Communications of the Association for Information Systems

Volume 35 Article 11

11-2014

Making and Evaluating Participant Choice in Experimental Research on Information Technology: A Framework and Assessment

Nancy K. Lankton
Marshall University, lankton@marshall.edu

Joan Luft
Department of Accounting and Information Systems, Michigan State University

Communications of the Association for Information Systems

CAIS

Making and Evaluating Participant Choice in Experimental Research on Information Technology: A Framework and Assessment

Nancy Lankton

Division of Accountancy and Legal Environment, Marshall University

lankton@marshall.edu

Joan Luft

Department of Accounting and Information Systems, Michigan State University

Abstract:

Evaluations of participant samples for experiments in information systems research often appear to be informal and intuitive. Appropriate participant choice becomes a more salient issue as the population of information technology professionals and users grows increasingly diverse, and the distribution of relevant characteristics in participant samples such as age, gender, nationality, and experience can often be unrepresentative of the characteristics' distribution in target populations. In this paper, we present a framework based on widely accepted standards for evaluating participant choice and providing rationale that the choice is appropriate. Using a step-by-step approach, we compare current practice in experimental studies from top information systems journals to this framework. Based on this comparison, we recommend how to improve the treatment of participant choice when evaluating the validity of study inferences and how to discuss the tradeoffs involved in choosing participant samples.

Keywords: Experimental Research, Sample Choice, Individual Characteristics, Experimental Participant.

Volume 35, Article 11, pp. 199-224, November 2014

The manuscript was received 15/02/2013 and was with the authors 8 months for 3 revisions.

I. INTRODUCTION

As the population of information system (IS) designers and users grows increasingly diverse—from leading-edge scientists to children, third-world villagers, and executives—questions about the validity of results obtained from using limited samples become increasingly salient. This is particularly the case for experimental research in IS, which often uses convenience samples of employees from a particular organization or students from a particular university, who are arguably often not representative of a study's target population. That is, individuals in the sample differ from the study's target population with respect to characteristics such as age, gender, nationality, or experience that (a) are not the variables of primary interest in the study and (b) could pose threats to the validity of the study's inferences about the variables of primary interest. Hereafter, we refer to these characteristics as "sample characteristics".

A recent paper (Compeau, Marcolin, Kelley, & Higgins, 2012) expresses strong concerns about how researchers are treating sample choices in IS research. Examining studies that use individual-level human subject data in two IS journals, Compeau et al. (2012) found that only a minority of these studies provided *any* justification for their sample choice or discussion of the resulting limitations. Moreover, in the authors' view, when justifications were provided, they were often inadequate. As such, they recommend that future research treat sample-choice issues much more carefully by identifying specific similarities between characteristics of the sample and "important population characteristics" and identifying specific limitations to generalizability that result from any dissimilarities between sample group and target population.

IS researchers are thus faced with the questions: how do we form and justify beliefs about which characteristics are important? If there are dissimilarities between sample and population, which of them limit the inferences that can be drawn from the research, and *how* do they limit these inferences? In this paper, we provide a framework for answering these questions by integrating best-practice recommendations from multiple sources. This framework includes a step-by-step approach that considers the statistical properties of the sample characteristics, threats to validity resulting from these properties, and judgments about which sample and empirical-modeling choices will result in the best tradeoffs—that is, which choices will reduce some important threats to validity with no increase, or an acceptably small increase, in other threats.

We find that current practice in IS experimental research rarely discusses experiment participant choices in terms of these tradeoffs. Researchers often omit the statistical tests that could help determine the appropriateness of a study's participant choices and the magnitude of the remaining threats to validity, even when they have collected the sample characteristic measures that would enable them to perform these tests (see Section 5 below). Discussions about how participant choice influences the validity of a study's inferences are perhaps too often limited to formulaic remarks such as "Results should be replicated with other groups" or "The use of student participants is justified by their use in other studies"[1]. Additionally, researchers often do not identify specific threats to validity that they were not able to eliminate in their study. Nor do they explain why their study's design represents a good tradeoff between reducing some threats and failing to reduce others.

The paper proceeds as follows. In Section 2, we describe our data collection methods. In Section 3, we show, consistent with Compeau et al. (2012), that participant samples in IS research are not prima facie representative of target populations. In Sections 4 to 8, because samples that are imperfectly representative do not always pose unacceptable threats to a study's validity, we provide a five-step approach for addressing validity issues related to imperfectly representative samples: (1) identify and measure potentially influential sample characteristics, (2) examine the variability of these sample characteristics, (3) test or otherwise assess statistical properties of the sample characteristics to determine what specific threats they could pose to the validity of the study's inferences, (4)

[1] With this paper, we do not intend to contribute to the extensive literature that debates whether and how students as participants are different from various subpopulations of nonstudents (e.g., Rosenthal & Rosnow, 2009; Sears, 1986; Gordon, Slade, & Schmitt, 1986; Henrich, Heine, & Norenzayan, 2010). Rather, we provide criteria for determining when and how these differences—*or any other differences between participant samples and target populations*—matter to IS research. Differences undoubtedly exist—for example, students have less-formulated senses of self and stronger tendencies to comply with authority than the population at large (Sears, 1986). These differences clearly matter to social psychology and behavioral science research that measures such characteristics in the population at large and in subpopulations. But, often, these differences do not matter to the questions of interest to IS research. For example, individuals with more- and less-formulated senses of self may be similarly influenced by graphic interfaces.

report these specific threats, and (5) explain the trade-offs among them. For each step, we also examine current practice in IS experimental research and recommend improvements.

II. DATA COLLECTION

To examine current practice in IS research, we reviewed papers published from 2000-2012 in four premier IS journals, *MIS Quarterly* (*MISQ*), *Information Systems Research (ISR)*, *Journal of Management Information Systems* (*JMIS*), and *Journal of the Association for Information Systems* (*JAIS*). These journals are in the Senior Scholars' Basket of Six, and were recently ranked as the top four information systems journals (Lowry et al., 2013), which makes them primary sources of "best practice" observations. Our final sample includes 184 experimental studies, including 53 from *MISQ*, 46 from *ISR*, 57 from *JMIS*, and 28 from *JAIS* (Appendix and Table 1). These studies account for 11 percent of the total research published in these journals over the 13-year period (Table 1). To collect the studies, we first defined "experimental studies" as those involving human subjects and in which at least one independent variable (IV) was manipulated and randomly assigned to participants to test one or more hypotheses. We excluded mathematical-simulation experiments, surveys, and usability studies that primarily validate the functionality of new software rather than test theory-based hypotheses. Next, the first author examined all papers published in the four journals during this period by reading abstracts, and skimming/reading the paper if needed. Then, the second author and a graduate assistant also examined papers in these journals using this same method. The two authors discussed those papers that only one identified as experimental, or that were borderline cases[2], to determine whether to include them in the study.

Table 1: Experimental Research Studies Published 2000-2012											
	MISQ papers		ISR papers		JMIS papers		JAIS papers		Total		
Year	Exp.	Total	Exp.	Total	Exp.	Total	Exp.	Total	Exp.	Total	%
2000	3	23	5	24	3	34	2	12	13	93	6%
2001	2	16	3	23	4	35	3	8	12	82	5%
2002	2	17	2	23	7	36	1	7	12	83	5%
2003	2	22	3	18	2	34	2	15	8	89	5%
2004	2	24	2	20	5	35	2	18	10	97	6%
2005	2	28	2	21	5	42	0	13	9	104	6%
2006	8	42	7	23	3	41	2	33	20	139	9%
2007	3	30	1	23	4	40	0	33	8	126	8%
2008	3	34	5	25	5	41	5	31	18	131	8%
2009	6	43	5	29	6	37	3	32	20	141	9%
2010	7	37	3	53	2	40	2	32	14	162	10%
2011	9	50	5	47	5	39	5	30	24	166	10%
2012	4	60	3	74	6	40	1	35	14	209	13%
Total	53	426	46	403	57	494	28	299	184	1,622	11%

For each study, we identified information about the sample, the target population, and the experimental task. We also collected information about any sample characteristics measured and how these sample characteristics were used in the study's data analysis. The first author and a graduate assistant did this coding. A random check of 20 percent of the coding of sample-characteristic information by the second author yielded no systematic differences in coding outcomes and no differences in conclusions about reporting practice in these experiments. Tables 2 through 6, which form the basis for the following analyses, summarize this information.

III. SAMPLES AND TARGET PARTITICIPANTS IN IS EXPERIMENTS

To provide evidence on the apparent lack of representativeness in IS experiment-participant samples, we replicated and extended Compeau et al.'s (2012) findings of questionable matches between sample and target populations. We used a broader and newer sample of research[3] than theirs, and we compared participant samples not only with explicitly stated target populations (which, as Compeau et al. (2012) observe, are relatively infrequent) but also with the target populations that were implicit in a study's research question or in the experimental task or technology employed. For example, if a paper's research question was how website design influenced consumers' willingness to transact with an online seller, then, in the absence of further qualifications (e.g., targeting only consumers in certain income or age ranges), we assumed the target population was online consumers. Similarly, if the task was to

[2] For example, we included some studies in which randomization of participants to treatments was limited (e.g., different treatments could not be assigned to students in the same class section).
[3] We examined four journals rather than two. We also began our sample in 2000 rather than 1990 to help exclude possibly outdated practices.

decide whether an organization should invest in particular software, then, in the absence of further qualifications, we assumed the target population was individuals who make software-investment decisions for organizations.

Table 2 overviews the sample and target population choices in the 184 studies we analyzed. The first column describes the participant sample in each study. The second column describes the target population. The third column describes the tasks and technologies used in the experiment (as a check on our judgment of implicit target populations). The fourth column presents the number of studies that include the sample/target population/tasks combination identified in each row.

Data in Table 2 shows that target populations—for example, “software application users needing training” or “organizational members performing group tasks”—were often highly diverse in terms of characteristics that might influence their IS-related behavior, such as age, education, nationality, and personality traits. Often, however, samples were students drawn from a single course or degree program at a single university or employees drawn from a single organization or organizational unit. These samples can be quite homogeneous with respect to some potentially influential characteristics. They can also be diverse with respect to some of these characteristics, but, in many cases, they are likely to lack *representative* diversity. For example, ages or nationalities in the sample might be quite diverse in both the sample and target population but also quite differently distributed in the two groups.

Thus, in these IS experiments, samples are often not prima facie representative of the target populations with respect to potentially influential sample characteristics; that is, characteristics that might influence the dependent and/or independent variables (DVs and IVs) in the experiment. Although representativeness is a desirable property of samples, we do not simply argue for more representative samples here. Rather, we provide an approach for evaluating threats to validity that arise from imperfect representativeness when full representativeness is not practically attainable. Sometimes it is not attainable—or at least cannot be verified—because the distributions of potentially influential characteristics in the target population are not known with sufficient exactitude. Sometimes, when multiple characteristics are potentially relevant, full representativeness is not attainable because an available sample that is a good match to the population with respect to some characteristics is a poor match with respect to others, and no available sample is an equally good match on all characteristics. Sometimes obtaining a fully representative sample is costly, and researchers may question whether the resulting increase in validity is significant enough to justify the cost. As we argue below, imperfect representativeness—that is, lack of matching on potentially influential characteristic distributions between sample and population—does not always create a threat to valid inference, even when the sample characteristics have significant effects on the variables in the study.

When mismatches between sample and population characteristics do create threats to validity, researchers often face tradeoffs in research design because a choice that reduces one threat can increase another. Researchers then need to make and explain judgments about their sample choice and its consequences for the validity of their research in terms of these tradeoffs. As Compeau et al. (2012) point out, the explanations of sample choice that appear in the literature are often rather perfunctory: they are based on prior practice or simple assertion rather than any more rigorous, theory-based approach. In the following sections, based on widely accepted standards for assessing validity in empirical research (e.g., Shadish, Cook, & Campbell, 2002), we propose a more systematic approach to explaining sample choice and its consequences,. We also compare existing practice to these standards and, in the process, often document a substantial gap between current and best-practice treatment of sample-characteristic issues in the IS literature.

IV. STEP ONE: IDENTIFY AND MEASURE POTENTIALLY INFLUENTIAL SAMPLE CHARACTERISTICS

Figure 1 illustrates a five-step approach to identifying and analyzing threats to validity arising from potentially influential sample characteristics that may not be representative of a study’s target population. Dotted-line boxes on the right-hand side of the figure present summary statistics on particularly visible inconsistencies between current and best practice. We present this as an ex post approach in which researchers have already (at least tentatively) selected a sample that they believe is appropriate and now must verify that belief to satisfy themselves and others of their research’s validity. However, the steps in Figure 1 can also be used as part of the ex ante judgment process in selecting a sample while researchers consider the sample’s likely consequences on validity.

The first step in Figure 1 is to identify and measure potentially influential sample characteristics and to explain why these characteristics are potentially influential. Researchers usually cannot identify with certainty a priori all the characteristics of the individuals in the target population and the sample (e.g., specialized training or willingness to take risks) that might influence the variables of interest in their studies. But IS and other social-science literature can often provide a basis for identifying a set of potentially influential characteristics. These characteristics should then be measured in the sample.

Table 2: Participants, Tasks, and Target Populations for IS Experiments			
Parti-cipants	**Target population**	**Tasks/technology identified**	**#**
Undergrad students	Stakeholders in the system development process	Reading and developing data models, developing queries, buying software	16
	Group decision-makers using DSS tools	Generating alternatives, choosing among alternatives, completing projects, identifying deception	24
	Online consumers	Evaluating/buying products/websites, using online agents	20
	Software application users	Undergoing training, completing skill tests	5
		Subtotal studies using undergraduate students	65
Undergrad and graduate students	Group or individual decision-makers using DSS	Choosing alternatives, developing a business plan, negotiating prices	7
	Stakeholders in the system development process	Reading conceptual models, programming, making project continuance decisions, reusing software	9
	Online consumers	Browsing websites/products, using mobile devices, participating in virtual worlds and online communities	15
	Organizational members logging on to systems, using information from various formats/displays	Making personality judgments, interpreting graphical data, assessing security issues	5
		Subtotal studies using undergraduate and graduate students	36
Graduate students	Managers using DSS	Making decisions, choosing alternatives, negotiating	4
	Stakeholders in the system development process	Making technology investment decisions	1
	Managers using a supply chain management system	Procuring goods	1
	Electronic market participants	Examining seller information, bidding	1
		Subtotal studies using graduate students	7
Unspe-cified Students	Stakeholders in the system development process	Reading conceptual models, performing systems analysis, making software project decisions, querying	7
	Online consumers	Providing information, browsing websites, evaluating products, examining seller profiles, using virtual reality	18
	Software application users	Training, completing skill tests	1
	Organizational members and professionals making decisions, performing group work	Making real-time dynamic decisions and solving problems using a DSS or other collaborative software	4
		Subtotal studies using unspecified students	29
Students and profess-ionals or other non-student samples	Managers using DSS	Choosing alternatives, decision making, choosing alternatives using a graphical DSS, analyzing deception	5
	Online consumers	Browsing websites, bidding for and evaluating products	12
	B2B e-commerce participants	Entering transactions using an exchange technology	1
	Software application users	Completing skill tests	1
	Professional and administrative organizational workers	Entering transactions using an exchange technology, examining fear appeals about computer security	2
		Subtotal studies using students and professionals or other non-student samples	21
Profess-ionals or other non-student samples	Online consumers	Searching sites and making purchasing decisions, using e-negotiations, using online agents, reviewing privacy messages, examining seller information, bidding	13
	Professionals using a learning management system	Using system for continuing learning	1
	Users of various database systems	Using different systems to generate ideas, make decisions, analyze problems, deceive, reach consensus	8
	Stakeholders in the system development process	Using conceptual models, making software project continuance decisions, giving requirements for a system development project, modifying code	4
		Total studies using professionals or other non-student samples	26
		Total Studies	184

Current Practice

Twenty-one percent of the IS experiments[4] we examined do not report measuring sample characteristics. In studies that did report information on sample characteristics, only 8 percent explain the choice of all reported characteristics, another 21 percent explain the choice of some of the reported characteristics, and the final 71 percent do not explain the choice of any reported characteristics.

Recommendation

Researchers should briefly report what sample characteristics they have measured and what their theory- and evidence-based reasons were for believing that these characteristics and not others might influence the IVs and/or DVs in the study. The measures should be used to analyze effects of sample choice as the following steps recommend.

V. STEP TWO: EXAMINE VARIABILITY OF SAMPLE CHARACTERISTICS

The second step in our approach (Figure 1) is to examine the variability of these sample characteristics. Variability information plays two important roles in addressing questions about participant sample choices. First, it can support prima facie judgments about the likelihood that the sample is appropriate. If the target population is relatively homogeneous (diverse) with respect to a potentially influential characteristic (e.g., nationality or IS expertise), a representative sample will be similarly homogeneous (or its diversity will be similarly distributed) with respect to this characteristic.

Second, especially when the distribution of some potentially influential characteristic differs between the target population and the sample, the variability of sample characteristics plays an important role in the testing that enables researchers to provide persuasive support for their sample choices and their judgments about the resulting limitations (see steps 3 and 4 for more detail on these tests and their uses). Too little variability in a characteristic can make it impossible to conduct meaningful tests of association between the characteristic and the variables of interest in the study. Even when variation in the characteristic is adequate for testing, the results of the test (e.g., an assurance that the characteristic has no association with the variables of interest, and thus a mismatch between sample and target population with respect to the characteristic is unproblematic) apply only to the sample range[5]. Researchers should therefore report on the variability and the means of potentially influential sample characteristics.

Current Practice

Variability information was limited. In the experiments we examined, only 22 percent of those that reported measures of sample characteristics also provided variability information for all these characteristics; 14 percent provided no information at all about the variability of sample characteristics, and 64 percent provided variability information for some but not all of the sample characteristics reported in the study (38% of the time, this partial information consisted only of the percentages of male and female participants in the sample). The types of variability information provided were quite diverse. We counted any of the following items as providing variability information: standard deviations, ranges, percentage distributions, and threshold values (e.g., “all participants had at least three years of work experience”). The types of variability information provided often differed across sample characteristics in and across experiments. Because variability reporting in existing studies has been infrequent, it is difficult for readers to judge how diverse the participant sample in a given study is, let alone whether the choice of diversity level is appropriate.

Recommendation

Researchers should report how variable the participant sample is with respect to all potentially influential sample characteristics. Relevant variability information includes standard deviations and marked non-normality of distribution shapes, such as bimodality or strongly skewed distributions, which might make mean values non-representative. For example, a participant group with mean software-development experience of ten years may appear solidly experienced but (to take an extreme case), if one-third of the participants have around thirty years of experience and two-thirds have virtually none, the mean of ten years is misleading. Range information alone, which is often provided, has limited usefulness: for example, one or both end points of the range could be outliers that provide little information about the location of most of the observations.

---

[4] A paper can report more than one experiment and can report differently on sample characteristics for each experiment. In our sample, 25 papers reported on more than one experiment. Taking this into account, we calculated this percent based on the number of experimental studies (222).

[5] For example, a finding of no effect of age on the variables of interest, when ages in the sample are broadly distributed between 25 and 45, does not allow researchers to conclude that there is no effect of age on these variables in the 45-75 range.

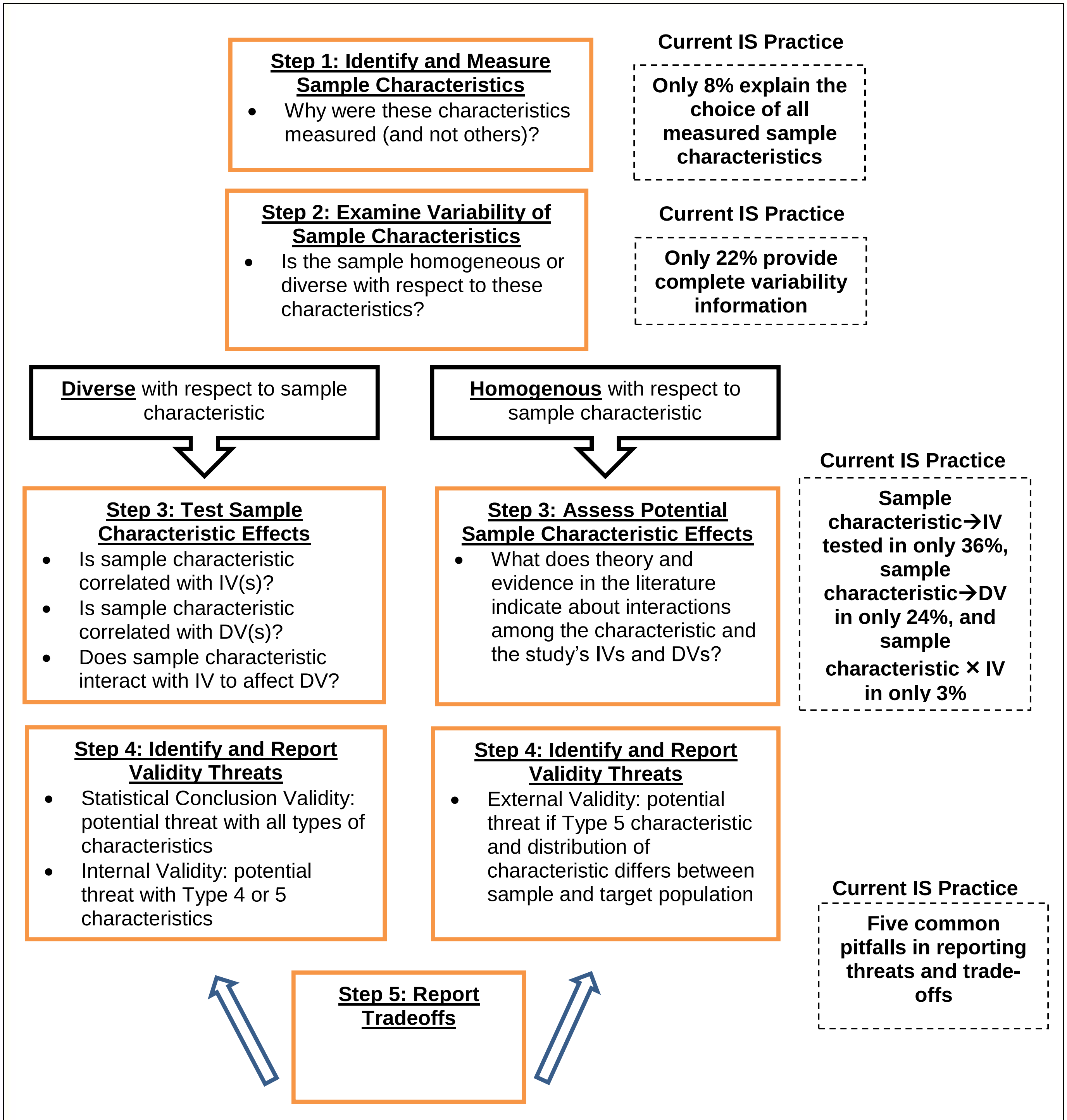

Figure 1. Five-Step Approach to Testing for and Reporting on Threats to Validity

VI. STEP 3: TEST EFFECTS OF SAMPLE CHARACTERISTICS

Step 3 is to test or otherwise assess the associations among potentially influential sample characteristics and the DVs and IVs in the study (Figure 1). The specific types and magnitudes of threats to validity posed by population and sample characteristic mismatches depend on whether and how the variance in the sample characteristic is associated with variance in the DVs and IVs in the study (see Table 4 and accompanying text below for details). These tests may, for example, allow researchers to conclude that a mismatch between sample and target population poses no threat to the validity of results. Or, conversely, the tests may indicate that the sample choice significantly limits the inferences that can be drawn from the study. Without some knowledge of the associations between sample characteristics and a study's DVs and IVs, an informed discussion of sample choice and its consequences is not possible.

When a sample is homogeneous with respect to a certain characteristic, meaningful testing for associations between the characteristic and the DVs and IVs is not possible. In such cases, a more judgmental assessment is required. If the characteristic that is homogeneous in the sample is also homogeneous in the target population, and if mean levels are similar, then the sample is representative and sample homogeneity on this characteristic poses no threats to validity. If the target population is homogeneous at a different mean level or is more diverse, then the primary threat to validity arises from interactions between the relevant characteristic and the IVs of interest (see Table 5 below and accompanying text). In such cases, prior literature, both theoretical and empirical, can sometimes offer evidence about whether such interactions exist and thus whether the mismatch between sample and target population poses a threat to the validity of the study's results.

Current Practice

Table 3 summarizes the reported testing of sample characteristics in the experimental studies we examined. Column a presents the number of experiments that measured each characteristic. Some characteristics were measured in more than one way in a single experiment. Because results of tests can differ depending on which measure was tested, we also report the total number of measures of each characteristic in square brackets in column a[6]. The remaining columns present the number of times researchers performed tests to determine whether the characteristic was correlated with the IV(s) or influenced the DV(s) (i.e., had an incremental statistical association with the DV(s), controlling for other variables in the model) either additively or interactively.

Table 3 shows that key tests were relatively infrequently performed. For example, 340 experience measures were collected in the studies summarized in Table 3; tests of experience effects on the DV were reported for 107 of the 340 measures. The infrequency of testing and lack of significant effects that appear in Table 3 (e.g., no significant effects of experience on the DV in 82 of the 107 cases) could simply be due to the absence of variation in the characteristics. The infrequency and inconsistency of variability reporting makes it difficult to judge how often this was the case.

Table 3: Testing of Sample Characteristics

Characteristic	(a) Number of experiments that reported a measure of the characteristic (# of measures)	Number of reports of tests		
		(b) Mean difference in characteristics across experimental treatments (# tests significant)	(c) Additive effect of characteristics on dependent variable (# tests significant)	(d) Interaction of characteristics with independent variables (# tests significant)
Gender	147 (147)	43 (0)	32 (4)	5 (2)
Age	137 (137)	43 (1)	17 (2)	3 (0)
Experience	130 (340)	147 (2)	98 (20)	9 (5)
Class or grade level	34 (35)	6 (0)	6 (0)	0 (0)
Education level	16 (16)	8 (0)	5 (1)	1 (0)
Personality type variables	13 (13)	6 (0)	7 (2)	0 (0)
Academic major	15 (15)	3 (0)	1 (0)	0 (0)
Trust disposition/risk propensity	12 (14)	8 (1)	6 (4)	0 (0)
Grade point average	10 (10)	5 (0)	2 (2)	0 (0)
Ethnic background	8 (8)	2 (0)	0 (0)	0 (0)
Personal relevance of task	7 (8)	2 (0)	3 (2)	1 (1)
Income	6 (6)	4 (0)	2 (0)	0 (0)
English as a second language	5 (5)	2 (0)	1 (1)	1 (1)
Citizenship	4 (4)	0 (0)	0 (0)	0 (0)
Country of birth	2 (2)	0 (0)	1 (1)	1 (1)
Voice quality	2 (6)	0 (0)	6 (5)	2 (2)
Motivation to learn	1 (1)	1 (0)	0 (0)	0 (0)
Chronic illness	1 (1)	0 (0)	0 (0)	0 (0)
Totals	550 (768)	280 (4)	187 (44)	23 (12)

[6] "Number of measures" represents the total number of measurements, not the number of types of measures. Thus, if experience is measured as "months of work experience" in thirty studies, this counts as thirty measures.

However, it is unlikely that absence of variation accounts for most of the low level of reported testing, which we can see by examining sample characteristics that are likely to vary in experimental samples (i.e., gender and, for student samples, academic major). Because 147 experiments in our sample reported gender information, 441 tests could, in principle, have been performed (147 experiments x the 3 tests in columns b–d of Table 3). However, only 80 tests (18% of the possible number) were reported. Similarly, 15 experiments recorded participants' academic major, which results in 45 possible further tests, but the studies reported only four such tests (9%).

Insofar as tests on sample characteristics were reported at all, the most common test was for mean difference in the characteristics across experimental treatments (i.e., for correlation between the sample characteristic and levels of the IV). Thirty-six percent (280) of the 768 instances of measured characteristics were tested for correlation with the IVs in the study. When researchers randomly assigned participants to the treatments, no such correlations would be expected, although differences can occur with a sufficiently bad draw from the random distribution. For example, in Sia, Tan, and Wei (2002), mean age differs significantly between groups of participants randomly assigned to face-to-face and dispersed conditions of computer-mediated communication. Age is then used as a control variable (covariate) in this study's hypothesis tests to provide assurance that the apparent effect of different settings on group processes is not due instead to differences in age. Over half the studies in our sample provided no assurance on this point.

Sample characteristics are more likely to be correlated with measured IVs than with manipulated, randomly assigned IVs. For example, when experience of a particular type is the IV of interest, it may well be correlated with age, amount and type of education, and various personality or attitude characteristics that influence individuals' dispositions to take certain jobs and acquire certain experience. When such correlations exist, they can raise questions about the validity of the tests of IV effects on the DV. Although 36 of the 184 papers in our sample included measured IVs, only two reported testing for correlations between the sample characteristics and the IVs.

Sample characteristics that influence the DV can provide alternative explanations for DV variance, which competes with the explanation provided by the IVs in which the researcher is interested. Or they can limit the results' range of generalizability (see Section 7 for details). These effects of sample characteristic on DVs can be either additive (independent of IV effects) or interactive. Ideally, researchers should test for both. Among the 768 measures of sample characteristics we identified, we found only 187 (24%) instances of tests for additive effects on the DVs. Testing for interactive effects of sample characteristics was even rarer, being reported for only 23 (3%) of the 768 measured characteristics.

Recommendations

When there is sufficient variance in sample characteristics to make such tests meaningful, researchers should report tests of sample characteristics' correlations with IV(s) and their additive and interactive effects on the DV(s). This information about the relations between sample characteristics and other IS-related variables can be valuable not only for analyzing threats to the validity of a study's results, but also for guiding future participant choices and assisting researchers who want to build and test theory related to these characteristics. When a potentially influential characteristic is homogeneous within the sample, researchers should report their reasons for believing that the characteristic either is similarly homogeneous in the population or does not interact with the IVs in the study. Such tests and assessments were relatively infrequent in the studies we examined, and the absence of variability information often made it difficult to determine when each approach would have been appropriate.

VII. STEP 4: IDENTIFY AND REPORT VALIDITY THREATS

Step 4 is to identify and report validity threats resulting from the distribution of potentially influential characteristics in the sample (Figure 1). We focus on three of the four validity types that Shadish et al. (2002) define:[7]

1. Statistical conclusion validity: do the DVs and IVs actually covary (or not) when the statistical tests in the study indicate that they do (or do not)?

2. Internal validity*:* is the observed covariance causal[8]?

[7] We do not list all the substantive sources of threats to validity that appear in Shadish et al. (2002), such as "selection", "history", etc. Instead, we provide a compact, structured framework for thinking about these substantive issues in terms of statistical inference problems.

[8] This view of internal validity assumes that the hypotheses being tested are causal in intent (e.g., computer-mediated communication causes different behavior than face-to-face communication), rather than hypotheses about parameter values (e.g., the mean return on investment in IT is greater than 10%).

3. External validity (generalizability): do the study's conclusions generalize beyond the sample and experimental setting employed?

The fourth validity type that Shadish et al. (2002) identify is construct validity, which is whether the measured or manipulated variables capture the theoretical constructs of interest in the study. We do not include construct validity here because characteristics of the individuals included in the sample typically do not influence the construct validity of the variables of interest in the study. For example, if the DV is a questionnaire measure of trust in an information system or intention to use it, there is typically little reason to believe that the questionnaire will be a good measure of the construct for men but not for women, or will be a good measure for individuals in their twenties but not individuals in their forties[9].

Table 4 defines and presents examples of five sample characteristic types with differing statistical properties. We first describe these five types and then show how researchers can use this typology to identify the specific threats to validity that are posed—and often *not* posed—by imperfectly representative samples.

Table 4: Five Types of Sample Characteristics

Sample characteristic	Relation to IV and DV
Type 1: uncorrelated	Uncorrelated with any IVs and does not influence DV
Type 2: IV-correlated only	Correlated with one or more IVs but does not influence DV
Type 3: DV-correlated only	Influences DV but is not correlated with IVs and does not interact with any IVs to influence DV
Type 4: IV- and DV-correlated	Influences DV and is correlated with one or more IVs, but does not interact with any IVs to influence DV
Type 5: interacting	Interacts with an IV to influence DV, and may or may not be correlated with IVs

Types 1 and 2 are characteristics that are potentially associated with the DVs and/or IVs of the study. That is, researchers believe that the characteristics might be associated with the DVs or IVs a priori. But testing reveals that, even with reasonable variation in the characteristic, the characteristics are not associated with the DV(s)—and in type 1, with the IVs. For example, in their investigation of e-commerce trust, Kim and Benbasat (2006) found that a sample characteristic, online shopping frequency, was uncorrelated with trust-assuring argument displays (the IV), and had no statistically significant association with trust beliefs (the DV). It was reasonable ex ante to suppose that such correlations might exist and hence the researchers tested for them. But the tests demonstrated no relation between the sample characteristic and the variables of interest in the study.

Type 2 sample characteristics are correlated with one or more IVs but do not influence the DV. That is, controlling for the IVs, the characteristic has no significant incremental association with the DV. For example, in Sia et al. (2002), age was a type 2 sample characteristic. Age was significantly higher in one of the experiment's computer-mediated communication treatments than in the other (thus it was correlated with the computer-mediated communication IV). However, age had no significant incremental association with the DVs (choice shift and preference change).

Type 3 sample characteristics influence one or more DVs, but do not correlate with any IVs or interact with them to influence the DV (i.e., their effect on the DV is additive). As an example of a type 3 sample characteristic, Piccoli, Ahmad, & Ives (2001) found that gender had a significant influence on the DVs (performance, satisfaction, and self-efficacy) and did not differ between learning environment treatments (virtual vs. traditional). While the authors did not report the results of any interaction tests, we assume for the convenience of this example that there are no interactions between gender and the learning environments.

Type 4 sample characteristics are identical to type 3 sample characteristics except that they are correlated with one or more IVs. Mennecke, Crossland, and Killingsworth (2000) provide an example of a type 4 sample characteristic. In their study, the sample characteristic "task interest" significantly influenced the DV (time spent on the task) and was also correlated with one of the IVs, expertise.

[9] However, errors in measuring the sample characteristics (a problem analogous to construct validity issues for IVs and DVs) can threaten the validity of the study's inferences via threats to statistical conclusion validity or internal validity. These threats are described in the following sections.

Type 5 sample characteristics interact with an IV to influence the DV, as illustrated in Allen and March (2006). They found that a sample characteristic, comfort level in writing queries, interacted with the treatment (ontological foundation) to significantly influence the DV (prediction of accuracy)[10].

Researchers can use the five sample characteristic types shown in Table 4 to identify different threats to validity arising from potentially influential sample characteristics[11]. Table 5 presents the potential threats to the three validity types that can occur for each sample characteristic type. The specific threats that occur depend on whether the characteristic is relatively diverse or homogeneous within the sample and, if diverse, whether the characteristic is included in the empirical model used to test hypotheses[12].

Table 5: Validity Threats Arising from Variation in Sample Characteristics

	Experiment participant and empirical model choices		
Sample characteristic types for which threat can occur (see Table 4)	Diverse sample characteristics not measured and not included in model	Diverse sample characteristics measured and included in model	Homogeneous sample characteristics
Type 1		Statistical conclusion validity threats (reduced degrees of freedom)	
Type 2		Statistical conclusion validity threats (reduced degrees of freedom, multicollinearity	
Type 3	Statistical conclusion validity threats (unexplained variability in Y)	Statistical conclusion validity threats (measurement error)	
		Internal Validity Threats (measurement and specification error)	
Type 4	Statistical conclusion validity threats (unexplained variability in Y)	Statistical conclusion validity threats (multicollinearity, measurement error)	
	Internal Validity threats (correlated omitted variable)	Internal validity threats (measurement and specification error)	
Type 5	Statistical conclusion validity threats (unexplained variability in Y)	Statistical Conclusion validity threats (multicollinearity, measurement error)	
	Internal validity threats (correlated omitted variable, aggregation error)	Internal validity threats (measurement and specification error)	
	External validity threats		External validity threats

[10] Type 5 sample characteristics may or may not be correlated with IVs. If they are correlated, then the correlation raises issues similar to those for Type 4 sample characteristics. Therefore, in discussing type 5 sample characteristics, we focus only on the interaction implications.

[11] Each threat to validity is represented independently in our discussion. For example, in presenting threats to external validity (generalizability), we assume that a statistically and internally valid inference has been drawn. The focus is then on whether the inference is also valid for settings outside the laboratory and individuals other than the participants actually used in the experiment.

[12] When a characteristic is homogeneous in the sample, its lack of variance will insure that it has no significance in the model. Hence, including it will be uninformative.

Diverse Samples: Sample Characteristics Not Measured and Not Modeled

It is sometimes difficult to be certain about what the relevant sample characteristics and their statistical properties are in a population of interest, and good measures and models for relevant characteristics are not always available. Perhaps because of these difficulties, sample characteristics are often omitted from empirical models used in hypothesis testing, which Table 3 indicates. With diverse samples, this can pose threats to all three of the types of validity we consider here.

First, if the omitted characteristics influence the DV (characteristic Types 3, 4, or 5), they will create unexplained variability in the DV (a large error term in the model). As such, they will weaken the power of statistical tests and create threats to statistical conclusion validity. Larger samples and/or stronger manipulations that produce larger mean effects are straightforward ways of dealing with this threat. Other threats resulting from unmeasured diversity are not so easily mitigated, however.

Second, if the omitted characteristics are types 4 or 5, then they are correlated omitted variables, which can pose important threats to internal validity[13]. Omitting a variable that is positively correlated with an IV inflates the estimated coefficient on the IV in the empirical model, which potentially results in a significant coefficient even when the IV has no causal influence on the DV. Conversely, omitting a variable that is negatively correlated with an IV reduces the estimated coefficient on the IV, which potentially results in a non-significant coefficient even when the IV actually has a significant influence on the DV (see MacKinnon, Krull, and Lockwood (2000) for a detailed discussion of the effect of correlated variables).

The third threat to validity from diverse samples with omitted sample characteristics is a threat to generalizability (external validity) that occurs with type 5 (interacting) characteristics. If sample characteristics are not measured, it is difficult to judge how representative a diverse sample is. The proportion of individuals with high and low values on a particular characteristic may very well differ between the participant sample and the target population. The numerical example in Table 6 illustrates how this can lead to generalizability problems for type 5 (but not other) characteristics.

Table 6: Example of Additive Effects versus Interaction Effects

Types 3 and 4 sample characteristics: additive effects on Y				Type 5 sample characteristic: interaction effect on Y			
Forecasting performance (Y)	Low problem-solving ability	High problem-solving ability	Means	Forecasting performance (Y)	Low statistical knowledge	High statistical knowledge	Means
Without DSS	40	70	55	Without DSS	70	40	55
With DSS	60	90	75	With DSS	40	70	55
Means	50	80	65	Means	55	55	55
Cell entries are forecasting performance on a 0–100 scale							

In this example, the sample characteristics (problem-solving ability and statistical knowledge) are characterized for simplicity as either high or low, and both have significant effects on the DV, forecasting performance. In the additive example (a type 3 or 4 characteristic), mean forecasting performance is higher by twenty points with a decision support system (DSS) than without it—and this is true for both low-ability and high-ability individuals. In the interaction example (a type 5 characteristic), in contrast, the mean effect of the DSS on performance is not the same for individuals with high and low values of the characteristic. Those with high statistical knowledge forecast more accurately when they use a DSS than when they do not, but those with low statistical knowledge are more accurate without the DSS, perhaps because the DSS requires statistical knowledge for effective use and confuses individuals with low knowledge.

If the type 5 characteristic, statistical knowledge, is not included in the model, and if roughly equal numbers of high- and low-knowledge individuals are in the sample, then mean forecasting performance will appear identical with and without the DSS, which the marginal means on the right-hand side of Table 6 show. Not only is this result an incorrect inference about the real (non-zero) effect of the DSS (an internal validity problem due to the omitted interaction variable), but it also has limited external validity. The conclusion that the DSS has no mean effect on forecasting performance will not generalize to any population that is not a 50-50 mix of low- and high-knowledge

[13] Even if type 5 sample characteristics are not themselves correlated with an IV, an interaction between a characteristic and an IV will be correlated with the IV. Thus, when the characteristic is not measured and included (as an interaction term) in the empirical model, the interaction will be a correlated omitted variable.

individuals. Moreover, if there is a clear separation between individuals who have enough knowledge or perform well with the DSS and those who do not, then the mean effect of DSS use in a mixed-experience group does not generalize to any individual because no individual has 50 percent high knowledge and 50% percent low knowledge. Thus, the null main effects in this example are aggregation errors and generalize neither to different samples or populations, nor to any specific individuals (see Lynch, 1999 for a discussion of aggregation errors).

Diverse Samples: Sample Characteristics Measured and Modeled

The threats to validity that are posed by omitting potentially relevant sample characteristics can, in principle, be obviated by measuring these characteristics and including them in the empirical models used for testing hypotheses. However, especially when sample-characteristic measures are imperfect and/or the relations of the characteristics with IVs and DVs are uncertain, including the characteristics in empirical models can create other threats to statistical conclusion validity and internal validity.

First, if a number of additional sample characteristics with little explanatory value (types 1 and 2) are included in a model “just in case”, then the model can lose power because the additional variance explained by the sample characteristics is not sufficient to compensate for losing degrees of freedom. Although it is straightforward to solve this problem by re-estimating the model without the uninformative characteristics, other threats are not so easily mitigated.

A second threat, multicollinearity, arises when characteristics are correlated with one or more IVs (type 2, 4, and 5). Multicollinearity inflates standard errors of the IV coefficient estimates, which results in threats to statistical conclusion validity due to imprecise coefficient estimates and low-power hypothesis tests. Widely used rules-of-thumb for identifying multicollinearity problems tend to understate these problems. For example, it is common to regard variance inflation factors (VIFs) greater than 5 or 10 as indicators of multicollinearity problems. But lower VIFs can still require a doubling or tripling of sample size, relative to a sample without correlations among the IVs and sample characteristics, to maintain adequate power (see Hsieh, Bloch, and Larsen (2003) for further information on VIFs and sample sizes). Note that, if a sample characteristic is highly correlated with the IV, it can be difficult for researchers to determine whether it has incremental explanatory power for the DV (a type 4 or 5 characteristic rather than a type 2 characteristic) using only sample data because the correlation will make it difficult to disentangle effects of the sample characteristic and the IV. In such cases, theory that helps to judge the likelihood of a causal relation between the sample characteristic and the DV, and empirical evidence from other studies with different correlation structures in the sample, can provide a basis for classifying the characteristic as type 2 or a type 4 or 5.

Third, measurement and specification errors related to the sample characteristics can pose threats to valid inference when researchers use diverse samples and include the characteristics in empirical models including characteristic types 3, 4, and 5. For example, measures of individuals’ experience and knowledge often capture the underlying constructs with some error, especially when the construct researchers want to control for is task-relevant knowledge and the measure is years of work experience. This error can create both inconsistency and bias in the estimated coefficients, which results in threats to both statistical conclusion validity and internal validity (Greene, 2000; Wooldridge, 2006)[14].

Errors in specifying the functional form of the relation between a sample characteristic (types 3, 4, and 5) and the DV can also bias the coefficients on the IVs. For example, suppose that there are diminishing returns to experience (a curvilinear relation between task experience and task performance that can be represented by an experience-squared term in the model). Suppose further that the relation between experience and performance is modeled as linear in the empirical analysis, with the quadratic term omitted from the model. If experience (and thus the omitted experience-squared term) is correlated with an IV in the empirical model, then the experience-squared term is a correlated omitted variable, with the potential to bias the estimated coefficient on the IV. Moreover, when a sample characteristic has a curvilinear effect on the DV and the quadratic term is omitted from the model, tests for interaction effects can be distorted: they can show no effect when an interaction actually exists or show a different form of interaction than actually exists (Ganzach, 1997). Because interaction effects are important both for understanding how an IV affects the DV in the sample and for identifying limits to the generalizability of the effect, this specification problem can be significant.

Thus, using diverse, representative samples and controlling for sample-characteristic effects in empirical models is not always an effective strategy for avoiding significant threats to valid inference. In particular, multicollinearity and

[14] The existence and nature of these threats depends on the structure of the correlations among the measurement error and the IVs. Even when this correlation structure is known, “the sizes and even the directions of the biases (in coefficient estimates) are not easily derived” (Wooldridge, 2006, p. 320).

sample-characteristic measurement or model-specification errors can result in invalid inferences about the IV–DV relations a study investigates.

Homogeneous Samples

Participant samples that are relatively homogeneous with respect to relevant characteristics[15] avoid the threats to valid inference that arise with more diverse samples. Characteristics that vary little in the sample will not create unexplained variance in the DV, nor will they create variance incorrectly attributed to an IV if they are omitted from the model. Because the sample characteristics do not need to be included in the model, characteristic measurement and model-specification errors do not threaten the validity of hypothesis tests.

The usual concern about homogeneous samples is external validity, but, as Table 4 indicates, only one type of sample characteristic—type 5, which interacts with an IV—actually creates external validity threats. External validity concerns do not arise with non-interacting characteristics, even when they have significant effects on the DVs, and the sample and target population differ significantly with respect to the characteristic. In the hypothetical example of additive effects in Table 6, individuals with low problem-solving ability do not forecast as well as individuals with high forecasting ability. But the effect of DSS use on forecasting performance for low-ability individuals (a 20-point improvement) generalizes to high-ability individuals and vice versa. Thus, the failure to match target population and sample on characteristics with additive effects only (types 3 and 4) has no effect on external validity[16]. An experiment using only low-ability, only high-ability, or any mix of low- and high-ability individuals will show a mean DSS effect of 20 points on forecasting performance. In consequence, when effects are additive, the researcher defining a target population does not need to specify its ability level or find a participant sample that has the same ability level or mix as the target population. Only the type 5 (interacting) characteristic, illustrated on the right-hand side of Table 5, poses a threat to external validity when sample and target population are not matched with respect to the characteristic.

Current Practice and Recommendations

Because of the close relation between identifying and reporting threats to validity (step 4) and reporting trade-offs based on the importance of the threat(s) (step 5), we provide current practice and recommendations for both steps at the end of Section 8.

VIII. STEP 5: ANALYZING AND REPORTING SAMPLE-CHOICE TRADEOFFS

Sample choice, like other elements of research design, often cannot minimize all threats to validity simultaneously at a reasonable cost. Even well-conducted studies will often fail to eliminate some threats to validity. When researchers have identified the threats to validity that their study has not eliminated, their task is then to explain how their sample choice represents an appropriate tradeoff; that is, how they are accepting some smaller threats to reduce other larger ones.

Researchers who consider the statistical characteristics summarized in Table 4 and the resulting specific threats to validity (Table 5) can help themselves judge which threats are likely to be large and which are not. For example, the magnitude of threats to statistical conclusion validity resulting from power limitations will be a relatively large threat when specialized study requirements keep sample sizes small, but not when large samples can overcome the power problems. The magnitude of threats to external validity resulting from unrepresentative samples will be large when existing literature indicates a likelihood of IV x sample characteristic interactions (or this likelihood is altogether unknown), but not when existing literature provides evidence against interactions.

Current Practice

Many studies report on participant choice and its consequences with qualitative comments in their methods or conclusion sections. To some extent these qualitative comments address validity threats, but often not in ways that can clearly be matched to each of the specific threats identified in Table 5 or to the tradeoffs among them. In this section, we identify the five types of qualitative comments that we found most frequently reported in the studies we analyzed. For each, we specify common limitations of the comments, which, in many cases, could be readily overcome, which makes these comments more informative and persuasive.

[15] That is, sample characteristics that are statistically associated with an IV and/or DV in the population or in a diverse sample.
[16] Recall also that failure to match sample and target population on type 2 characteristics, which have no effect on the DV and are correlated with an IV, will have no effect on external validity.

Comment 1. Homogeneity: IS researchers occasionally cite the value of a homogeneous sample in reducing extraneous variation in behavior. This is an important point, consistent with the value of homogeneity in reducing threats to validity as represented in Table 5.

Limitations: This type of qualitative comment will be more persuasive when it is supported by two kinds of evidence:

- variability information on relevant sample characteristics that confirms the homogeneity, and
- evidence (so far as it is available, either from other research or from such variability of the sample characteristic as exists in the study) that the IVs in the study do not interact with the sample characteristics over the range of the characteristic observed in the target population.

Comment 2. Sample size: For studies that require large sample sizes, the participant choice is sometimes justified by the observation that large numbers were easy to obtain with a particular participant group. Sufficiently large sample sizes can reduce the threats to statistical conclusion validity that arise from the various sources of reduced statistical power that are summarized in Table 5.

Limitations: If sample-size considerations are the only defense provided for participant choice—as they sometimes are—they are insufficient. A large sample size does not help with the internal validity (biased hypothesis tests) or external validity (generalizability) concerns summarized in Table 5. Hence, authors should provide some assurance about these threats to persuade readers that the study's participant choice represents a favorable tradeoff between statistical power and other validity concerns.

Comment 3. Prior valid use of a participant group: IS researchers sometimes justify the choice of a particular participant group based on the fact that the group has been used in prior high-quality IS research, and results have tended to generalize to diverse populations.

Limitations:. This justification is not very informative unless it is specific. For example, suppose that the concern about participant choice in a particular study is whether the results are generalizable from a low-experience participant sample to a population that includes higher-experience individuals. The concern is whether experience interacts with the IVs of interest in the study. The fact that experience does not interact with different IVs used in prior literature is not very informative on this point. To provide a defense for the choice of participants, the prior research that is cited needs to share IVs with the study in question, and the prior research needs to indicate that experience does not interact with these IVs.

Comment 4. Participants are real users: IS researchers sometimes present their participant choice as appropriate because the study is about (for example) online consumers or decision support system users, and the participants are in fact online consumers or decision support system users.

Limitations: Results from one group of "real" users do not necessarily generalize to other groups of "real" users. For example, Ko and Dennis (2011) find that how much knowledge workers in an organization benefit from a knowledge management system depends on the workers' specific prior job experience and when the research data are gathered (shortly after the introduction of the system or later). If the authors had been able to gather data only at one point in time, or only from knowledge workers with a narrow range of experience, then the fact that their participants were "real" knowledge workers would have been no guarantee of generalizability. All the validity tradeoff concerns summarized in Table 5 apply equally to all samples, and observing that the participants are "real" is not a substitute for analyzing these tradeoffs.

Comment 5. Recommending replications: Some IS experimental studies acknowledge that their participant samples are limited and therefore their results should be replicated with different samples.

Limitations: Although replications can be valuable, general recommendations for replication can be uninformative, in that they can leave the impression that the key to generalizability is simply testing over and over again with differing samples. Building theory- and evidence-based arguments identifying the sample characteristics that are likely to interact with the variables of interest and focusing on testing these interactions is likely to be a more fruitful strategy. Hence, recommendations for replications are likely to be more meaningful when they are more clearly focused on potential interaction issues.

Recommendation

A good research design (including sample, measurement, and modeling choices) is one that avoids large threats, though possibly at the expense of incurring smaller threats. These tradeoffs should be considered when researchers embark on a study, and we recommend that IS researchers explicitly discuss the tradeoffs between specific validity threats that result from the participant choices made in their studies to persuade readers that their sample choice is appropriate. The smaller threats should be appropriately identified as limitations.

For example, if theory and previous research (using diverse samples) demonstrate that there are no interactions (type 5) between a certain sample characteristic and the independent variables in the study, then researchers can select a sample that is homogenous with respect to that characteristic (and measure the characteristic to document the homogeneity) because this avoids the other potential threats to validity. Conversely, if theory supports the existence of an interaction, or interaction effects have been demonstrated, then researchers should select a representative sample with regard to that characteristic and carefully measure it because this is the only way to identify the limits to generalizability that are created by interacting (type 5) characteristics. When it is uncertain ex ante whether a characteristic interacts with a study's independent variables, then researchers need to make—and explain—their judgments about the likelihood of interaction, the importance of broad generalizability in their study, and the likely magnitude of the threats to validity that can result from using a diverse sample. For example, if a characteristic can be reliably measured and modeled (and this is documented in the study), then the threats to validity that arise from using a diverse sample are lower, and a diverse sample can be attractive even when the likelihood of interaction is only moderate. If well-validated measures and models are not available and the reasons to expect interaction are weaker, then the validity tradeoffs can favor a homogeneous sample.

V. CONCLUSION

Based on our proposed framework, we analyze experimental studies appearing in four major IS journals. We identify and analyze their participants, target populations, and tasks (Table 2), and show that there is not always a prima facie match between samples and target populations. We provide a prescriptive, step-by-step approach for testing and reporting on potential threats to validity resulting from participant choice (Figure 1). We show that current practice dealing with sample characteristic variation in IS experiments has some features that are valuable and should be maintained. For example, researchers sometimes collect sample characteristic data and perform tests (e.g., for equality of means across experimental treatments) that help to allay concerns about the validity of study inferences. We also show where current practice in IS research could be improved.

As our analysis illustrates, the questions that arise about sample characteristic influences are often both methodological questions about how best to conduct and analyze experiments and substantive questions about the identity and form of significant influences on IT-related behavior. Future experiments can test and refine theory in this area, and more refined theory can be used to support more informative experimental tests in the future. Efforts toward "realism" in participant choice by simply using more diverse or more experienced participants are likely to be less helpful than building a better theory-based understanding of the relations between sample characteristics and IV/DVs in IS research. As Lynch (1999, p. 368) observes:

> *The only path to understanding the generality of one's findings is to have a theory that specifies moderator (interacting) variables and boundary conditions and specifies what variables should not moderate the findings reported, and to test for the asserted pattern of interactions. If one's theory is impoverished, no degree of adherence to methodological prescriptions will help 'ensure' external validity.*

REFERENCES

Author's note: The complete list of all papers in our sample is in Appendix A.

Editor's Note: The following reference list contains hyperlinks to World Wide Web pages. Readers who have the ability to access the Web directly from their word processor or are reading the paper on the Web, can gain direct access to these linked references. Readers are warned, however, that:

1. These links existed as of the date of publication but are not guaranteed to be working thereafter.
2. The contents of Web pages may change over time. Where version information is provided in the References, different versions may not contain the information or the conclusions referenced.
3. The author(s) of the Web pages, not AIS, is (are) responsible for the accuracy of their content.
4. The author(s) of this article, not AIS, is (are) responsible for the accuracy of the URL and version information.

Allen, G. N., & March, S. T. (2006). The effects of state-based and event-based data representation on user performance in query formulation tasks. *MIS Quarterly, 30*(2), 269-290.

Compeau, D., Marcolin, B., Kelley, H., & Higgins, C. (2012). Generalizability of information systems research using student subjects—a reflection on our practices and recommendations for future research. *Information Systems Research*, *23*(4), 1093-1109.

Ganzach, Y. (1997). Misleading interaction and curvilinear terms. *Psychological Methods, 2*(3), 235-247.

Gordon, M. E., Slade, L. A., & Schmitt, N. (1986). The "science of the sophomore" revisited: From conjecture to empiricism. *The Academy of Management Review, 11*(1), 191-207.

Greene, W. H. (2003). *Econometric analysis* (4th Ed.). Upper Saddle River, NJ: Prentice Hall.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*, 61-135.

Hsieh, F. Y., Bloch, D. A., & Larsen, M. D. (1998). A simple method of sample size calculation for linear and logistic regression. *Statistics in Medicine*, *17*, 1623-1634.

Kim, D., & Benbasat, I. (2006). The effects of trust-assuring arguments on consumer trust in Internet stores: Application of Toulmin's model of argumentation. *Information Systems Research, 17*(3), 286-300.

Ko, D. G., & Dennis, A. R. (2011). Profiting from knowledge management: The impact of time and experience. *Information Systems Research, 22*(1), 134-152.

Lowry, P. B., Gaskin, J., Humpherys, S. L., Moody, G. D., Galletta, D. F., & Barlow, J. B. (2013). Evaluating journal quality and the Association of Information Systems Senior Scholars' Journal Basket via bibliometric measures: Do expert journal assessments add value? *MIS Quarterly*, 37(4), 993-1012.

Lynch, J. G. (1999). Theory and external validity. J*ournal of the Academy of Marketing Science, 27*(3), 367-376.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science, 1*(4), 173–181.

Mennecke, B. E., Crossland, M. D., & Killingsworth, B. L. (2000). Is a map more than a picture? The role of SDSS technology, subject characteristics, and problem complexity on map reading and problem solving. *MIS Quarterly, 24*(4), 601-629.

Piccoli, G., Ahmad, R., & Ives, B. (2001). Web-based virtual learning environment: A research framework and a preliminary assessment of effectiveness in basic IT skills training. *MIS Quarterly, 25*(4), 401-426.

Rosenthal, R., & Rosnow, R. L. (2009). *Artifacts in behavioral research*. Oxford: Oxford University Press.

Sears, D. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology, 51*(3), 515-30.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton-Mifflin.

Sia, C. L., Tan, B. C. Y., & Wei, K. K. (2002). Group polarization and computer-mediated communication: Effects of communication cues, social presence and anonymity. *Information Systems Research, 13*(1), 70-90.

Wooldridge, J. M. (2006). *Introductory econometrics: A modern approach* (3rd Ed.). Cincinnati, Ohio: South-Western College Publishing.

APPENDIX A: EXPERIMENTAL STUDIES

Adipat, B., Zhang, D., & Zhou, L. (2011). The effects of tree-view based presentation adaptation and mobile web browsing. *MIS Quarterly*, *35*(1), 99-121.

Alavi, M., Marakas, G. M., & Yoo, Y. (2002). A comparative study of distributed learning environments on learning outcomes. *Information Systems Research*, *13*(4), 404-415.

Allen, G., & Parsons, J. (2010). Is query reuse potentially harmful? Anchoring and adjustment in adapting existing database queries. *Information Systems Research, 21*(1), 56-77.

Allen, G. N., & March, S. T. (2006). The effects of state-based and event-based data representation on user performance in query formulation tasks. *MIS Quarterly, 30*(2), 269-290.

Allen, G. N., & March, S. T. (2012). A research note on representing part-whole relations in conceptual modeling. *MIS Quarterly*, *36*(3), 945-964.

Al-Natour, S., Benbasat, I., & Cenfetelli, R. T. (2006). The role of design characteristics in shaping perceptions of similarity: The case of online shopping assistants. *Journal of the Association for Information Systems, 7*(12), 821-861.

Alnuaimi, O. A., Robert, L. P., & Maruping, L. M. (2010). Team size, dispersion, and social loafing in technology-supported teams: A perspective on the theory of moral disengagement. *Journal of Management Information Systems*, *27*(1), 203-230.

Anderson, C. I., & Agarwal, R. (2010). Practicing safe computing: A multimethod empirical examination of home computer user security behavioral intentions. *MIS Quarterly, 34*(3), 613-643.

Andres, H. P., & Zmud, R. W. (2001-2). A contingency approach to software project coordination. *Journal of Management Information Systems*, *18*(3), 41-70.

Angst, C. M., & Agarwal, R. (2009). Adoption of electronic health records in the presence of privacy concerns: The elaboration likelihood model and individual persuasion. *MIS Quarterly, 33*(2), 339-370.

Appan, R., & Browne, G. J. (2010). Investigating retrieval-induced forgetting during information requirements determination. *Journal of the Association for Information Systems, 11*(5), 250-275.

Appan, R., & Browne, G. J. (2012). The impact of analyst-induced misinformation on the requirements elicitation process. *MIS Quarterly*, *36*(1), 85-106.

Arnold, V., Clark, N., Collier, P. A., Leech, S. A., & Sutton, S. G. (2006). The differential use and effect of knowledge-based system explanations in novice and expert judgment decisions. *MIS Quarterly, 30*(1), 79-97.

Ba, S., & Pavlou, P. A. (2002). Evidence of trust building technology in electronic markets: Price premiums and buyer behavior. *MIS Quarterly, 26*(3), 243-268.

Balijepally, V., Mahapatra, R., Nerur, S., & Price, K. H. (2009). Are two heads better than one for software development? The productivity paradox of pair programming. *MIS Quarterly, 33*(1), 91-118.

Barkhi, R. (2001-2). The effects of decision guidance and problem modeling on group decision-making. *Journal of Management Information Systems*, *18*(3), 259-282.

Benlian, A., Titah, R., & Hess, T. (2012). Differential effects of provider recommendations and consumer reviews in e-commerce transactions: An experimental study. *Journal of Management Information Systems*, *29*(1), 237-272.

Bera, P., Burton-Jones, A., & Wand, Y. (2011). Guidelines for designing visual ontologies to support knowledge identification. *MIS Quarterly*, *35*(4), 883-908.

Biocca, F., Owen, C., Tang, A., & Bohil, C. (2007). Attention issues in spatial information systems: Directing mobile users' visual attention using augmented reality. *Journal of Management Information Systems*, *23*(4), 163-184.

Biros, D. P., George, J. F., & Zmud, R. W. (2002). Inducing sensitivity to deception in order to improve decision making performance: A field study. *MIS Quarterly, 26*(2), 119-144.

Bodart, F., Patel, A., Sim, M., & Weber, R. (2001). Should optional properties be used in conceptual modeling? A theory and three empirical tests. *Information Systems Research, 12*(4), 384-405.

Bolton, G., Loebbecke, C., & Ockenfels, A. (2008). Does competition promote trust and trustworthiness in online trading? An experimental study. *Journal of Management Information Systems*, *25*(2), 145-169.

Bowen, P. L., O'Farrell, R. A., & Rohde, F. H. (2006). Analysis of competing data structures: Does ontological clarity produce better end user query performance. *Journal of the Association for Information Systems, 7*(8), 514-544.

Bowen, P. L., O'Farrell, R. A., & Rohde, F. H. (2009). An empirical investigation of end-user query development: The effects of improved model expressiveness vs. complexity. *Information Systems Research, 20*(4), 565-584.

Browne, G. J., & Rogich, M. B. (2001). An empirical investigation of user requirements elicitation: Comparing the effectiveness of prompting techniques. *Journal of Management Information Systems*, *17*(4), 223-249.

Browne, G. J., Pitts, M. G., & Wetherbe, J. C. (2007). Cognitive stopping rules for terminating information search in online tasks. *MIS Quarterly, 31*(1), 89-104.

Burton-Jones, A., & Meso, P. (2008). The effects of decomposition quality and multiple forms of information on novices' understanding of a domain from a conceptual model. *Journal of the Association for Information Systems, 9*(12), 748-802.

Burton-Jones, A., & Meso, P. N. (2006). Conceptualizing systems for understanding: An empirical test of decomposition principles in object-oriented analysis. *Information Systems Research, 17*(1), 38-60.

Chen, L., Goes, P., Mardsen, J. R., & Zhang, Z. (2009-10). Design and use of preference markets for evaluation of early stage technologies. *Journal of Management Information Systems*, *26*(3), 45-70.

Chen, Y., Ramamurthy, K., & Wen, K.-W. (2012-13). Organizations' information security policy compliance: Stick or carrot approach? *Journal of Management Information Systems*, *29*(3), 157-188.

Chidambaram, L., & Tung, L. L. (2005). Is out of sight, out of mind? An empirical study of social loafing in technology-supported groups. *Information Systems Research, 16*(2), 149-168.

Chiravuri, A., Nazareth, D., & Ramamurthy, K. (2011). Cognitive conflict and consensus generation in virtual teams during knowledge capture: Comparative effectiveness of techniques. *Journal of Management Information Systems*, *28*(1), 311-350.

Chung, W., Chen, H., & Nunamaker, J. F. (2005). A visual framework for knowledge discovery on the web: An empirical study of business intelligence exploration. *Journal of Management Information Systems*, *21*(4), 57-84.

Cyr, D., Head, M., Larios, H., & Pan, B. (2009). Exploring human images in website design: A multi-method approach. *MIS Quarterly, 33*(3), 539-566.

Dabbish, L., & Kraut, R. (2008). Awareness displays and social motivation for coordinating communication. *Information Systems Research, 19*(2), 221-238.

Dang, Y., Zhang, Y., Chen, H., Brown, S. A., Hu, P. J.-H., & Nunamaker, J. F. (2012). Theory-informed design and evaluation of an advanced search and knowledge mapping system in nanotechnology. *Journal of Management Information Systems*, *28*(4), 99-127.

Deng, L., & Poole, M. S. (2010). Affect in web interfaces: A study of the impacts of web page visual complexity and order. *MIS Quarterly, 34*(4), 711-730.

Dennis, A. R., Robert, L. P., Curtis, A. M., Kowalczyk, S. T., & Hasty, B. K. (2012). Trust is in the eye of the beholder: A vignette study of postevent behavioral controls' effects on individual trust in virtual teams. *Information Systems Research*, *23*(2), 546-558.

Dimoka, A. (2010). What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. *MIS Quarterly, 34*(2), 373-396.

Dou W., Lim, K. H., Su, C., Zhou, N., & Cui, N. (2010). Brand positioning strategy using search engine marketing. *MIS Quarterly, 34*(2), 261-279.

Dunn, C. L., Gerard, G. J., & Grabski, S. V. (2011). Diagrammatic attention management and the effect of conceptual model structure on cardinality validation. *Journal of the Association for Information Systems, 12*(8), 585-605.

Everard, A., & Galletta, D. F. (2005-6). How presentation flaws affect perceived site quality, trust, and intention to purchase from an online store. *Journal of Management Information Systems*, *22*(3), 55-95.

Fang, X., Hu, P. J.-H., Chau, M., Hu, H.-F., Yang, Z., & Sheng, O. R. L. (2012). A data-driven approach to measure web site navigability. *Journal of Management Information Systems*, *29*(2), 173-212.

Fisher, C. W., Chengalur-Smith, I., & Ballou, D. P. (2003). The impact of experience and time on the use of data quality information in decision making. *Information Systems Research, 14*(2), 170-188.

Forgionne, G. A., & Kohli, R. (2000). Management support system effectiveness: Further empirical evidence. *Journal of the Association for Information Systems, 1*(1), 1-37.

Fuller, R. M., & Dennis, A. R. (2009). Does fit matter? The impact of task-technology fit and appropriation on team performance in repeated tasks. *Information Systems Research, 20*(1), 2-17.

Galletta, D. F., Henry, R., McCoy, S., & Polak, P. (2004). Web site delays: How tolerant are users? *Journal of the Association for Information Systems, 5*(1), 1-28.

Galletta, D. F., Henry, R. M., Mccoy, S., & Polak, P. (2006). When the wait isn't so bad: The interacting effects of website delay, familiarity, and breadth. *Information Systems Research, 17*(1), 20-37.

Garfield, M. J., Taylor, N. J., Dennis, A. R., & Satzinger, J. W. (2001). Modifying paradigms: Individual differences, creativity techniques, and exposure to ideas in group idea generation. *Information Systems Research, 12*(3), 322-333.

Geissler, G., Zinkhan, G., & Watson, R. T. (2001). Web home page complexity and communication effectiveness. *Journal of the Association for Information Systems, 2*(1), 1-46.

Gemino, A., Parker, D., & Kutzschan, A. O. (2005-6). Investigating coherence and multimedia effects of a technology-mediated collaborative environment. *Journal of Management Information Systems*, *22*(3), 97-121.

George, J. F., Marett, K., & Giordano, G. (2008). Deception: Toward an individualistic view of group support systems. *Journal of the Association for Information Systems, 9*(10/11), 653-676.

Goswami, S., Chan, H. C., & Kim, H. W. (2008). The role of visualization tools in spreadsheet error correction from a cognitive fit perspective. *Journal of the Association for Information Systems, 9*(6), 321-343.

Gregg, D. G., & Walczak, S. (2008). Dressing your online auction business for success: An experiment comparing two eBay businesses. *MIS Quarterly, 32*(3), 653-670.

Hender, J. M., Dean, D. L., Rodgers, T. L., & Nunamaker, J. F. (2002). An examination of the impact of stimuli type and GSS structure on creativity: Brainstorming versus non-brainstorming techniques in a GSS environment. *Journal of Management Information Systems*, *18*(4), 59-85.

Heninger, W. C., Dennis, A. R., & Hilmer, K. M. (2006). Individual cognition and dual-task interference in group support systems. *Information Systems Research, 17*(4), 415-424.

Hess, T. J., Fuller, M. A., & Mathew, J. (2005-6). Involvement and decision-making performance with a decision aid: The influence of social multimedia, gender, and playfulness. *Journal of Management Information Systems*, *22*(3), 15-54.

Hess, T. J., Fuller, M., & Campbell, D. (2009). Designing interfaces with social presence: Using vividness and extraversion to create social recommendation agents. *Journal of the Association for Information Systems, 10*(12), 889-919.

Hilmer, K. M., & Dennis, A. R. (2000-01). Stimulating thinking: Cultivating better decisions with groupware through categorization. *Journal of Management Information Systems*, *17*(3), 93-114.

Hinz, O., Hann, I.-H., & Spann, M. (2011). Price discrimination in e-commerce? An examination of dynamic pricing in name-your-own price markets. *MIS Quarterly*, *35*(1), 81-98.

Hinz, O., & Spann, M. (2008). The impact of information diffusion on bidding behavior in secret reserve price auctions. *Information Systems Research, 19*(3), 351-368.

Ho, S. Y., Bodoff, D., & Tam, K. Y. (2011). Timing of adaptive web personalization and its effects on online consumer behavior. *Information Systems Research*, *22*(3), 660-679.

Hong, W., Thong, J. Y. L., & Tam, K. Y. (2004-5). The effects of information format and shopping task on consumers' online shopping behavior: A cognitive fit perspective. *Journal of Management Information Systems, 21*(3), 149-184.

Hong, W., Thong, J. Y. L., & Tam, K. Y. (2004). Does animation attract online users' attention? The effects of flash on information search performance and perceptions. *Information Systems Research, 15*(1), 60-86.

Huang, W. W., & Wei, K. K. (2000). An empirical investigation of the effects of group support systems (GSS) and task type on group interactions from an influence perspective. *Journal of Management Information Systems*, *17*(2), 181-206.

Hui, K.-L., Teo, H. H., & Lee, S.-Y. T. (2007). The value of privacy assurance: An exploratory field experiment. *MIS Quarterly, 31*(1), 19-33.

Irwin, G. (2002). The role of similarity in the reuse of object-oriented analysis models. *Journal of Management Information Systems*, *19*(2), 219-248.

Jensen, M. L., Lowry, P. B., Burgoon, J. K., & Nunamaker, J. F. (2010). Technology dominance in complex decision making: The case of aided credibility assessment. *Journal of Management Information Systems*, *27*(1), 175-201.

Jensen, M. L., Lowry, P. B., & Jenkins, J. L. (2011). Effects of automated and participative decision support in computer-aided credibility assessment. *Journal of Management Information Systems*, *28*(1), 201-233.

Jiang, Z., & Benbasat, I. (2004-5). Virtual product experience: Effects of visual and functional control of products on perceived diagnosticity and flow in electronic shopping. *Journal of Management Information Systems*, *21*(3), 111-147.

Jiang, Z., & Benbasat, I. (2007). Investigating the influence of the functional mechanisms of online product presentation. *Information Systems Research, 18*(4), 454-470.

Jiang, Z., & Benbasat, I. (2007). The effects of presentation formats and task complexity on online consumers' product understanding. *MIS Quarterly, 31*(3), 475-500.

Jiang, Z., Chan, J., Tan, B. C. Y., & Chua, W. S. (2010). Effects of interactivity on website involvement and purchase intention. *Journal of the Association for Information Systems,* 11(1), 34-59.

Johnson, N. A., & Cooper, R. B. (2009). Power and concession in computer-mediated negotiations: An examination of first offers. *MIS Quarterly, 33*(1), 147-170.

Johnson, R. D., & Marakas, G. M. (2000). The role of behavioral modeling in computer skills acquisition—Toward refinement of the model. *Information Systems Research, 11*(4), 402-417.

Johnston, A. C., & Warkentin, M. (2010). Fear appeals and information security behaviors: An empirical study. *MIS Quarterly, 34*(3), 549-566.

Kahai, S. S., & Cooper, R. B. (2003). Exploring the core concepts of media richness theory: The impact of cue multiplicity and feedback immediacy on decision quality. *Journal of Management Information Systems*, *20*(1), 263-299.

Kamis, A., Koufaris, M., & Stern, T. (2008). Using an attribute-based decision support system for user-customized products online: An experimental investigation. *MIS Quarterly, 32*(1), 159-177.

Kayande, U., De Bruyn, A., Lilien, G. I., Rangaswamy, A., & van Bruggen, G. H. (2009). How incorporating feedback mechanisms in a DSS affects DSS evaluations. *Information Systems Research, 20*(4), 527-546.

Keil, M., Tan, B. C. Y., Wei, K.-K., Saarinen, T., Tuunainen, V., & Wassenaar, A. (2000). A cross-cultural study on escalation of commitment behavior in software projects. *MIS Quarterly, 24*(2), 299-325.

Keith, M., Shao, B., & Steinbart, P. (2009). A behavioral analysis of passphrase design and effectiveness. *Journal of the Association for Information Systems, 10*(2), 63-87.

Khatri, V., Vessey, I., Ramesh, V., Clay, P., & Park, S.-J. (2006). Understanding conceptual schemas: Exploring the role of application and IS domain knowledge. *Information Systems Research, 17*(1), 81-99.

Kim, D., & Benbasat, I. (2006). The effects of trust-assuring arguments on consumer trust in Internet stores: Application of Toulmin's model of argumentation. *Information Systems Research, 17*(3), 286-300.

Kim, D., & Benbasat, I. (2009-10). Trust-assuring arguments in B2C e-commerce: Impact of content, source, and price on trust. *Journal of Management Information Systems*, *26*(3), 175-206.

Kim, J., Hahn, J., & Hahn, H. (2000). How do we understand a system with (so) many diagrams? Cognitive integration processes in diagrammatic reasoning. *Information Systems Research, 11*(3), 284-303.

Kohler, C. F., Breugelmans, E., & Dellaert, B. G. C. (2011). Consumer acceptance of recommendations by interactive decision aids: The joint role of temporal distance and concrete versus abstract communications. *Journal of Management Information Systems*, *27*(4), 231-260.

Komiak, S. Y. X., & Benbasat, I. (2006). The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Quarterly, 30*(4), 941-960.

Kuechler, W. L., & Vaishnavi, V. (2006). So talk to me: The effect of explicit goals on the comprehension of business process narratives. *MIS Quarterly, 30*(4), 961-979.

Kumar, N., & Benbasat, I. (2004). The effect of relationship encoding, task type, and complexity on information representation: An empirical evaluation of 2D and 3D line graphs. *MIS Quarterly, 28*(2), 255-281.

Kumar, N., & Benbasat, I. (2006). The influence of recommendations and consumer reviews on evaluations of websites. *Information Systems Research*, *17*(4), 425-439.

Kwok, R. C.-W., Ma, J., & Vogel, D. R. (2002-3). Effects of group support systems and content facilitation on knowledge acquisition. *Journal of Management Information Systems*, *19*(3), 185-229.

Lang, T.-P., Lai, H.-J., & Ku, Y.-C. (2006-7). Personalized content recommendation and user satisfaction: Theoretical synthesis and empirical findings. *Journal of Management Information Systems, 23*(3), 45-70.

Lankton, N., & Luft, J. (2008). Uncertainty and industry structure effects on managerial intuition about information technology real options. *Journal of Management Information Systems*, *25*(2), 203-240.

Lee, J. S., Keil, M., & Kasi, V. (2012). The effect of an initial budget and schedule goal on software project escalation. *Journal of Management Information Systems*, *29*(1), 53-77.

Lee, Y., Chen, A. N. K., & Ilie, V. (2012). Can online wait be managed? The effect of filler interfaces and presentation modes on perceived waiting time online. *MIS Quarterly*, *36*(2), 365-394.

Lee, Y. E., & Benbasat, I. (2011). The influence of trade-off difficulty caused by preference elicitation methods on user acceptance of recommendation agents across loss and gain conditions. *Information Systems Research, 22*(4), 867-884.

Lerch, F. J., & Harter, D. E. (2001). Cognitive support for real-time dynamic decision making. *Information Systems Research, 12*(1), 63-82.

Lilien, G. L., Rangaswamy, A., Van Bruggen, G. H., & Starke, K. (2004). DSS effectiveness in marketing resource allocation decisions: Reality vs. perception. *Information Systems Research, 15*(3), 216-235.

Lim, K. H., & Benbasat, I. (2000). The effect of multimedia on perceived equivocality and perceived usefulness of information systems. *MIS Quarterly, 24*(3), 449-471.

Lim, K. H., & Benbasat, I. (2002). The influence of multimedia on improving the comprehension of organizational information. *Journal of Management Information Systems*, *19*(1), 99-127.

Lim, K. H., Benbasat, I., & Ward, L. M., (2000). The role of multimedia in changing first impression bias. *Information Systems Research, 11*(2), 115-136.

Lim, K. H., Sia, C. L., Lee, M. K. O., & Benbasat, I. (2006). Do I trust you online, and if so, will I buy? An empirical study of two trust-building strategies. *Journal of Management Information Systems*, *23*(2), 233-266.

Limayem, M., & DeSanctis, G. (2000). Providing decisional guidance for multicriteria decision making in groups. *Information Systems Research, 11*(4), 386-401.

Lowry, P. B., Vance, A., Moody, G., Beckman, B., & Read, A. (2008). Explaining and predicting the impact of branding alliances and web site quality on initial consumer trust of e-commerce web sites. *Journal of Management Information Systems*, *24*(4), 199-224.

Lowry, P. B., Romano, N. C., Jenkins, J. L., & Guthrie, R. W. (2009). The CMC interactivity model: How interactivity enhances communication quality and process satisfaction in lean-media groups. *Journal of Management Information Systems*, *26*(1), 155-195.

Lowry, P. B., Roberts, T. L., Dean, D. L., & Marakas, G. (2009). Toward building self-sustaining groups in PCR-based tasks through implicit coordination: The case of heuristic evaluation. *Journal of the Association for Information Systems, 10*(3), 170-195.

Marcolin, B. L., Compeau, D. R., Munro, M. C., & Huff, S. L. (2000). Assessing user competence: Conceptualization and measurement. *Information Systems Research, 11*(1), 37-60.

Mennecke, B. E., Crossland, M. D., & Killingsworth, B. I. (2000). Is a map more than a picture? The role of SDSS technology, subject characteristics, and problem complexity on map reading and problem solving. *MIS Quarterly, 24*(4), 601-629.

Milton, S. K., Rajapakse, J., & Weber, R. (2012). Ontological clarity, cognitive engagement, and conceptual model quality evaluation: An experimental investigation. *Journal of the Association for Information Systems, 13*(9), 657-694.

Miranda, S. M., & Saunders, C. S. (2003). The social construction of meaning: An alternative perspective on information sharing. *Information Systems Research,* 14(1), 87-106.

Murray, K. B., & Häubl, G. (2011). Freedom of choice, ease of use, and the formation of interface preferences. *MIS Quarterly*, *35*(4), 955-976.

Nah, F. F.-H., Eschenbrenner, B., & DeWester, D. (2011). Enhancing brand equity through flow and telepresence: A comparison of 2D and 3D virtual worlds. *MIS Quarterly*, *35*(3), 731-747.

Nah, F. F.-H., & Benbasat, I. (2004). Knowledge-based support in a group decision making context: An expert-novice comparison. *Journal of the Association for Information Systems*, *5*(3), 125-150.

Nicolaou, A. I., & McKnight, D. H. (2006). Perceived information quality in data exchanges: Effects on risk, trust, and intention to use. *Information Systems Research, 17*(4), 332-351.

Nicolaou, A. I., & McKnight, D. H. (2011). System design features and repeated use of electronic data exchanges. *Journal of Management Information Systems*, *28*(2), 269-304.

Nissen, M. E., & Sengupta, V. (2006). Incorporating software agents into supply chains: Experimental investigation with a procurement task. *MIS Quarterly*, *30*(1), 145-166.

Nunamaker, J. F., Derrick, D. C., Elkins, A. C., Burgoon, J. K., & Patton, M. W. (2011). Embodied conversational agent-based kiosk for automated interviewing. *Journal of Management Information Systems*, *28*(1), 17-48.

Panko, R. R., & Halverson, R. P. (2001). An experiment in collaborative development to reduce spreadsheet errors. *Journal of the Association for Information Systems*, *2*, 1-29.

Parboteeah, D. V., Valacich, J. S., & Wells, J. D. (2009). The influence of website characteristics on a consumer's urge to buy impulsively. *Information Systems Research, 20*(1), 60-78.

Park, C. W., Im, G., & Keil, M. (2008). Overcoming the mum effect in IT project reporting: Impacts of fault responsibility and time urgency. *Journal of the Association for Information Systems, 9*(7), 409-431.

Parsons, J. (2002-3). Effects of local versus global schema diagrams on verification and communication in conceptual modeling. *Journal of Management Information Systems*, *19*(3), 155-183.

Parsons, J. (2011). An experimental study of the effects of representing property precedence on the comprehension of conceptual schemas. *Journal of the Association for Information Systems, 12*(6), 441-462.

Paul, S., Samarah, I. M., Seetharaman, P., & Mykytyn, P. P. (2004-5). An empirical investigation of collaborative conflict management style in group support system-based global virtual teams. *Journal of Management Information Systems, 21*(3), 185-222.

Pennington, R., Wilcox, H. D., & Grover, V. (2003-4). The role of system trust in business-to-consumer transactions. *Journal of Management Information Systems*, *20*(3), 197-226.

Piccoli, G., & Ives, B. (2003). Trust and the unintended effects of behavior control in virtual teams. *MIS Quarterly, 27*(3), 365-395.

Piccoli, G., Ahmad, R., & Ives, B. (2001). Web-based virtual learning environments: A research framework and a preliminary assessment of effectiveness in basic IT skills training. *MIS Quarterly, 25*(4), 401-426.

Poston, R. S., & Speier, C. (2005). Effective use of knowledge management systems: A process model of content ratings and credibility indicators. *MIS Quarterly, 29*(2), 221-244.

Potter, R. E., & Balthazard, P. (2004). The role of individual memory and attention processes during electronic brainstorming. *MIS Quarterly, 28*(4), 621-643.

Price, R., & Shanks, G. (2011). The impact of data quality tags on decision-making outcomes and process. *Journal of the Association for Information Systems, 12*(4), 323-346.

Qiu, L., & Benbasat, I. (2009). Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems. *Journal of Management Information Systems*, *25*(4), 145-181.

Rafaeli, S., & Raban, D. R. (2003). Experimental investigation of the subjective value of information in trading. *Journal of the Association of Information Systems*, *4*, 119-139.

Raghu, T. S., Sinha, R., Vinze, A., & Burton, O. (2009). Willingness to pay in an open source software environment. *Information Systems Research, 20*(2), 218-236.

Reinig, B. A., & Shin, B. (2002). The dynamic effects of group support systems on group meetings. *Journal of Management Information Systems*, *19*(2), 303-325.

Reinig, B. A., Briggs, R. O., & Nunamaker, J. F. (2007). On the measurement of ideation quality. *Journal of Management Information Systems*, *23*(4), 143-161.

Ren, Y., Harper, F. M., Drenner, S., Terveen, L., Kiesler, S., Riedl, J., & Kraut, R. E. (2012). Building member attachment in online communities: Applying theories of group identity and interpersonal bonds. *MIS Quarterly*, *36*(3), 841-864.

Rice, S. C. (2012). Reputation and uncertainty in online markets: An experimental study. *Information Systems Research*, *23*(2), 436-452.

Riedl, R., Hubert, M., & Kenning, P. (2010). Are there neural gender differences in online trust? An fMRI study on the perceived trustworthiness of eBay offers. *MIS Quarterly, 34*(2), 397-428.

Robert, L. P., Dennis, A. R., & Ahuja, M. K. (2008). Social capital and knowledge integration in digitally enabled teams. *Information Systems Research, 19*(3), 314-334.

Robert, L. P., Dennis, A. R., & Hung, Y.-T. C. (2009). Individual swift trust and knowledge-based trust in face-to-face and virtual team members. *Journal of Management Information Systems*, *26*(2), 241-279.

Roberts, T. L., Cheney, P. H., Sweeney, P. D., & Hightower, R. T. (2004-5). The effects of information technology project complexity on group interaction. *Journal of Management Information Systems*, *21*(3), 223-247.

Santanen, E. L., Briggs, R. O., & De Vreede, G.-J. (2004). Causal relationships in creative problem solving: Comparing facilitation interventions for ideation. *Journal of Management Information Systems*, *20*(4), 167-197.

Santhanam, R., Sasidharan, S., & Webster, J. (2008). Using self-regulatory learning to enhance e-learning-based information technology training. *Information Systems Research, 19*(1), 26-47.

Sarker, S., & Valacich, J. S. (2010). An alternative to methodological individualism: A non-reductionist approach to studying technology adoption by groups. *MIS Quarterly, 34*(4), 779-808.

Saunders, C., Rutkowski, A. F., van Genuchten, M., Vogel, D., & Orrego, J. M. (2011). Virtual space and place: Theory and test. *MIS Quarterly*, *35*(4), 1079-1098.

Scheffel, T., Pikovsky, A., Bichler, M., & Guler, K. (2011). An experimental comparison of linear and nonlinear price combinatorial auctions. *Information Systems Research, 22*(2), 346-368.

Shaft, T. M., & Vesseym I. (2006). The role of cognitive fit in the relationship between software comprehension and modification. *MIS Quarterly*, *30*(1), 29-55.

Shanks, G., Tansley, E., Nuredini, J., Tobin, D., & Weber, R. (2008). Representing part-whole relations in conceptual modeling: An empirical evaluation. *MIS Quarterly, 32*(3), 553-573.

Sheng, H., Nah, F. F.-H., & Siau, K. (2008). An experimental study on ubiquitous commerce adoption: Impact of personalization and privacy concerns. *Journal of the Association for Information Systems, 9*(6), 344-376.

Sia, C.-L., Tan, B. C. Y., & Wei, K.-K. (2002). Group polarization and computer-mediated communication: Effects of communication cues, social presence, and anonymity. *Information Systems Research, 13*(1), 70-90.

Sia C. L., Lim, K. H., Leung, K., Lee, M. K. O., Huang, W. W., & Benbasat, I. (2009). Web strategies to promote Internet shopping: Is cultural-customization needed?. *MIS Quarterly, 33*(3), 491-512.

Smith, H. J., Keil, M., & Depledge, G. (2001). Keeping mum as the project goes under: Toward an explanatory model. *Journal of Management Information Systems*, *18*(2), 189-227.

Smith, S. P., Johnston, R. B., Howard, S. (2011). Putting yourself in the picture: An evaluation of virtual model technology as an online shopping tool. *Information Systems Research*, *22*(3), 640-659.

Speier, C., & Morris, M. G., (2003). The influence of query interface design on decision-making performance. *MIS Quarterly, 27*(3), 397-423.

Stewart, K. J. (2006). How hypertext links influence consumer perceptions to build and degrade trust online. *Journal of Management Information Systems*, *23*(1), 183-210.

Storey, V. C., Burton-Jones, A., Sugumaran, V., & Purao, S. (2008). Conquer: A methodology for context-aware query processing on the world wide web. *Information Systems Research, 19*(1), 3-25.

Suh, K.-S., & Lee, Y. E. (2005). The effects of virtual reality on consumer learning: An empirical investigation. *MIS Quarterly, 29*(4), 673-697.

Suh, K-S., Kim, H., & Suh, E. K. (2011). What if your avatar looks like you? Dual-congruity perspectives for avatar use. *MIS Quarterly*, *35*(3), 711-729.

Swaab, R. I., Postmes, T., Neijens, P., Kiers, M. H., & Dumay, A. C. M. (2002). Multiparty negotiation support: The role of visualization’s influence on the development of shared mental models. *Journal of Management Information Systems*, *19*(1), 129-150.

Tam, K. Y., & Ho, S. Y. (2005). Web personalization as a persuasion strategy: An elaboration likelihood model perspective. *Information Systems Research, 16*(3), 271-291.

Tam, K. Y., & Ho, S. Y. (2006). Understanding the impact of web personalization on user information processing and decision outcomes. *MIS Quarterly, 30*(4), 865-890.

Tan, C.-H., Teo, H.-H., & Benbasat, I. (2010). Assessing screening and evaluation decision support systems: A resource-matching approach. *Information Systems Research, 21*(2), 305-326.

Te'eni, D., & Feldman, R. (2001). Performance and satisfaction in adaptive websites: An experiment on searches within a task-adapted website. *Journal of the Association for Information Systems, 2*(1), 1-28.

Tingling, P., & Parent, M. (2002). Mimetic isomorphism and technology evaluation: Does imitation transcend judgment? *Journal of the Association for Information Systems, 3*(1), 113-143.

Tsai, J. Y., Egelman, S., Cranor, L., & Acquisti, A. (2011). The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research, 22*(2), 254-268.

Tung, Y. A., & Mardsen, J. R. (2000). Trading volumes with and without private information: A study using computerized market experiments. *Journal of Management Information Systems*, *17*(1), 31-57.

Turel, O., Yuan, Y., & Connelly, C. E. (2008). In justice we trust: Predicting user acceptance of e-customer services. *Journal of Management Information Systems*, *24*(4), 123-151.

Wang, W., & Benbasat, I. (2007). Recommendation agents for electronic commerce: Effects of explanation facilities on trusting beliefs. *Journal of Management Information Systems*, *23*(4), 217-246.

Wang, W., & Benbasat, I. (2009). Interactive decision aids for consumer decision making in e-commerce: The influence of perceived strategy restrictiveness. *MIS Quarterly, 33*(2), 293-320.

Webster, J., & Ahuja, J. S. (2006). Enhancing the design of web navigation systems: The influence of user disorientation on engagement and performance. *MIS Quarterly*, *30*(3), 661-678.

Wei, K.-K., Tan, B. C. Y., & Heng, C.-S. (2003). Willingness to continue with software projects: Effects of feedback direction and optimism under high and low accountability conditions. *Journal of the Association for Information Systems, 4*(4), 171-194.

Wells, J. D., Valacich, J. S., & Hess, T. J. (2011). What signal are you sending? How website quality influences perceptions of product quality and purchase intentions. *MIS Quarterly*, *35*(2), 373-396.

Wells, J. D., Parboteeah, V., & Valacich, J. S. (2011). Online impulse buying: Understanding the interplay between consumer impulsiveness and website quality. *Journal of the Association for Information Systems, 12*(1), 32-56.

Wolfe, C. J., & Murthy, U. S. (2005-6). Negotiation support systems in budget negotiations: An experimental analysis. *Journal of Management Information Systems*, *22*(3), 351-381.

Xu, H., Teo, H.-H., Tan, B. C. Y., & Agarwal, R. (2009-10). The role of push-pull technology in privacy calculus: The case of location-based services. *Journal of Management Information Systems, 26*(3), 135-173.

Xu, H., Teo, H.-H., Tan, B. C. Y., & Agarwal, R. (2012). Effects of individual self-protection, industry self-protection, and government regulation on privacy concerns: A study of location-based services. *Information Systems Research*, *23*(4), 1342-1363.

Xu, J., Benbasat, I., & Cenfetelli, R. (2011). The effects of service and consumer product knowledge on online customer loyalty. *Journal of the Association for Information Systems, 12*(11), 741-766.

Xu, P., & Ramesh, B. (2008-9). Impact of knowledge support on the performance of software process tailoring. *Journal of Management Information Systems*, *25*(3), 277-314.

Yang, Y., Singhal, S., & Xu, Y. (2012-13). Alternate strategies for a win-win seeking agent in agent-human negotiations. *Journal of Management Information Systems*, *29*(3), 223-255.

Yi, M. Y., & Davis, F. D. (2003). Developing and validating an observational learning model of computer software training and skill acquisition. *Information Systems Research, 14*(2), 146-169.

Yoo, Y., & Alavi, M. (2001). Media and group cohesion: Relative influences on social presence, task participation, and group consensus. *MIS Quarterly, 25*(3), 371-390.

Zhang, D., Lowry, P. B., Zhou, L., & Fu, X. (2007). The impact of individualism-collectivism, Social presence, and group diversity on group decision making under majority influence. *Journal of Management Information Systems*, *23*(4), 53-80.

Zhang, P. (2000). The effects of animation on information seeking performance on the world wide web: Securing attention or interfering with primary tasks?", *Journal of the Association for Information Systems, 1*(1), 1-28.

Zhang, T., Agarwal, R., & Lucas, H. C. (2011). The value of IT-enabled retailer learning: Personalized product recommendations and customer store loyalty in electronic markets. *MIS Quarterly*, *35*(4), 859-881.

Zhu, L., Benbasat, I., & Jiang, Z. (2010). Let's shop online together: An empirical investigation of collaborative online shopping support. *Information Systems Research, 21*(4), 872-891.

ABOUT THE AUTHORS

Nancy Lankton is an Associate Professor in the Division of Accountancy and Legal Environment at Marshall University. She teaches accounting information systems and information systems auditing. Nancy's main research interests include trust's impacts on individual and organizational use of information technology, privacy behaviors related to social networking websites, and decision making in organizations. She has published in journals such as *Contemporary Accounting Research, IEEE Transactions on Engineering Management, Journal of Management Information Systems, and the Journal of the American Medical Informatics Association*. She is a member of the Association for Information Systems, and the Information Systems Auditor and Control Association, and serves as an Associate Editor for Communications of the Association for Information Systems.

Joan Luft is Eli Broad Professor (emerita) of Accounting at Michigan State University. Her research, which focuses on experimental studies in accounting, has appeared in *The Accounting Review*, *Journal of Accounting and Economics*, *Accounting, Organizations and Society*, *Contemporary Accounting Research*, and other journals. She has been Editor or Associate Editor of *The Accounting Review*, *Journal of Management Accounting Research*, and *Accounting Horizons*.

Communications of the Association for Information Systems

ISSN: 1529-3181