitudes toward IT (Schumacher & Morahan-Martin, 2001; Simon, 2001). Some research has also noted that women tend to be less skilled in IT (Harrison & Rainer, 1992). In the light of this substantial body of finding, we expect males to have higher accuracy than females in their knowledge of web. Higher accuracy along with higher confidence results in no difference in calibration between genders. Hence,

**H4.** Males will have higher confidence in their knowledge of the web, compared to females.

Following the examination of individual-specific factors, we examine the antecedent effects of knowledge-specific factors (refer Fig. 2).
2.3.5. Knowledge type — procedural versus declarative

Based on the literature from cognitive sciences and marketing, a distinction can be drawn between procedural and declarative knowledge. Declarative knowledge is defined as the ‘factual information that is somewhat static in nature which is usually describable’ (Best, 1989, p. 7); for instance, information about the attributes, terminology, evaluative criteria, facts, and usage situations of the domain of interest (Brucks, 1986). Procedural knowledge, by contrast, refers to the dynamic information underlying skillful actions (Best, 1989; p 7); that is, the knowledge of rules or procedures for taking action believed to be stored and organized into production systems (Brucks, 1986).

Anderson (1983) discussed the impact of these different types of knowledge content on behavior, suggesting that although both types can guide behavior, procedural knowledge content is considered to be of greater influence on actual behavior. For example, to send an e-mail, knowing how to send an e-mail (procedural) will have more influence on the final behavior than knowing what an e-mail is (declarative).

Research on procedural and declarative information in software manuals has found that users prefer procedural to declarative information (Ummelen, 1997). Such preference for procedural information is likely to result in greater familiarity with procedural information and better realism in perceptions regarding procedural knowledge, compared to declarative knowledge.

Procedural knowledge has been recognized as a link between experience and performance (e.g., Bonner & Walker, 1994). Anderson’s (1983) ACT* theory (ACT is not an acronym, though it has been translated as Adaptive Control of Thought), a general theory of cognition that focuses on memory processes, offers a view that practice of a cognitive task results in procedural knowledge, which enhances performance. In the context of knowledge of the web, practice will also help verify the correctness or incorrectness of the knowledge, since incorrect knowledge will hinder effective navigation of the web. That is, web users operate in an environment where immediate and real-time feedback on their procedural knowledge is available. Such feedback is not available in the case of declarative knowledge of the web. Research has recorded the role of feedback in a feedback is not available in the case of declarative knowledge, whereas immediate feedback on their procedural knowledge is available. Such knowledge of rules or procedures for taking action believed to be stored and organized into production systems (Brucks, 1986).

Anderson (1983) discussed the impact of these different types of knowledge content on behavior, suggesting that although both types can guide behavior, procedural knowledge content is considered to be of greater influence on actual behavior. For example, to send an e-mail, knowing how to send an e-mail (procedural) will have more influence on the final behavior than knowing what an e-mail is (declarative).

Procedural knowledge has been recognized as a link between experience and performance (e.g., Bonner & Walker, 1994). Anderson’s (1983) ACT* theory (ACT is not an acronym, though it has been translated as Adaptive Control of Thought), a general theory of cognition that focuses on memory processes, offers a view that practice of a cognitive task results in procedural knowledge, which enhances performance. In the context of knowledge of the web, practice will also help verify the correctness or incorrectness of the knowledge, since incorrect knowledge will hinder effective navigation of the web. That is, web users operate in an environment where immediate and real-time feedback on their procedural knowledge is available. Such feedback is not available in the case of declarative knowledge of the web. Research has recorded the role of feedback in improving calibration (e.g., Lichtenstein et al., 1982). Higher calibration has been observed in certain professions, such as weather forecasting, where immediate feedback is available. Given the availability of feedback for procedural knowledge of the web, we posit that web users will be better calibrated in their procedural knowledge compared to declarative knowledge. Therefore,

H5. Calibration will be higher for procedural knowledge of the web, compared to declarative knowledge.

2.3.6. Knowledge type — common versus specialized

Knowledge literature also distinguishes between experts and novices, and the familiar and the unfamiliar (Alba & Hutchinson, 1987; Page & Uncles, 2004). This distinction leads to two types of knowledge — common and specialized. Common knowledge can be defined as general and/or publicly known information of the domain of interest required to perform general and common domain related tasks successfully. Conversely, specialized knowledge is defined as skilled and/or extraordinary information about facts, terms, attributes, and dynamic information underlying skillful actions about a domain required to perform skilled domain related tasks successfully (Page & Uncles, 2004). The definition of common knowledge is a non-technical one and should be distinguished from the technical definition of the term in economics and game theory, where it denotes a piece of information that is available to agents A and B, and both A and B know that the other has the information, and both A and B know that the other knows that the other has the information, and so on ad infinitum (Lewis, 1969).

Given the widespread usage of the web, consumers are likely to possess a high degree of common knowledge. They are also likely to have higher familiarity with common knowledge, enabling realistic assessments of confidence. Specialized knowledge questions are more technical in nature and consumers are likely to be less familiar with these. Also, given the widespread use of computers and the web, consumers could have been exposed to some of these technical aspects during formal training or in their own usage of the web. We posit that the modicum of familiarity arising from such exposure leads to a perception that they know these topics, even when they do not. Research on the feeling of knowing has recorded how such an intuition often produces dissociation between accuracy and confidence (Koriat, 1998; Nelson et al., 1986). Such dissociations will lead to a miscalibration of knowledge. It is argued that this phenomenon is more likely in the case of specialized knowledge than in the case of common knowledge, leading to higher calibration in the latter, compared to the former. Hence,

H6. Calibration will be higher for common knowledge, compared to specialized knowledge.

3. Method

3.1. Sample

We used two samples to test the hypotheses. Sample 1 comprised 151 students enrolled in undergraduate business courses in a large university in the southeastern United States, who participated in the study for course credit. Sample 2 comprised 153 adults in a British city. The adults included (a) community members (40%) who were contacted in libraries, parks, residential apartments, and retail stores, (b) nonacademic staff members (25%) and faculty members (5%) in a leading British university and (c) postgraduate and undergraduate students in the same university (30%). All respondents were contacted individually and were paid 2 pounds for participation. The respondents filled out a structured questionnaire that measured consumer knowledge, involvement with the web, experience, and usage. In order to control for order effects, roughly half of the sample answered questionnaires which had procedural knowledge questions first, followed by declarative knowledge questions, while the other half answered
questionnaires which had declarative knowledge questions first, followed by procedural knowledge questions. Common and specialized knowledge questions were distributed within these. We conducted t-tests to assess differences between the samples. No difference was observed in mean calibration ($p=.85$) while a difference was observed in confidence ($p=.00$). The latter could be attributed to the difference in age, in that the first sample comprised only undergraduate students, whereas the second comprised a distribution of adults of varying age groups. A simple regression model in the second sample showed a significant effect (beta=.21; $p=.01$) of age on confidence. Moreover, regression models between confidence and the independent variables showed comparable beta parameter estimates and significances across the two samples. Consequently, we decided to combine the two samples, with sample included as a covariate. Results are reported for the combined sample as well as the separate samples. Given the similarity between American and British cultures, and the technical nature of the questions, cross-cultural issues are discounted.

Information on gender was missing from 28 questionnaires in the first sample and hence the sample size for multivariate tests was correspondingly lower. In addition, in some cases web involvement could not be computed due to missing data, which also resulted in a few dropouts. When calibration was computed for subscales, there were a few instances where correlation could not be computed because one variable was constant, which resulted in slight reductions in sample sizes for the specific analyses using the subscales.

3.2. Measures

3.2.1. Consumer knowledge calibration

As mentioned earlier, consumer knowledge of the web was measured using a 38-item scale developed by Page and Uncles (2004). Based on qualitative and quantitative analyses, these authors developed four scales that measure common declarative knowledge, specialized declarative knowledge, common procedural knowledge, and specialized procedural knowledge. The number of items for each of the aforementioned scales was 10, 11, 6, and 11. Most of the items employed a true/false format to elicit responses, while a few had multiple choice responses. A combined scale, comprising all 38 items (21 declarative knowledge items and 17 procedural knowledge items) was used to measure consumer knowledge of the web in this study. The few multiple choice items (3 items) were converted to dichotomous (true/false) format to ensure correspondence. Respondents indicated their choice of answer for each item, and then indicated their confidence that the answer is correct, on a scale ranging from 50% to 100%. We chose a restricted scale ranging from 50% to 100%, as against a scale ranging from 0% to 100%, because of the dichotomous response format (true or false) which provides a 50% chance of success even in a random choice.

Knowledge calibration was measured through point biserial correlation between accuracy and confidence (e.g., Lin, Moore, & Zabrucky, 2001). Calibration was measured at the within subjects level. Correlation between accuracy (either 0 or 1) and confidence rating for each item provided the measure of calibration at the respondent level. In keeping with the definition of calibration as ranging from 0 (nil calibration) to 1 (perfect calibration), we restricted the lower bound of the correlation measure to zero. That is, negative correlations observed in a few cases (which were not statistically significant) were restricted to zero.

3.2.2. Involvement

Involvement was measured using the ten-item scale proposed by Zaichkowsky (1994). Respondents were asked to indicate their perceptions of the web on a 7-point bipolar scale anchored by adjectives such as important–unimportant, fascinating–mundane, relevant–irrelevant, etc. Involvement was measured as the mean of the ten items, after recoding the negatively framed items.

3.2.3. Experience

Experience was measured by a single item that asked respondents to specify the number of years they have been using World Wide Web.

3.2.4. Usage

Usage was measured by a single item that asked respondents to specify the number of hours per week they spend browsing or using the web. It should be noted here that online researchers have employed single-item measures to measure experience/usage (e.g., Rodgers, Negash, & Suk, 2005). Also, the use of self-reported usage measures have been validated by research that suggests that self-reported usage measures correlate well with actual usage measures (Venkatesh & Davis, 2000).

4. Results

4.1. Scale reliability

The ten-item involvement scale had acceptable reliability (alpha=.875). The reliability of the knowledge of the web scale, computed using KR 20, fell below the threshold of .7 (overall .52; specialized=.83; common=.13; procedural=.35; declarative=.39). This is not deemed to be a cause for concern due to the following reasons: (a) the construct of knowledge of the web is necessarily multidimensional and a scale that attempts to tap into the various dimensions will have to compromise on reliability indices and (b) the focal construct is calibration of knowledge of the web and not knowledge of the web per se. As long as the scale has face and content validity, it serves to be an acceptable instrument to measure calibration. It should also be noted that previous research on consumer knowledge has used a set of items chosen judgmentally by researchers, for which reliabilities have not been reported (e.g., Capraro, Broniarczyk, & Srivastava, 2003).

4.2. Descriptive statistics

Table 1 shows the descriptive statistics. Calibration ranged from 0 to .72, with a mean of .34. Mean level of involvement
with the web was 6.15, indicating a high level of involvement with the web. Experience with the web ranged from 1 to 15 years, with a mean of 8.3 years. Usage ranged from 1 to 40 h per week, with a mean of 11.26 h. Accuracy ranged from 55% to 97%, with an average of 80%. Hence, the sample exhibited high levels of accuracy. This finding leads to the discounting of instances of calibration arising from low accuracy and low confidence.

4.3. Tests of hypotheses

We employed a general linear model procedure to test the hypotheses, with calibration and confidence as the dependent variables and web-involvement, experience, usage, and gender as the independents. Sample was a covariate, when the overall sample was tested. Comparison of means was employed to test hypotheses 5 and 6.

The first hypothesis stated that higher involvement will lead to higher calibration of knowledge. This hypothesis was supported by the overall sample (Wilks’ lambda = .96; \( F = 5.11; p = .01 \); \( b = .03; p = .03 \)) and the American sample (Wilks’ lambda = .93; \( F = 4.13; p = .002; b = .05; p = .03 \)), but not by the British sample (Wilks’ lambda = .97; \( F = 1.94; p = .15; b = .01; p = .33 \)). Table 2a shows the multivariate test results (Wilks’ lambda, \( F \), and \( p \) values), while Table 2b shows the parameter estimates and their \( p \) values.

The second hypothesis stated a positive relationship between experience and confidence. Though no hypothesis was offered, we examined the relationship between experience and calibration. The overall sample supported H2 (Wilks’ lambda = .96; \( F = 6.16; p = .00; b = .69; p = .00 \)), while no relationship was found between experience and calibration (Wilks’ lambda = .96; \( F = 6.16; p = .00; b = .00; p = .69 \)). Similarly, American and British samples also supported H2 (American — Wilks’ lambda = .94; \( F = 3.98; p = .02; b = .78; p = .02 \); British — Wilks’ lambda = .96; \( F = 3.25; p = .04; b = .62; p = .02 \) and found no relationship between experience and calibration (American — Wilks’ lambda = .94; \( F = 3.98; p = .02; b = .01; p = .16 \); British — Wilks’ lambda = .96; \( F = 3.25; p = .04; b = .00; p = .65 \)).

The third hypothesis stated that usage will have a positive effect on calibration. The overall sample did not support this hypothesis (Wilks’ lambda = .92; \( F = 10.92; p = .01; p = .17 \)). The American sample also did not support the hypothesis (Wilks’ lambda = .98; \( F = 1.36; p = .26; b = .00; p = .95 \)). The British sample showed an effect, albeit in the negative direction (Wilks’ lambda = .84; \( F = 12.73; p = .00; b = .004; p = .04 \)).

Hypothesis 4 stated that males will have higher confidence in their knowledge of the web, compared to females. We also examined the relationship between gender and calibration as an empirical issue. The overall sample supported H4 (Wilks’ lambda = .84; \( F = 24.71; p = .00; b = -6.65; p = .00 \)) and found no relationship between gender and calibration (Wilks’ lambda = .84; \( F = 24.71; p = .00; b = .02; p = .29 \)). Similar effects were obtained in American (H4 — Wilks’ lambda = .90; \( F = 6.65; p = .00; b = -4.73; p = .00 \); gender—calibration — Wilks’ lambda = .90; \( F = 6.65; p = .00; b = .00; p = .96 \)) and British samples (H4 — Wilks’ lambda = .79; \( F = 18.42; p = .00; b = -8.12; p = .00 \); gender—calibration — Wilks’ lambda = .79; \( F = 18.42; p = .00; b = .03; p = .24 \)). It is to be noted that gender was entered as a dummy variable in the general linear model with values of 1 for males and 2 for females.

The fifth hypothesis stated that calibration will be higher for procedural knowledge compared to declarative knowledge. We computed calibration scores for declarative and procedural subscales for each respondent. This was done by measuring the correlations between accuracy and confidence in the subscales

Table 1
Descriptive statistics — means and standard deviations

<table>
<thead>
<tr>
<th>Effect</th>
<th>Overall N</th>
<th>American N</th>
<th>British N</th>
<th>Overall mean (S.D.)</th>
<th>American mean (S.D.)</th>
<th>British mean (S.D.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibration</td>
<td>304</td>
<td>151</td>
<td>153</td>
<td>.34 (.16)</td>
<td>.34 (.16)</td>
<td>.33 (.15)</td>
</tr>
<tr>
<td>Web involvement</td>
<td>295</td>
<td>149</td>
<td>146</td>
<td>6.15 (.84)</td>
<td>6.32 (.71)</td>
<td>5.97 (.93)</td>
</tr>
<tr>
<td>Experience</td>
<td>302</td>
<td>150</td>
<td>152</td>
<td>8.31 (2.52)</td>
<td>8.66 (2.15)</td>
<td>7.97 (2.80)</td>
</tr>
<tr>
<td>Usage</td>
<td>301</td>
<td>150</td>
<td>151</td>
<td>11.26 (7.85)</td>
<td>12.45 (8.04)</td>
<td>10.07 (7.51)</td>
</tr>
<tr>
<td>Accuracy</td>
<td>304</td>
<td>151</td>
<td>153</td>
<td>80% (8%)</td>
<td>81% (7%)</td>
<td>79% (9%)</td>
</tr>
<tr>
<td>Confidence</td>
<td>304</td>
<td>151</td>
<td>153</td>
<td>81% (9%)</td>
<td>83% (8%)</td>
<td>79% (10%)</td>
</tr>
</tbody>
</table>

Table 2a
Multivariate tests for H1, H2, H3, and H4

<table>
<thead>
<tr>
<th>Effect</th>
<th>Overall sample</th>
<th>American sample</th>
<th>British sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value ( F )</td>
<td>Error ( df )</td>
<td>Sig.</td>
</tr>
<tr>
<td>Intercept</td>
<td>Wilks’ Lambda</td>
<td>.43</td>
<td>172.68</td>
</tr>
<tr>
<td>Experience</td>
<td>Wilks’ Lambda</td>
<td>.96</td>
<td>6.16</td>
</tr>
<tr>
<td>Gender</td>
<td>Wilks’ Lambda</td>
<td>.84</td>
<td>24.71</td>
</tr>
<tr>
<td>Web involvement</td>
<td>Wilks’ Lambda</td>
<td>.96</td>
<td>5.11</td>
</tr>
<tr>
<td>Sample</td>
<td>Wilks’ Lambda</td>
<td>.96</td>
<td>5.44</td>
</tr>
</tbody>
</table>
for each respondent. The hypothesis was tested by comparing the mean calibration of procedural knowledge (.38) among all respondents with the mean calibration of declarative knowledge (.31). A paired sample t-test indicates that the difference is significant (p=.00). Table 3 shows the test results. These show that respondents are better calibrated in their procedural knowledge compared to their declarative knowledge. The American sample also supported the hypothesis (p=.00), while the British sample fell marginally short of significance (p=.055).

The sixth hypothesis stated that calibration will be higher for common knowledge compared to specialized knowledge. Similar to the procedure mentioned above, paired sample t-tests were used to test the hypothesis. The overall sample supported the hypothesis (mean calibration common=.39; specialized=.28; p=.00). The American as well as the British samples provided similar effects (American p=.00; British p=.01) (Table 4).

To throw further light on the gender differences, we also examined the differences in accuracy. It was found that the accuracy levels are higher for males compared to females (p=.00). Hence, although no difference was observed in calibration, the results lend support to previous findings regarding gender differences in accuracy and confidence.

We also examined whether the hypothesized relationships between individual-specific factors and calibration and confidence differ between the various types of knowledge. Four models, where calibration and confidence were computed for the subscales (procedural knowledge, declarative knowledge, common knowledge, and specialized knowledge), were examined for the entire sample (including sample as a covariate).

Results were similar to those for the entire scale for hypotheses 2–4, as well as the empirical relationships between (a) experience and calibration and (b) gender and calibration. Regarding hypothesis 1, it was found that web involvement led to high calibration in the cases of declarative (b=.03; p=.04) and specialized knowledge (b=.04; p=.01), but not in the cases of procedural (b=.02; p=.33) and common knowledge (b=.00; p=.89).

Since it could be argued that the constructs of experience (number of years the person has been using the web) and usage (number of hours spent browsing the web) are inherently related, we ran the model with (a) calibration and (b) confidence as dependents, and (a) web involvement, (b) gender and (c) the product of experience and usage as independents, with sample as the covariate. The results remained the same. Only web involvement was related to calibration (p=.04).

5. Discussion

We organize the discussion section into two parts. In the first part, we continue the discussion of the model, examining some potential antecedents and, more importantly, the consequences of calibration of consumer knowledge of the web. In the second part, we examine the implications of the study and outline future research directions.

5.1. Extension of the model-antecedent factors

Several other antecedents of consumer knowledge of the web can be identified. Two potential antecedents are distorted memory and counterfactual thinking. Notions of “availability,”

### Table 2b

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Parameter</th>
<th>Overall sample</th>
<th></th>
<th>American sample</th>
<th></th>
<th>British sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>Standard error</td>
<td>t</td>
<td>Sig.</td>
<td>B</td>
</tr>
<tr>
<td>Calibration</td>
<td>Intercept</td>
<td>.16</td>
<td>.08</td>
<td>1.94</td>
<td>.05</td>
<td>-.04</td>
</tr>
<tr>
<td></td>
<td>Experience</td>
<td>.00</td>
<td>.00</td>
<td>.40</td>
<td>.69</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>Usage</td>
<td>-.00</td>
<td>.00</td>
<td>-.13</td>
<td>.17</td>
<td>-.00</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>.02</td>
<td>.02</td>
<td>1.05</td>
<td>.29</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>Web involvement</td>
<td>.03</td>
<td>.01</td>
<td>2.15</td>
<td>.03</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td>Sample</td>
<td>-.00</td>
<td>.02</td>
<td>-.14</td>
<td>.91</td>
<td>-.00</td>
</tr>
<tr>
<td>Confidence</td>
<td>Intercept</td>
<td>75.02</td>
<td>4.04</td>
<td>18.57</td>
<td>.00</td>
<td>72.45</td>
</tr>
<tr>
<td></td>
<td>Experience</td>
<td>.02</td>
<td>.02</td>
<td>-.07</td>
<td>.95</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>-.66</td>
<td>.96</td>
<td>-.69</td>
<td>.00</td>
<td>-.78</td>
</tr>
<tr>
<td></td>
<td>Sample</td>
<td>-.32</td>
<td>.98</td>
<td>-.33</td>
<td>.00</td>
<td>-.36</td>
</tr>
</tbody>
</table>

O — Overall; A — American; B — British.
“dominance of the given,” “salience effect” and so forth formalize the idea that memory and judgment have a simple direct relationship. Empirical studies support the expectation of a direct relationship between memory and judgment. Beyth-Marom and Fischhoff (1977) provided strong evidence on the strength of memory–judgment relationship, and Gabrielpick and Fazio (1984) demonstrated that ease of retrieval exerts a causal (not just correlational) influence on frequency estimates. Based on the evidence, we reason that distorted and incomplete memory diminishes calibration of knowledge of the web. Conversely, undistorted and complete memory promotes calibration. Another potential antecedent factor is counterfactual thinking. It has been suggested that individuals who are required to consider contradictory evidence have lower over-confidence than individuals who are not required to do so (Hoch, 1985). Experimental studies have demonstrated that individuals automatically generate supporting reasons as a byproduct of the question–answering process, but must be explicitly instructed to consider the counterfactual (Koriatt, Lichenstein, & Fischhoff, 1980). Reviewing the relevant psychology literature, Faust (1986) identifies a series of rules or necessary conditions that help improve calibration. Among the rules suggested are to (1) present evidence that contradicts, disconfirms, and refutes one’s position and (2) generate competing alternative hypotheses. It follows that those who are given to considering opposite and alternative perspectives regarding their knowledge of the web are likely to be more calibrated in their knowledge than those who do not consider such counterfactual perspectives (Pillai & Goldsmith, 2006). Hence, counterfactual thinking could lower overconfidence and enhance calibration of knowledge of the web. Several other antecedent variables could be identified. The interested reader is referred to the review article by Alba and Hutchinson (2000).

5.2. Consequence of calibration of knowledge of the web

We discuss two potential consequences of calibration of web knowledge — frustration with web and flow experience.

5.2.1. Frustration with web

A lack of calibration could contribute to the frustrations ordinary people experience when dealing with the paradoxes of web technology (Mick & Fournier, 1998). Frustration is the natural outcome when the technology challenge is too great (Hoffman & Novak, 1996), or to restate slightly, when a firm’s web interface is not sufficiently easy to use. In addition to the key role that ease of use has with the TAM described previously, numerous other studies (e.g., Collier & Bienstock, 2006; Yoo & Donthu, 2001; Yang & Jun, 2002) have uncovered a critical role for ease of use in perceptions of online service quality. Overconfidence could make the consumer more likely to conclude that the site is hard to use and to attribute e-service failure to the site, not to him or herself, with potentially negative consequences for the retention of that customer.

Several firms plan to move customer service to the web in order to lower costs and provide faster service. Frustration arising from poor calibration could lead to the disconfirmation of expected service quality. This in turn will lead to lower levels of exclusive loyalty and repeat-purchase rates. The frustration could be attributed to self (“I am not as good as I thought”) or to the firm (“the website is not good”). Neither of these augurs well for the firm. Thus, calibration of web knowledge could be a mediator between firm’s e-service offering and customer satisfaction/loyalty.

While calibration of knowledge leads to optimal information search, overconfident people (cell 1) might act presumptuously and commit errors in their web use. Underconfident people (cell 4) might spend considerable time searching for the relevant information, even while they are in the know. Both of these are likely to result in frustration with web use. Calibrated people who possess the right amount of knowledge (cell 2) spend little time searching for such information, while calibrated people who don’t possess the right kind of knowledge spend appropriate and optimal amount of time searching for the relevant information. Some research has examined specific antecedents of user frustration, including lack of navigation standards (Neerincx, Lindenberg, & Pemberton, 2001), disorientation in hyperspace and associated cognitive overload (Nielsen, 2000; Patel, Drury, & Shalin, 1998), and speed, accuracy, and preference of navigational systems (Kiger, 1984). Apart from the direct effect of miscalibration on frustration as argued above, miscalibrated individuals are likely to feel disoriented, which also contributes to frustration.

5.2.2. Flow

Flow is the process of optimal experience (Csikszentmihalyi, 1977). Flow experience in a computer mediated environment has been defined as the state occurring during network
navigation which is (1) characterized by a seamless sequence of response facilitated by machine interactivity, (2) intrinsically enjoyable, (3) accompanied by a loss of self-consciousness, and (4) self-reinforcing (Hoffman & Novak, 1996). In the flow experience, consumers are so acutely involved in the act of network navigation in the hypermedia environment that "nothing else seems to matter" (Csikszentmihalyi, 1990, p.4).

Flow is assumed to have a number of positive outcomes for the firm (Hoffman & Novak, 1996).

Calibration of knowledge of the web will facilitate the flow experience. As noted earlier, flow occurs when perceived skills match perceived challenges. A realistic assessment of the skill is required to achieve flow. Since knowledge of the web is a skill in the context of navigation, better calibration leads to better flow. Overconfident individuals will face bottlenecks in their navigation and wouldn’t know how to remedy the situation since they are unaware of their ignorance. They will also undertake tasks that are too challenging, which results in a reduction in flow experience. Underconfident individuals will dither, which inhibits their navigation. They are likely to engage in navigation that is too simple to be enjoyable. Calibrated individuals are likely to experience flow in the strict sense of the term.

5.3. Implications of the findings

The results of the study point to the role of involvement with the web in promoting calibration of knowledge of the web. Higher calibration implies greater realism of judgment in the level of one’s knowledge. Greater realism enables one to judge correctly what one is capable of doing. It also facilitates optimal information search, which in the context of knowledge of the web could include attending training programs, reading manuals, or seeking information from colleagues. And, as argued above, calibration lowers frustration with the web in promoting calibration of knowledge of the web.

It is also shown that calibration is higher for procedural knowledge, compared to declarative knowledge. This finding can add to the research on designing computer and software manuals. In particular it adds a new dimension to the debate on the optimum contents of software manuals. Some researchers (Carroll, 1990; Van der Maj & Carroll, 1995) have asserted that manuals should be reduced to contain only procedural, task-supporting information. The finding regarding the differences in calibration of procedural and declarative knowledge of the web can inform this debate. If people are more likely to be miscalibrated in their declarative knowledge, there is a case to include such knowledge in manuals. Similar reasoning would lead to the argument that, apart from common knowledge, manuals should focus on specialized knowledge as well. These findings will also contribute to the design of contents for training modules on the World Wide Web.

The study found no difference in calibration between the genders. However, the previously reported finding that males are more confident than females was supported. Though we did not offer formal hypotheses, data on age and education was collected in the British sample. A simple regression model indicated that neither age ($p=.17$) nor education ($p=.65$) had any effect on calibration. Education, however, was found to be positively correlated with confidence ($r=.22$).

Generalizing from the above findings, it can be stated that calibration of knowledge is more a function of type of knowledge and less a function of demographic variables. This is an important finding to the program of research on consumer knowledge calibration.

This study contributes to the development of a program of research on knowledge calibration in consumer sciences. A large body of work on knowledge and confidence calibration, its antecedents and consequences informs the dialogue in several areas of enquiry such as psychology, education, etc. Surprisingly, little research on knowledge calibration exists in consumer research. It is hoped that this study acts as a catalyst for enquiry in this important area.

Future research needs to extend the proposed model by examining other antecedents and consequents. As mentioned earlier, the mediating role of calibration between e-service offering and satisfaction/loyalty is of considerable importance and needs to be examined. Research also needs to investigate the roles of other antecedent factors in influencing calibration of knowledge of the web. Further research can also distinguish between the different types of high and low calibration (cells 1 to 4) and examine the differential effects of antecedent factors and consequences arising from these different instances. In particular, the two types of calibration will have different effects on consequences such as information search. As discussed before, high accuracy and high confidence were observed in this study, rendering such distinctions redundant.

At a general level, the model can be applied to the domain of specific knowledge of consumers and their use of the web.
Considerable research has examined online information search. Researchers have suggested several antecedents to online information search, such as income, experience with the web, and education (Klein & Ford, 2003). Models of online search that combine the benefit–cost perspective and the ability–motivation paradigm have been proposed (Kulviwat, Guo, & Engchana, 2004). Calibration of knowledge of the domain (product category) could be an antecedent to online search (those in cells 1 and 2 perform lower search compared to those in cells 3 and 4). More importantly, calibration could be a determinant of the effectiveness of search, whereby calibrated individuals will perform more effective search compared to misclassified individuals.

Acknowledgments

The authors thank Professor Hubert Gatignon and three anonymous reviewers for their valuable comments throughout the review process.

References


