Cognitive Bias in the Peer Review Process: Understanding a Source of Friction between Reviewers and Researchers

Chris Street
University of Regina

Kerry W. Ward
University of Nebraska, Omaha

Abstract

In a recent critique of reviewers, Ralph (2016) stated that "Peer review is prejudiced, capricious, inefficient, ineffective and generally unscientific" (p. 274). Our research proposes that one way the peer review process could appear flawed is if those involved had different beliefs about what was important in evaluating research. We found evidence for a cognitive bias where respondents to a survey asking about the importance of particular validity and reliability method practices gave different answers depending on whether they were asked to answer the survey as a researcher or as a reviewer. Because researchers have higher motivation to publish research than reviewers do to review research, we theorize that motivational differences between researchers and reviewers leads to this bias and contributes to the perception that the review process is flawed. We discuss the implications of our findings for improving the peer review process in MIS.

Keywords: Methodology; Peer Review; Case Study.

Introduction

In a recent critique of reviewers, Ralph (2016) stated, "Peer review is prejudiced, capricious, inefficient, ineffective and generally unscientific" (p. 274). He argued that the peer review process is biased, and citing low inter-rater reliability (Cole, Cole, & Simon, 1981; Lock, 1991; Peters & Ceci, 1982; Rothwell & Martin, 2000), concluded that the peer review process is "mostly random" (Ralph, 2016, p. 275), corrupt, and lacks validity.

We find his contention interesting and important. It is as if the peer reviewers being criticized belong to some group that is separate from researchers, when in fact reviewers are drawn from the very same community of academics that conduct the research. If the peer review process is appearing to be random and lacking validity it could be 'reviewer hypocrisy' (i.e., "do-as-I-say, not-as-I-do") or a problem with incentives for authors (if research continues to be published with low quality reporting, why should I expend the effort to improve?[1]). Thinking the issue might somehow lie in variable beliefs about research methodologies, we considered whether the framing that occurs when an academic adopts either the researcher or reviewer role changes the expectations that person has about what is important in applying and assessing research methods. If that is true, it could account for why peer reviews appear random or lacking validity *from the perspective of the researcher*. Our research explores this question and asks, "To what extent do researchers and reviewers

share the same expectations about how to demonstrate validity and reliability in a research study?"

We theorize that researchers and reviewers may give lip service to the idea that validity and reliability are important, but they may not share the same expectations about *how* important it is. These differences may create a disconnect between the two groups. Drawing on the rich literature about problems in the peer review process, we propose that even though the same group of people alternately perform the roles of researcher and reviewer, these roles have different motivations and therefore lead to different conclusions about the application of methodological validity and reliability. If expectations about research practices are influenced by whether the person takes the researcher or the reviewer role, it may help to explain the "mostly random" problems discussed by Ralph (2016).

This paper makes four contributions to the methodology literature in MIS by (1) providing an empirical framework describing a peer review process critique as it exists in MIS and supporting literatures, (2) providing support for a theoretical explanation of why researchers and reviewers have different expectations about research quality, (3) making suggestions for future research on how to improve the peer review process and, (4) proposing an area of focus (reviewer incentives) as an effective first step.

This paper is organized as follows: we first discuss problems that have been identified with the peer review process from an MIS perspective and also provide comparative literature from other disciplines. We then explain what we propose is a new problem in the peer review process (role bias) and present our hypotheses. Next, we describe our method for testing for differences in attitudes between researchers and reviewers regarding the issues of validity and reliability in case studies. We then present our survey results. The paper concludes with a discussion of the findings, the implications for practice, and suggestions for future research.

What Is Wrong with Peer Review?

There are two views in the MIS literature about the seriousness of the problems in the peer review process. First, some argue that there are systemic problems that make peer review inefficient, ineffective, and fatally flawed. It is therefore in need of a complete re-orientation (Ralph, 2016) because incremental changes are not sufficient to solve the problems that are evident (Mora, 2016). A similar view in other management disciplines argues that peer review does not operate in practice the way we

think it does in theory (MacDonald, 2015) and proposes that a renaissance is required (Frey, 2003). Reawakening requires reviewers to shift from being 'coal miners' digging through tailings (Straub, 2009) to becoming 'diamond cutters' who help develop the research into polished gems (Saunders, 2005a). The diamond cutter analogy has taken root outside the MIS literature (Ragins, 2015) and has resulted in recommendations that journal editors should insist on developmental reviews as opposed to simply stressing the limitations of a paper (Van Lange, 1999). The systemic perspective on improving the peer review process proposes that a revolutionary cultural shift from gatekeeping to developmental reviewing is required.

The second view is that the peer review process can be improved without a total makeover. Problems can instead be resolved by taking a components perspective that allows us to understand and mitigate the "Big Four" root causes of inefficiency and ineffectiveness in peer review: unmet obligations, misaligned incentives, reviewer consistency, and cognitive bias. Table 1 summarizes the literature underlying the systemic and component perspectives. The following paragraphs explain each component.

Unmet obligations refer to shirking or otherwise avoiding the professional responsibility that researchers have to not just write papers but to review the papers of others (Petter, 2016). Blind review ensures that once an article is published, the authors will be known but the reviewers will not. This creates an environment where shirking is more likely and is more difficult to detect (Iivari, 2016). That, in turn, means that authors who want a good (i.e., thorough) review are going to be disappointed because few reviewers are willing to do the work that is required to produce a good review (Stafford, 2018; Grainger, 2007). The typical antidote for this problem is to say that if researchers act on the 'Golden Rule' (Petter, 2016), giving what they want to get (Glen, 2014), then fewer 'bad reviews' (Iivari, 2016) will occur. Reminding scholars of their responsibility is a common prescription in this category, but its efficacy is unknown.

Misaligned incentives refer to an imbalance in the implicit and explicit rewards that motivate academics to provide peer reviews. Recognizing and responding to professional responsibility is an example of an implicit incentive, although most editorials and opinions deal with explicit incentives such as peer and institutional rewards and recognition for being a reviewer (Iivari, 2016; Recker, 2016). Stated more bluntly, professors are not rewarded as well for reviewing as they are for publishing (Jennex, 2016), and this leads to less motivation for completing

reviews. This argument was recently made in a panel discussion where one participant described the reward asymmetry between junior scholars and senior scholars and between researchers and reviewers (Chua et al., 2017). Similar arguments have been made in the broader management literature (Northcraft, 2011; Pitsoulis & Schnellenbach, 2012; Kachewar & Sankaye, 2013; Knudson, Morrow, & Thomas, 2014). The most common (and obvious) advice that is heard on the issue of addressing misalignment is to create more meaningful and explicit rewards such as the inclusion of reviewing work in a person's career evaluation (e.g., Jennex, 2018), and paying for reviews (Copiello, 2018). Unfortunately, at least one study found that material incentives do not necessarily lead to high quality, constructive reviews (Squazzoni, Bravo, & Takacs, 2013).

Reviewer consistency refers to the level of agreement between reviewers regarding the contribution and quality of a research article; it is essentially a measure of inter-reviewer reliability. Agreement between reviewers is typically quite low (Lee et al., 2013), with large empirical studies showing that consistency among reviewers is lower than expected (Mulligan, Hall, & Raphael, 2013; Bornmann, 2015). Poor training and inexperience are often cited as the reasons for low reliability (Iivari, 2016) so it is not surprising that most of the suggestions for resolving the problem involve training, practice, and agreeing on reviewing guidelines (Iivari, 2016; Sipior, 2018; Janze, 2017). Mutually understood and consistently applied criteria for understanding research contribution and quality is among the most-cited recommendations for improving reviewer consistency (Epstein, 1995; Clair, 2015; Miller, 2006).

Table 1. What Is Wrong with Peer Review?

Systemic Perspective		Component Perspective	
MIS	*Supporting Literature*	**MIS**	*Supporting Literature*
Structurally Flawed		Unmet Obligations	
• Ralph (2016) • Mora (2016)	• *Frey (2003)* • *MacDonald (2015)*	• Petter (2016) • Iivari (2016) • Stafford (2016)	• *Grainger (2007)* • *McPeek et al. (2009)* • *Glen (2014)*
Not Enough Diamond Cutters		Misaligned Incentives	
• Saunders (2005a) • Saunders (2005b) • Straub (2009)	• *Van Lange (1999)* • *Ragins (2015)* • *Ragins (2018)*	• Iivari (2016) • Recker (2016) • Jennex (2016) • Chua (2017) • Stafford (2018a) • Stafford (2018b)	• *Northcraft (2011)* • *Pitsoulis and Schnellenbach (2012)* • *Squazzoni, Bravo, and Takacs (2013)* • *Kachewar and Sankaye (2013)* • *Knudson, Morrow, and Thomas (2014)* • *Copiello (2018)*
		Reviewer Consistency	
		• Lee et al (2013) • Mulligan Hall, and Raphael (2013) • Bornmann (2015) • Iivari (2016) • Sipior (2018) • Janze (2017)	• *Mitchell, Beach, and Smith (1985)* • *Epstein (1995)* • *Starbuck (2003)* • *Smith (2006)* • *Miller (2006)* • *Cooper (2009)* • *Clair (2015)*
		Cognitive Bias	
		• Weber (1999) • Mandviwalla, Patnayakuni, and Schuff (2009) • Nandhakumar (2010) • Agarwal (2013) • Davison (2015)	• *Mahoney (1977)* • *Beyer, Chanove, and Fox (1995)* • *Justice et al (1998)* • *Tomkins, Zhang, and Heavlin (2017)* • *King et al (2018)*

Cognitive bias refers to prejudicial judgments by reviewers who evaluate research using factors other than research quality and academic contribution (Agarwal, 2013). Four types of bias are commonly discussed in MIS research: (1) ad hominem: well-known researchers receive more favorable reviews (Agarwal, 2013; Davison, 2015), (2) affiliational: authors from well-known institutions receive more favorable reviews than authors from less well-known institutions (Agarwal, 2013), (3) ideological: well-known theories and orthodoxies are favored over less well-known ones (Nandhakumar, 2010; Mandviwalla, Patnayakuni, & Schuff, 2009; Agarwal, 2013), and (4) aesthetic: writing style and prose are used as the basis for evaluation instead of research quality and scientific contribution (Ralph, 2016; Agarwal, 2013). Common advice for mitigating the effects of bias in peer review is to increase transparency and awareness in the publishing process by naming reviewers (Weber, 1999; Mandviwalla, Patnayakuni, & Schuff, 2009) and allotting publication space to diverse perspectives (Nandhakumar, 2010). The conclusion appears to be that bias exists and we must continually work to ensure that quality and contribution of the work are the key criteria when making evaluations. Given the number of cognitive biases shown to influence decision-making, we wonder if there is evidence for more than four types of cognitive bias lurking within the peer review process.

We think there is evidence. Consider first how peer review is both a task ("can you complete this review?") and a role ("Reviewer 2 has asked for the following," "think like a reviewer") (Higgins, 2017; Dupps, 2017). Research shows that when individuals take on different roles, those roles can trigger decision biases that influence both judgment and behavior. In terms of judgment, a lab experiment on professional conduct showed that those asked to be an advisor used different decision criteria to decide what ethically acceptable behavior was compared to those who were asked to be an estimator in the experiment (Sah & Loewenstein, 2014). Another study showed that the decision criteria applied for solving a typical business problem was different depending on whether people were asked to approach the problem as an executive or as a project manager (Dilts & Pence, 2006). In terms of behavior, those playing a character with an impulsive personality in a video game played it differently from those playing other non-impulsive characters (Siebelink, van der Putten, & Kaptein, 2016), while another study showed that people asked to think like a consumer made different decisions than those who were asked to think like a business person. Based on this work we similarly propose that acting as a researcher or as a reviewer can trigger a role bias

that changes the behavior and judgment of an academic; whether an academic is acting as the reviewer or as the researcher in the peer review process can result in a difference in opinion over the importance of criteria for evaluating research quality.

Criteria for Evaluating Research

While there is great diversity in research philosophies, research methodology texts generally describe quality using terms like validity and reliability (Yin, 1994; Kirk, Miller, & Miller, 1986), or credibility and dependability (Lincoln & Guba, 1985). Many perspectives exist about which set of criteria best represents research quality, but for our purposes we chose the positivist perspective and will therefore focus on the evaluative criteria of validity and reliability. [2] One of the most influential references (based on the number of times it is cited) is the methodology text by Yin (1994), which explains how assessing and reporting reliability and validity is essential when conducting rigorous research. Four factors—internal validity, external validity, construct validity, and reliability—communicate quality (Yin, 1994); these factors are widely accepted in the literature (e.g., Bornmann, 2008; Hackett & Chubin, 2003, Pare & Dube, 2003).

Internal validity refers to how well the internal elements of a research project's design, data collection, and data analysis procedures were carried out; the greater the internal validity, the more confident a researcher can be in believing that the results are accurate. *External validity* refers to how well the results from research in one setting also apply to research in other settings; the higher the external validity, the greater the confidence that the results will be applicable in other environments or situations and the greater the likelihood the findings are indicative of a general truth (Brewer, 2000). *Construct validity* refers to how well the researchers are able to measure what they claim to be measuring (Cronbach & Meehl, 1955). *Reliability* refers to whether the results in a particular study can be reproduced by other researchers in other settings (Yin, 1994). Research that has internal, external, and construct validity, as well as reliability is considered by the scientific community to be of 'high quality.' In the scientific world, researchers are responsible for designing and conducting high quality research and reviewers are responsible for evaluating how well the research meets the requirements for quality.

We propose that while both sides can agree that validity and reliability are important, researchers will be biased to place relatively more importance on validity and reliability because, since they receive status and career success through publication, the outcome of the peer review process is vitally

important to them. Conversely, reviewers will be biased to place relatively less importance on validity and reliability because, since they usually remain anonymous and are often acknowledged by name only once a year by the journals that they reviewed for, the peer review process is considerably less important to their career. We therefore suggest the following hypotheses about differences between researchers and reviewers:

Hypothesis 1 (H1): Researchers will indicate that the internal validity of a research study is more important in demonstrating research quality than reviewers will.

Hypothesis 2 (H2): Researchers will indicate that the external validity of a research study is more important in demonstrating research quality than reviewers will.

Hypothesis 3 (H3): Researchers will indicate that the construct validity of a research study is more important in demonstrating research quality than reviewers will.

Hypothesis 4 (H4): Researchers will indicate that the construct validity of a research study is more important in demonstrating research quality than reviewers will.

These hypotheses build on Ralph's (2016) argument that scholars hold different views about the peer review process depending on which side they are on at any given time. The next sections describe the survey methodology and results.

Methodology

We used a mixed methods design (Venkatesh, Brown, & Bala, 2013), which consisted of a quantitative survey to test the hypotheses and qualitative interviews with mid-career and senior MIS scholars to further understand why we got the results we did. We needed a context for the survey questions and decided to use the case study methodology for two reasons: (1) because it is popular in MIS (Kohli & Lymanin (2006) found that 27 percent of IS papers published use this method), and (2) because it is often criticized for not properly describing validity and reliability characteristics (Dube & Pare, 2003). The case research method is criticized in other disciplines for not being sufficiently rigorous (Reige, 2003), not providing adequate details (Cepeda & Martin, 2005), not meeting quality standards (Cho & Trent, 2006), not following expected methods (Kim, Price, & Lee, 2014; Liao, 2015; Piekkari, Welch, & Paavilainen, 2009), and not meeting expectations for high quality research (da Mota Pedrosa, Näslund, & Jasmand, 2012; Kumar, 2015; Marques, Camacho, & Alcantara, 2015). Since

the case method receives so much criticism about quality we thought it would be a useful context in which to examine expectations about validity and quality in the peer review process.

Quantitative Method

Participant Selection

We tested our hypotheses using data from a survey of case study researchers who (1) had recent experience with the case study method, and (2) followed a reasonably consistent research epistemology. "Recent experience" was operationalized by searching the ABI-Proquest online database for management-related case studies that were published in the seven years prior to our survey being launched. "Reasonably consistent epistemology" was operationalized by authors who explicitly referred to their research as being "positivist" or "realist" OR who implicitly followed a positivist method by using either Eisenhardt (1989) or Yin (1994) as a methodology reference. A review of positivist case study research indicated that these two references were so popular as to be considered de facto positivist standards (Dube & Pare, 2003). Email addresses for authors and coauthors of the articles meeting these criteria were available in the articles; this resulted in a pool of 1466 participants.

Research Design

We used a parallel survey design where participants were randomly divided into two groups; we arbitrarily labeled one group as researchers and the other group as reviewers. We created two versions of our survey, one for the researcher group and one for the reviewer group (see Appendix 1). The invitations to participate mentioned researcher or reviewer in the wording of the invitations (i.e. "*Please complete this survey if you have authored or co-authored a case study research project,*" OR *"Please complete this survey if you have served as a reviewer on a case study research manuscript").* The surveys differed only in the framing of the instructions (i.e., *"Please complete this survey from your own perspective as a researcher who has conducted studies using the case method,"* OR *"Please complete this survey from your own perspective as a reviewer who has reviewed manuscripts employing the case method").*

Measures

The survey asked participants about their beliefs regarding the importance of specific practices associated with validity and reliability in case studies. For the purpose of this survey, practices were considered to be the tasks or procedures that were used to represent the different categories of validity

and reliability as presented in Yin (1994). For example, tasks like interviewing multiple informants or having key informants review a draft of the case study represent forms of construct validity. Recommended validity and reliability practices were identified by reviewing methodology sources (Cooper & Schindler, 2008; Kerlinger & Lee, 2000; Lee, 1989; Lee, 1999; Merriam, 1998; Eisenhardt, 1989; Vogt, 1999; Yin, 1994). We followed the descriptions in Yin (1994) to operationalize each category with the corresponding method practices. A total of 20 methodology practices were identified and categorized according to the classifications of internal validity, external validity, construct validity, and reliability. Each practice was treated as a single-item self-report (Dollinger & Malmquist, 2009), which was considered acceptable because single-item measures have been shown to be legitimate measures (Nichols & Webster, 2013) and are used in contemporary management research (Aladwani, 2013; Milewicz & Saxby, 2013; Susskind & Vicarri, 2011).

Pre-Test

We asked eight colleagues with case study or survey method experience to review the list of methodology practices and to make suggestions for improvement. We then pretested the parallel survey design with a sample of 43 scholars who had all published case study research; the goal was to determine if the list of concepts and practices was complete and to solicit comments about potentially missing items. We made minor clarity-related revisions to the wording of some questions but no new items were suggested. The 43 participants (researchers = 19; reviewers = 24) provided further comments and opinions, but no comments were evident indicating that inappropriate, missing, or unclear concepts or practices existed.

Procedure

Survey Monkey (www.surveymonkey.com) was used to host the survey. Separate invitations directing respondents to the appropriate survey were sent to the researcher and reviewer groups. Three reminders were sent at approximately 2 week intervals after the initial invitations were sent to participants. Data collection was completed within two months of the invitations first being sent.

Data Analysis

A series of ANOVAs were used to test for differences between researcher and reviewer responses to the matched questions regarding their beliefs about the importance of specific practices in demonstrating the validity or reliability of a case study.

Qualitative Method

Following the survey analysis we conducted a series of interviews with MIS scholars in January and February, 2018, to help us understand the quantitative results. Thematic analysis (Braun & Clarke, 2006) was selected as the data collection approach because it fit well with the intent of searching for basic explanations regarding a topic and because the approach was consistent with a positivist perspective (Payne & Payne, 2004). An interview protocol was developed in December, 2017, and tested with colleagues of the first author; this resulted in minor changes to the wording of some questions to improve clarity.

Fourteen interview participants were selected from a convenience sample that was created using a snowballing technique. It began with an interview with a colleague of the second author who was active in publishing qualitative research in the MIS literature. Triangulation of multiple viewpoints was sought by interviewing both mid-career (n = 7) and senior MIS academics (n = 7), and both male (n = 9) and female (n = 6) academics. One recommended interviewee was from the healthcare research community and the remaining participants were from the MIS discipline. All interviewees had *published* and *reviewed* at least one case study.

The interview protocol contained three questions: (1) what are your responsibilities when authoring a case study, (2) what are your responsibilities when reviewing a case study, and (3) why might there be differences between respondents when they answered the survey questions from the perspective of a researcher versus a reviewer. When participants were asked the third question, they were presented with the survey results in a table similar to Table 4 below. Interviews were recorded, note transcripts were provided to interviewees for member checks, and the transcripts were open-coded independently by both authors. Each author independently defined and labeled the themes in their analysis and then compared and discussed differences until agreement was reached. An independent evaluator, experienced in qualitative methods and blind to the research question and quantitative results, was recruited to compare the final coding with the original transcripts to check against coding bias on the part of the authors. The evaluator asked minor clarifying questions (e.g., *does this quote sufficiently signify this theme?*) but reported that no obvious bias was apparent and that the codes fairly represented the interview data.

Results

Quantitative Results

Email invitations were sent to 733 researchers and 733 reviewers. A total of 83 researcher invitations and 76 reviewer invitations "bounced back," indicating that the contact information was no longer valid and that the potential participants did not receive the invitations. Eleven respondents from the researcher pool and six from the reviewer pool responded that they would not participate for reasons related to definitional issues (e.g., *"a case study is qualitative, not positivist"*). This process left us with a researcher group of 639 participants and a reviewer group of 651 participants. We received 80 researcher responses and 74 reviewer responses; this resulted in a combined response rate of 11.9 percent overall (see Table 2).

We reviewed descriptive statistics for the demographic data to test for possible response bias between the two groups prior to hypothesis testing. There were no differences between the two groups in terms of area of most expertise (see Table 3) or in authoring and reviewing experience (see Table 4). Our conclusion was that the two randomly assigned samples were not significantly different in terms of experience with the case method.

We tested for non-response bias by ordering the results from first response to last and comparing responses to each question for differences. No significant differences were found, supporting the conclusion of no response bias. We also tested the ANOVA results for false positive bias due to multiple comparisons using the Benjamini and Hochberg (1995) method. Following McDonald (2014) who suggests using a Q value between 0.10 and 0.20 for this method in the early stages of a project when there is little additional data to refer to, we used a Q value of 0.20. Our results showed that 13 of the 20 comparisons were significantly different, supporting the conclusion that our ANOVA results are not significantly influenced by false positives. Finally, we used Q-Q plots and Levene's test to confirm that the data met the normality and homoscedasticity requirements for ANOVA and concluded that the data were sufficient for this analysis.

Table 5 shows the results from comparing researcher and reviewer beliefs toward common practices in reporting validity and reliability; it also shows which group placed higher importance on an individual practice. Overall, twelve of the 20 concepts (60 percent) showed significant differences in the direction hypothesized. All but one of the significant differences ("establishing a chain of evidence") pointed in the direction of the researcher. In order to test the individual hypotheses, we ordered the practices in Table 4 according to Yin's (1994) categories of internal validity (data analysis), external validity (research design), construct validity (data collection), and reliability (data collection). Internal validity and external validity showed many differences in attitude, while construct validity and reliability showed relatively fewer differences.

Table 2. Number of Usable Surveys: 154

	Researcher	*rate*	Reviewer	*Rate*	Overall	*Overall rate*
Responses	80	*12.5%*	74	*11.4%*	154	
Sample Size	639		651		1290	***11.9%***

Table 3. Methodology Expertise

Which research method do you have the <u>most</u> experience with?			
	'Researcher'	'Reviewer'	Total
Experiment	2	3	5
Case Study	47	48	95
Survey	13	13	26
Other	18	8	26
Total	80	72	152
n=152. No sig. differences between researchers and reviewers (chi-square independence test) Pearson chi-square (3.646, df=3)			

Table 4. Authoring and Reviewing Experience

In how many case study research articles have you acted as a:								
	Researcher			Reviewer				
	N	M	SD	N	M	SD	F	*P*
Authored / co-authored?	80	5.39	3.99	73	6.18	4.91	1.201	0.275
Reviewed	79	8.92	14.07	72	10.03	14.91	0.219	0.641
n=152. No sig. differences between researchers and reviewers (one-way ANOVA)								

Table 5. t-test Results Comparing Attitudes Toward Common Practices Regarding Validity and Reliability Between Researchers and Reviewers (Mean Importance Scores and Standard Deviations)

		'Researchers'			'Reviewers'			Who indicated it was of higher importance?	*p*	Cohen's d *effect size*
		n	M	SD	n	M	SD			
Internal Validity *(analysis, Yin, 1994)*	Development of rival hypothesis	**75**	**7.19**	**1.46**	**61**	**6.54**	**1.93**	**Researcher**	**0.008*****	**0.4214**
	Using explanation-building	**53**	**7.58**	**1.15**	**48**	**7.10**	**1.69**	**Researcher**	**0.048****	**0.3354**
	Address rival explanations	**77**	**7.84**	**1.29**	**71**	**7.45**	**1.57**	**Researcher**	**0.027****	**0.3182**
	Using triangulation of data	**79**	**8.19**	**1.19**	**66**	**7.73**	**1.92**	**Researcher**	**0.039****	**0.2960**
	Coding of data	**76**	**7.79**	**1.26**	**73**	**7.30**	**2.14**	**Researcher**	**0.045****	**0.2796**
	Using pattern matching in analyzing the data	70	7.27	1.56	63	7.19	1.80		0.391	0.0483
External Validity *(design, Yin, 1994)*	Having multiple cases	**79**	**6.54**	**1.63**	**72**	**5.88**	**2.03**	**Researcher**	**0.013****	**0.3656**
	Using theory to develop site selection criteria	**77**	**7.29**	**1.52**	**67**	**6.79**	**2.16**	**Researcher**	**0.056***	**0.2677**
	Use replication logic in multiple case studies	**63**	**7.11**	**1.52**	**62**	**6.69**	**2.10**	**Researcher**	**0.093***	**0.2382**
	Having a research question	80	8.21	1.38	73	8.14	1.68		0.381	0.0493
	Having propositions or hypotheses	76	6.18	2.15	71	6.04	2.66		0.361	0.0589
Construct Validity *(collection, Yin, 1994)*	Having multiple sources of evidence	**77**	**8.32**	**1.03**	**72**	**7.90**	**1.24**	**Researcher**	**0.012****	**0.3719**
	Key informants review draft case study report	**78**	**7.51**	**1.32**	**73**	**6.97**	**1.97**	**Researcher**	**0.024****	**0.3243**
	Establish chain of evidence	76	7.79	1.34	68	7.84	1.51		0.419	-0.0342
	Interviewing multiple informants	80	8.01	1.51	71	8.00	1.41		0.479	0.0085
Reliability *(collection, Yin, 1994)*	Develop case study database	**75**	**7.09**	**1.88**	**69**	**6.30**	**2.42**	**Researcher**	**0.011****	**0.3864**
	Determining inter-coder reliability	**72**	**7.36**	**1.43**	**68**	**6.99**	**2.09**	**Researcher**	**0.098***	**0.2201**
	Use case study protocol	79	7.59	1.46	70	7.47	1.89		0.327	0.0736
	Using an interview guide	78	7.86	1.42	70	7.79	1.69		0.388	0.0471
	Recording and transcribing interviews	78	7.86	1.38	72	7.57	1.66		0.123	0.1902

*Note: n = 154. ***p<0.01; **p<0.05; *p<0.1*

H1 predicted that researchers would attach significantly more importance to internal validity than reviewers would. Five of the six practices (83.3 percent) show significant differences in the direction expected, suggesting strong support for H1. With regard to H2, three of the five practices (60 percent) show significant differences in the expected direction, suggesting moderate support for H2. With regard to H3, two of the four practices (50 percent) show significant differences in the expected direction, suggesting modest support for H3. With regard to H4, two of the five practices (40 percent) show significant differences in the expected direction, suggesting only weak support for H4. With the exception of *"Establish chain of evidence,"* 19 of the 20 scores for the researcher role are higher than for the reviewer role, suggesting that the proportion of responses (95 percent versus 5 percent) is consistent with hypotheses one to four.

Overall, our analysis demonstrates at least three results. First, significant differences exist in the opinions of respondents depending on whether a problem is framed as "think like a researcher" versus "think like a reviewer." Second, the significant differences are *all* in the expected direction; in *all* cases it is researchers who see a given methodological practice as being more important in demonstrating validity and reliability. Third, some practices are considered as more important than others in signaling quality and consistency. Internal validity is the most important concept to demonstrate, followed by external validity and then construct validity. Reliability places a distant fourth in importance. The qualitative data in the next section help explain why these differences might exist.

Qualitative Results

Interviews were conducted with 14 MIS scholars to help us understand the differences between the researcher and reviewer results. Interview participants thought researchers had four primary responsibilities and that reviewers had three. They also thought that there were at least three explanations for why researchers would have stronger opinions about validity and reliability in case studies than reviewers would. We included a factor if it appeared multiple times and if it was conceptually distinct from the other factors. Initially there was a

single theme labeled "*Getting Published*" under researcher responsibilities; however, after the independent coding evaluator suggested that there were two distinct concepts in this category, "*surviving the first review*" and also subsequently "*to be a published author,*" we split this theme into two. Distinctions between these two factors, and the rest of the categories, are explained below. These results are summarized in Table 6 and elaborated with illustrative quotes in Tables 7-9.

The most common researcher responsibility was in contributing high quality work to the discipline. "Contribution" referred to adding something new to ongoing scholarship while "High quality" typically meant accurate, representative work. The next most common responsibility was "Communication" which typically meant the quality and clarity of scientific writing. The third theme that "Surviving the first review" requires submitting high quality work and not alienating reviewers, while "Being a published author" requires the motivation to publish research and to be flexible during the review process.

The relative frequency of reviewer responsibilities were essentially the same for the three themes shown in Table 8. The most common was in ensuring "quality control," which means evaluating whether the accuracy and reliability/believability of the work meets the standards of the journal. The second most commonly described role was to "help develop the contribution," which refers to a shared expectation that reviewers should be constructive in how they evaluate others, help authors to "*pull that nugget out,*" and help authors to communicate the unique contribution in their work. The third role, "Be a gatekeeper," was mentioned slightly less than the first two, but it was recognized as distinct from quality control. One interviewee, a senior scholar and well-respected senior editor of a journal, explained it this way: "*A gatekeeper looks at a paper and looks for a problem and focuses on that instead of focusing on if there's a gem there . . . quality control is more about development . . . instead of trying to shut down papers, that's the gatekeeper approach . . . good stories will sell weak data, development reviewers will lean that way, gatekeepers won't.*" Provocative theories supported by relatively weak data may be considered by developmental reviewers but are likely to be rejected immediately by gatekeepers.

Table 6. Role Differences and Possible Reasons Why

Researcher Responsibilities	n	Reviewer Responsibilities	N	Reasons why opinions on method are different	n
Contributing high quality work	22	Quality control	16	Reviewers emphasize the contribution	12
Communication	15	Help develop the contribution	15	Authors fear an early reject	12
Surviving first review	5	Be a gatekeeper	13	Misinterpretation of expected standards	7
To be a published author	4				

Table 7. Researcher Responsibilities

Perceived Expectation	Supporting Evidence
Contributing high quality work	"The role of an author is to make a contribution and to demonstrate the case method rigor. A reviewer wants the same thing, to see a contribution and to be convinced the contribution is based on rigorous methods." "The methodology is a necessary but not sufficient condition. As an author, it is important to have a good method, a very rigorous method, but it's not enough. The author has to go above and beyond for a contribution." Sometimes half-baked stuff gets submitted and we shouldn't do that; researchers have a responsibility to complete their work well. "A lot of research I have done has the objective of proposing theory and for papers that I have had accepted ... my role as an author was to demonstrate to reviewers that I did make a contribution."
Communication	"It's a persuasive exercise; the author must show the reviewer that what they did and how they presented it was adequate. Authors need to ask themselves to more clearly explain what they've done. Emphasis is on making that believable, and more so than for a quantitative researcher." "The role of an author is to find an interesting, valid question that is relevant to the field. Proper design and reasonable, and to write it up well, be sure the writing is sufficiently clearly communicated. There is pressure on authors to communicate well." "To provide an unbiased account of a story. To do that you need to include as many people as you can to get all the stakeholder views." "I tend to think a lot about the writing; if you don't communicate it clearly then it won't get in."
Surviving first review	"As an author I need to publish, and to do that I need to establish the quality in the first review and I can develop the contribution over time." "My strategy is to not tick off too many people (reviewers); can I put something into the paper that only ticks off one reviewer, just enough to get in the first review, then I can focus in on working with the remaining reviewers to get something published." "As a researcher, if the method is insufficient that's a fatal flaw and you can't normally go back and fix that and you're done, there won't be a revision." "As the author, your role is to submit and pray! You hope you get a generous reviewer that wants to work with you."
To be a published author	"To get published is an end itself." "What's your motivation? Publish or perish! It's required for the academic position that I'm in." "As an author, it's always to get a line in the CV, promoted, a bonus at work." "...at times when I have something I think is a contribution but the editor disagrees, I'll change the main message if I have to in order to get published."

Table 8. Reviewer Responsibilities

Perceived Expectation	Supporting Evidence
Quality control	"It's not different depending on the type of review; as a reviewer the first role is quality control and shouldn't be a gatekeeper." "I see it more as quality control so my first job is to determine if there is a fatal flaw that makes it unacceptable, either empirical or conceptual." "Is there enough quality to be able to develop a theoretical contribution?" "Role of the reviewer is quality control; is the study framed well, motivation, conclusions reasonable, is there a contribution."
Help develop the contribution	"The main role is to determine if something is a contribution, and evaluators can sometimes contribute to the contribution through suggestions and observations in a significant way even if it's a minor comment." "As a reviewer I can take a longer view and if the paper is sufficient then I can help develop the paper." "What's the role of the reviewer? I need to be convinced there is a core component, is there a nugget that can be brought out and can it be brought out? Can my review help to pull that nugget out? Can I draw out the theoretical contribution?" "The role of a reviewer is to help authors make a contribution."
Be a gatekeeper	"Psychology is like us in MIS that they're problem focused, looking for reasons to reject. Maybe there's something about our field and we're always thinking about our identity and we don't take risks, we want papers to be perfect and we're afraid of letting something through that's not perfect." "My role is the policeman role; it's a role that says "is this accurate, is this a contribution, I play the devil's advocate role, did they actually do this, are these the actual results they got, have the results been manipulated, what are the other explanations." "As a reviewer, quality control means you look at the paper as a critic … you're looking for problems …. Reviewers are predisposed to find something wrong so you can reject it quickly." "There is a role of safeguarding and making sure you don't get manuscripts that have no value."

Table 9. Reasons Why Opinions on Method Are Different

Explanation	Supporting Evidence
Reviewers emphasize the contribution	*"As a reviewer I want to cut to the chase, I want to get to the discussion, what they contributed to the topic and to theory. From a review perspective if the front end is smooth enough, it's convinced me [about the quality]."* *"I'll even recommend in later reviews to cut back the method section for length reasons. Once I'm convinced the authors have quality methods I'm confident the paper can scale that back. We're [MIS] always focused on theory, need all the space to focus on that."* *"Maybe the reviewers are looking at the whole paper; the theoretical contribution is the most important and is more important than the method. So maybe as a reviewer I am putting more focus on that."*
Authors fear an early reject	*"As an author I need to punch the review team in the face with the paper; if I don't nail everything in the first review I won't get any further, and if I get a quirky reviewer it makes it even more iffy, so I need to make sure everything is over the top, on the first go."* *"As an author, I am so concerned about being careful, dotting all my I's so the reviewers don't get me. I am trying to think through, how are they going to criticize me."* *"We tried to not just look at Yin but look at all of them, and try to get an overall picture of what is the best thing, because you never know who you will get as a reviewer. Are you going to get a Yin lover, an Eisenhardt lover, a Miles and Huberman lover, so we tried to be really comprehensive in making sure we did it correctly."* *"Need to hit that first round of review hard to keep the reviewers on the hook if you haven't convinced them of the quality."*
Misinterpretation of expected standards	*"The reviewers don't always see all the steps but the authors do and they have to decide how to communicate what it is that the reviewers will need to see. Reviewers are looking for warning signs and authors are trying to avoid warning signs."* *"As the reviewer, positivist is one genre with another view being an interpretivist or constructivist; when they collide they'll hold each other to their own standards and that causes tension of course."* *"Without a protocol how would reviewers know what you did and how important it was?"* *"As an author you must make it as tight as possible because you don't know who you're going to get or what they're going to evaluate on. Authors are trying to incorporate field standards while reviewers can apply personal standards. Reviewers are even applying new and different standards in order to stand out."*

The last question we asked interviewees was whether they had theories about *why* differences might be evident in the researcher and reviewer roles. All interviewees were able to propose their own theories quickly after reading through the results from Table 4. Three explanations clearly emerged in the transcripts: (1) reviewers may overemphasize the contribution relative to researchers ("*As a reviewer I want to cut to the chase, I want to get to the discussion*"); (2) researchers fear an early reject ("*If I don't nail everything in the first review I won't get any further*"); and (3) misinterpretation of expected standards may be occurring between researchers and reviewers ("*As an author you must make it as tight as possible because you don't know who you're going to get or what they're going to evaluate on*").

Discussion

Our motivation for this research came from critical reviews of the peer review process in MIS research. Assessments of the review process show a great range of opinion, with some viewing it as structurally capricious (Ralph, 2016), and others seeing it as a problem of consistency (Iivari, 2016) or cognitive bias (Agarwal, 2013). In our research, we tested whether attitudinal differences exist between researchers and reviewers over the relative importance of validity and reliability techniques in the realm of case studies. We found that researchers and reviewers do, in fact, have different attitudes about the importance of validity and reliability. Differences were most apparent in two areas (internal validity and external validity), with five of six internal validity practices and three of five external validity practices being viewed as more important to researchers. In all cases except one ("establishing a chain of evidence") it was the researcher framing condition that reported higher importance.

We also tried to determine *why* such attitudinal differences might exist. We hypothesized that while both researchers and reviewers would agree that methods are important, researchers would say that techniques for establishing validity and reliability are more important than reviewers would. We thought we would see these differences because even though researchers and reviewers come from the same pool of scholars, framing oneself as either a researcher or as a reviewer would likely influence expectations related to validity and reliability.

Because the only difference in the study was how the question was framed in the survey, we propose classifying the results as a type of cognitive bias that affects the peer review process. In terms of choosing a name for this particular bias we feel that 'reviewer bias' or 'researcher bias' does not really fit because either option could suggest that one perspective or the other is responsible for the outcome. Similarly, using the term 'framing bias' could be misleading because there are many competing but unrelated frames such as 'positivist,' 'interpretivist,' 'feminist,' that could be involved in the peer review process. For our *quantitative* results, we chose the term 'role bias,' which we defined as a cognitive bias that arises from framing effects in the peer review process. Depending on whether individuals are acting as a researcher or a reviewer, role bias influences scholars to think differently about the relative importance of validity and reliability.

The *qualitative* analysis provided insights about why this bias might exist. When we asked interview participants about what researchers and reviewers are responsible for, the replies suggested that both researchers and reviewers are working toward the same goal, namely a consistent stream of high quality research publications. In some cases this connection was explicit, such as when one individual commented, "*The role of an author is to make a contribution and to demonstrate the case method rigor. A reviewer wants the same thing, to see a contribution and to be convinced the contribution is based on rigorous methods.*" Others agreed saying, "*The role of a reviewer is to help authors make a contribution,*" and "*evaluators can sometimes contribute … through suggestions and observations in a significant way even if it's a minor comment*" showing that both researchers and reviewers work toward a common goal. Researchers are responsible for combining rigorous research with effective communication to get through a review: "*find(ing) an interesting valid question that is relevant to the field.*" Reviewers are responsible for finding and polishing the contribution while evaluating the suitability of that work for publication: "*as a reviewer the first role is quality control.*"

While both researchers and reviewers work toward a common goal, the results also suggest that the motivation is different. Publishing is seen as a professional requirement by researchers, as evidenced by the following comments:

"*As an author I need to publish.*"

"*… your role is to submit and pray! You hope you get a generous reviewer that wants to work with you.*"

"*To get published is an end itself.*"

"*What's your motivation? Publish or perish!*"

"*As an author, it's always to get a line in the CV, promoted, a bonus at work.*"

Respondents indicated no similar responsibility on the reviewer side, only the tasks of helping to develop the contribution, gatekeeping, and quality control.

The difference between researchers and reviewers becomes more significant when looking at reasons why they might not fully agree about the relative importance of specific method practices. On the one hand, comments like, "*Maybe the reviewers are looking at the whole paper; the theoretical contribution is the most important and is more important than the method*" and "*As a reviewer I want to cut to the chase, I want to get to the discussion, what they contributed to the topic and to theory,*" suggest that relative to researchers, reviewers may downplay the method to emphasize the contribution. On the other hand, comments like, "*If I don't nail everything in the first review I won't get any further*" and "*As an author, I am so concerned about being careful, dotting all my i's so the reviewers don't get me,*" suggest that, relative to reviewers, researchers may overemphasize method to avoid early rejection in the review peer process.

Researchers experience something that reviewers do not: the risk of not being published. Concerns about "*if I get a quirky reviewer*" and "*dotting all my i's*" point out the substantial uncertainty of the peer review process. If a researcher survives the first round, the rest of the review process is a careful exchange between the researcher and the reviewers. As one interviewee reflected, "*[if] the editor disagrees, I'll change the main message if I have to in order to get published.*" This mirrors earlier criticism that researchers give up their voice too readily during the review process (Frey, 2003). Given the critical importance to researchers of surviving the peer review process, it's not surprising that extra emphasis is placed on an area that they have some measure of control over. Because researchers have more 'skin in the game' than reviewers in terms of the professional implications of publishing, we propose that role bias is the result of motivational differences. Researchers tend to place significantly more importance on validity and reliability related to research design and data analysis methods, while reviewers do not attribute the same level of importance to these areas. Researchers are more highly motivated than reviewers because they want to have their work successfully pass through peer review and be published.

Role bias exists as a complementary factor beyond the four cognitive biases that are discussed in the

peer review critique literature. Ad hominem (*well-known researchers are favored*) and affiliational (*well-known institutions are favored*) biases are grounded in the reputational effects of an individual or their workplace (Agarwal, 2013; Davison, 2015). Ideological (*well-known theories are favored*) and aesthetic (*engaging writing styles are favored*) biases are grounded in the stylistic effects of popularity and comfort (Ralph, 2016; Agarwal, 2013). Adding to reputational and stylistic effects, we propose that role bias is a motivational effect driven by differences between researchers and reviewers.

Implications

Our research provides a cognition-based perspective on the debate about the quality of the peer review process and suggests three implications. First, in terms of the systemic argument that there are not enough 'diamond cutters' (Saunders, 2005a) in the reviewer ranks, our qualitative results provide support for the suggestion that this constructive approach is gaining popularity. Comments that describe the reviewer role as helping to develop the contribution could also have been labeled as diamond cutting. Some comments like "*Can my review help to pull that nugget out?*" and "*Can I draw out the theoretical contribution?*" implies that reviewers are in some cases considering themselves as co-authors (Saunders, 2005b).

Second, in terms of the component perspective, concern about reviewer consistency appeared in our results as well. One possible explanation for inconsistency is misinterpretation of quality standards. Whether these occur through differences in research perspectives ("*As the reviewer, positivist is one genre with another view being an interpretivist … when they collide they'll hold each other to their own standards and that causes tension…*") or idiosyncratic reviewing ("*reviewers can apply personal standards … even applying new and different standards in order to stand out*"), our results support the inconsistency criticism of the peer review process. Building from our results, we suspect that part of the problem may be differences in what researchers and reviewers believe are the appropriate methods and how they should be evaluated. Recommendations to provide systematic feedback to reviewers (Iivari, 2016) could be supplemented with developmental information about what parts of a review were particularly well-applied (or mis-applied). A better applied understanding of methods could reduce the reliance on personal and idiosyncratic opinions in the peer review process and perhaps lessen cognitive bias effects.

Third, our results provide a possible re-interpretation of two other component problems—unmet

expectations and misaligned incentives—that plague the peer review process. Petter (2016) characterizes the *unmet expectations* problem by describing reviewing as a professional obligation for scholars. Failing to fulfill this obligation harms the research community because scholars who submit articles for review but do not provide reviews themselves are essentially free-riding (Stafford, 2016). Jennex (2016) describes the *misaligned incentives* problem as one where researchers are rewarded for publishing, but reviewers are not similarly rewarded for providing reviews; this reduces the motivation to act as a reviewer. Our results support both these perspectives. Building on the suggestion that misaligned incentives influence motivation, our results suggest that motivational differences may go on to influence how reviewers think about the assessment process and where they will focus their attention. Whether through unmet obligations, misaligned incentives, or cognitive bias, the motivation to provide high quality reviews becomes a central factor in explaining problems in the peer review process. Without resorting to 'silver bullet' prescriptions, we suspect that increasing the incentives for reviewing, and thereby increasing the motivation to review, would significantly improve the process.

If there is a motivational imbalance (as our research suggests), rebalancing incentives should reduce the differences between authors and reviewers. Rebalancing incentives might include such things as giving credit for reviewing activities when promotion and tenure decisions are made (Jennex, 2016), and making the reviewer's work more objectively visible. The latter suggestion has already been implemented in some situations. For example, the *Journal of Behavioral and Experimental Economics* has experimented with posting a ranking of reviewer response times (Matthews, 2016), and the website Publon.com provides a centralized public forum to promote a scholar's reviewing activity. Promoting reviewing could also be accomplished by extending a Google Scholar-style citation count system to show how many times a published paper that a reviewer was involved with has been cited.[3]

Limitations and Suggestions for Future Research

Our research examines the differences between researchers and reviewers during the review process, but it is not without limitations. First, our survey instrument focused only on respondents' views of how important each practice is to quality research. We did not have information about *why* respondents had those opinions or *why* the collection of practices themselves were seen as important. Using multiple items for each practice would have likely resulted in

better construct validity than using single-item indicators. We identified potential respondents based on an indication that their published case study took a positivist approach, but the lack of consensus about the conceptualization of a positivist case study may have contributed to our relatively low response rate (11.9 percent). We noted earlier that eleven respondents declined our survey due to issues related to the conceptualization of positivist case studies, and there may have been others who did not respond for the same reason.

Second, we recognize that no one practice stands on its own but is instead part of a greater whole. We need to know more about how scholars see collections of practices complementing one another and whether synergy in methods in some way influences attitudes about quality. While our interviews provided insights into the differences between researchers and reviewers (and why these differences might exist), we were not able to identify why specific items of validity and reliability exhibited significant differences with the research design we used. More research is therefore needed to explore the reason why these differences exist and what can be done about them.

Third, further research is needed to better understand how the peer review process can be improved. We believe that the most significant factor in improving the process is increasing the motivation to review. Research on how best to increase motivation for reviewers is necessary. For example, some research has shown that material incentives do not improve reviews (Squazzoni, Bravo, & Takacs, 2013), but other research suggests that recognition-based systems for rewarding reviewers are valuable (Kachewar, Ghanshyam, & Sankaye, 2013; Knudson, Morrow, & Thomas, 2014). Journal editors have

many opportunities to experiment with ways to balance incentives between researchers and reviewers. We have just scratched the surface of what could be an interesting area of research.

Conclusion

The conclusion from this research is that role bias driven by motivational differences between researchers and reviewers contributes to differences in the importance that researchers and reviewers attach to validity and reliability practices. This bias adds to the critique of the peer review process. This research helps us to understand recent criticism of the quality of the peer review process in MIS research and suggests a path forward in resolving the shortcomings. Considering the importance of publishing to all those involved, renewed attention and an improved process should be well appreciated.

Notes

[1] We thank Brent Gallupe for these insights.

[2] We recognize that many other perspective choices exist and would be equally applicable as a foundation for this research. The authors choose this perspective only because it was the most familiar to them.

[3] We would be interested to see an 'r-index' developed, similar in concept to the h-index, that represents reviewing activity weighted by an impact factor such as the citation count of the reviewed paper where each time a paper that someone reviewed is cited, their r-index increases. While there are many ways this measure can be manipulated (just like citation counts themselves), it would be interesting to see how well an r-index would correlate with the h-index. We suspect that existence of an r-index would motivate people to increase their reviewing activity because they would then have something external to show their Dean.

References

Agarwal, R. (2013). On the ethics of peer review. *Information Systems Research*, *24*(1), 1-2.

Aladwani, A. M. (2013). A contingency model of citizens' attitudes toward e–government use. *Electronic Government, an International Journal, 10*(1), 68-81.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, Series B (Methodological), 289-300.

Beyer, J. M., Chanove, R. G., & Fox, W. B. (1995). The review process and the fates of manuscripts submitted to AMJ. *Academy of Management Journal*, *38*(5), 1219-1260.

Bornmann, L. (2015). Interrater reliability and convergent validity of F 1000 Prime peer review. *Journal of the Association for Information Science and Technology, 66*(12), 2415-2426.

Bornmann, L. (2008). Scientific peer review: An analysis of the peer review process from the perspective of sociology of science theories. *Human Architecture: Journal of the Sociology of Self-Knowledge, 6*(2), 23-38.

Braun, V., & Clarke, V. (2008). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101.

Brewer, M. (2000). Research design and issues of validity. In H. Reis & C. Judd (eds.), *Handbook of research methods in social and personality psychology*. Cambridge: Cambridge University

Press.

Cho J., & Trent, A. (2006). Validity in qualitative research revisited. *Qualitative Research*, *6*, 319-327.

Chua, C. E.H., Thatcher, J., Niederman, F., Chan, Y.E., & Davidson, E.J. (2018). ICIS Panel Report: Break your shackles! Emancipating information systems from the tyranny of peer review. *Communications of the Association for Information Systems,* 43. https://doi.org/10.17705/1CAIS.04325

Clair, J. A. (2015). Toward a bill of rights for manuscript submitters. *Academy of Management Learning & Education, 14*(1), 111-131.

Cole, S., Cole, J. R., & Simon, G. A. (1981). Chance and consensus in peer review. *Science*, *20*(214), 881-886.

Cooper, M. L. (2009). Problems, pitfalls, and promise in the peer-review process: Commentary on Trafimow & Rice. *Perspectives on Psychological Science, 4*(1), 84-90.

Cooper, D.R., & Schindler, P. S. (2008). *Business research methods*, 10th ed. Boston, MA: McGraw Hill.

Copiello, S. (2018). On the money value of peer review. *Scientometrics, 115*(1), 613-620.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302

da Mota Pedrosa, A., Näslund, D., & Jasmand, C. (2012). Logistics case study based research: Towards higher quality. *International Journal of Physical Distribution & Logistics Management, 42*(3), 275 – 295.

Davison, R. M. (2015). Editorial: The art of constructive reviewing. *Information Systems Journal, 25*(5), 429-432.

Dilts, D. M., & Pence, K. R. (2006). Impact of role in the decision to fail: An exploratory study of terminated projects. *Journal of Operations Management, 24*(4), 378-396.

Dollinger, S. J., & Malmquist, D. (2009). Reliability and validity of single-item self-reports: With special relevance to college students' alcohol use, religiosity, study, and social life. *Journal of General Psychology, 136*(3), 231–241.

Dubé, L., & Paré, G. (2003) Rigor in information systems positivist case research: Current practices, trends, and recommendations. *MIS Quarterly, 27*(4), 597-636.

Dupps, W. J. (2017). Peer review: Get involved. *Journal of Cataract & Refractive Surgery, 43*(8), 997-998.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review, 14*(4), 532-550.

Epstein, S. (1995). What can be done to improve the journal review process? *American Psychologist, 50*(10) 883.

Frey, B. S. (2003). Publishing as prostitution?– Choosing between one's own ideas and academic success. *Public Choice, 116*(1-2), 205-223.

Grainger, D. W. (2007). Peer review as professional responsibility: A quality control system only as good as the participants. *Biomaterials, 28*(34), 5199-5203.

Glen, A. S. (2014). A new "golden rule" for peer review? *The Bulletin of the Ecological Society of America, 95*(4), 431-434.

Higgins, E. M. (2017). The state of peer review in criminology: Literary theory, perceptions, and the Catch-22 metaphor of peer review. *Journal of Criminal Justice Education*, *29*(4), 507-530.

Iivari, J. (2016). How to improve the quality of peer review? Three suggestions for system level changes. *Communications of the Association for Information Systems*, *38*(1), 263-273.

Janze, C. (2017). Design of a decentralized peer-to-peer reviewing and publishing market. In *Proceedings of the 25th European Conference on Information Systems* (ECIS), Guimarães, Portugal, June 5-10, 2017, 1713-1725.

Jennex, M. E. (2016). No free lunch: Suggestions for improving the quality of the review process. *Communications of the Association for Information Systems*, *38*(1), 307-313.

Justice, A. C., Cho, M. K., Winker, M. A., Berlin, J. A., Rennie, D., & Peer Investigators. (1998). Does masking author identity improve peer review quality?: A randomized controlled trial. *Jama*, *280*(3), 240-242.

Kachewar, S. G., & Smita B. S. (2013). Reviewer index: A new proposal of rewarding the reviewer. *Mens sana monographs, 11*(1), 274-284.

Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research,* 4th ed., United States: Thompson Learning.

King, E. B., Avery, D. R., Hebl, M. R., & Cortina, J. M. (2018). Systematic subjectivity: How subtle biases infect the scholarship review process. *Journal of Management*, *44*(3), 843-853.

Kim, J., Price, M., & Lau, F. (2014). The case study research method: Overview and proposed guidelines for reporting and evaluation illustrated with health informatics case studies. *International Journal of Health Information Management Research, 2*(1), 13-30.

Kirk, J., & Miller, M. L. (1986). *Reliability and validity in qualitative research*. Vol. 1, Sage.

Knudson, D. V., Morrow Jr, J. R. & Thomas, J. R. (2014). Advancing kinesiology through improved peer review. *Research Quarterly for Exercise and Sport, 85*(2), 127-135.

Kohli, R., & Limayem, M. (2006). Information systems research: Reference disciplines and theoretical contributions. *Proceedings of the Americas Conference on Information Systems.*

Kumar, S. S. (2015). Qualitative research: Examining the misconceptions. *South Asian Journal of Management*, *22*(3), 176-191.

Lee, A. S. (1989). A scientific methodology for MIS case studies. *MIS Quarterly, 13*(1), 33-50.

Lee, T. W. (1999). *Using qualitative methods in organizational research*. Thousand Oaks, CA: Sage Publications.

Lee, C. J., Sugimoto, C. R., Zhang, G. & Cronin, B. (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology, 64*(1), 2-17.

Liao, H. (2015). Reporting credibility in educational evaluation studies that use qualitative methods: A mixed methods research synthesis (Unpublished doctoral dissertation). Patton College of Education, Ohio University.

Lincoln, Y.S., & Guba, E.G. (1985). *Naturalistic Inquiry.* Newbury Park, CA: Sage Publications.

Macdonald, S. (2015). Emperor's new clothes: The reinvention of peer review as myth. *Journal of Management Inquiry, 24*(3), 264-279.

Mahoney, M. J. (1977). Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research, 1*(2), 161-175.

Mandviwalla, M., Ravi P., & Schuff, D. (2008). Improving the peer review process with information technology. *Decision Support Systems, 46*(1), 29-40.

Marques, K. C. M., Camacho, R. R., & Alcantara, C. C. V. (2015). Assessment of the methodological rigor of case studies in the field of management accounting. *Revista Contabilidade & Finanças, 26*(67), 27-42.

Matthews, D. (2016). Journal to rank peer reviewers to improve response times. Retrieved on January 16, 2019, from https://www.timeshighereducation.com/news/journal-rank-peer-reviewers-improve-response-times

McDonald, J.H. (2014). *Handbook of biological statistics* (3rd ed.). Baltimore, MD: Sparky House Publishing.

McPeek, M. A., DeAngelis, D. L., Shaw, R. G., Moore, A. J., Rausher, M. D., Strong, D, R., Ellison, A. M.,...Elgar, M. A. (2009). The golden rule of reviewing. *The American Naturalist, 173*(5), E155-E158.

Merriam, S. B. (1998). *Qualitative research and case study applications in education.* San Francisco, CA: Jossey-Bass, Inc.

Mitchell, T. R., Beach, L. R., & Smith, K. G. (1985). Some data on publishing from the authors' and reviewers' perspectives. *Publishing in the organizational sciences*, 248-264.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Newbury Park, CA: Sage Publications.

Milewicz, C., & Saxby, C. (2013). Leaders' social media usage intentions for in-bound customer communications. *Management Research Review, 36*(9), 849-867.

Miller, C. C. (2006). Peer review in the organizational and management sciences: Prevalence and effects of reviewer hostility, bias, and dissensus. *Academy of Management Journal*, *49*(3), 425-431.

Mora, M. (2016). Rejoinder to Iivari's (2016) paper: How to improve the quality of peer reviews? Three suggestions for system-level change. *Communications of the Association for Information Systems*, *38*(1), 284-291.

Mulligan, A., Hall, L., & Raphael, E. (2013). Peer review in a changing world: An international study measuring the attitudes of researchers. *Journal of the American Society for Information Science and Technology*, *64*(1), 132.

Nichols, A. L., & Webster, G. D. (2013). The single-item need to belong scale. *Personality and Individual Differences, 55*(2), 189–192.

Nandhakumar, J. (2010). Contrarian information systems studies. *European Journal of Information Systems*, *19*(6), 687-688.

Northcraft, G. B., & Tenbrunsel, A. E. (2011). Effective matrices, decision frames, and cooperation in volunteer dilemmas: A theoretical perspective on academic peer review. *Organization Science, 22*(5), 1277-1285.

Payne, G., & Payne, J. (2004). *Positivism and realism: Key concepts in social research*, Sage Publications, Newbury Park, CA.

Peters, D. P., & Ceci, S. J. (1982). Peer-review practices of psychological journals: The fate of published articles, submitted again. *Behavioral and Brain Sciences*, *5*(2), 187-195.

Petter, S. (2018). Embracing the golden rule of reviewing: A response to the tragedy of the scientific commons. *Communications of the Association for Information Systems, 42*(1), 654-662.

Pitsoulis, A., & Schnellenbach, J. (2012). On property rights and incentives in academic publishing. *Research Policy, 41*(8), 1440-1447.

Piekkari, R., Welch C., & Paavilainen, E. (2009). The case study as disciplinary convention: Evidence from international business journals. *Organizational Research Methods, 12*(3), 567-585.

Ragins, B. R. (2018). From boxing to dancing: Creating a developmental editorial culture. *Journal of Management Inquiry, 27*(2), 158-163.

Ragins, B. R. (2015). Editor's comments: Developing our authors. *Academy of Management, 40*(1), 1-8.

Ralph, P. (2016). Practical suggestions for improving scholarly peer review quality and reducing cycle times. *Communications of the Association for Information Systems*, *38*(13), 274-283.

Recker, J. (2016). Author responsibilities in improving peer reviews: A rejoinder to Iivari (2016). *Communications of the Association for Information Systems*, *38*(1), 314-320.

Rothwell, P. M., & Martyn, C. N. (2000). Reproducibility of peer review in clinical neuroscience. *Brain*, *123*(9), 1964-1969.

Sah, S., & Loewenstein, G. (2014). Nothing to declare: Mandatory and voluntary disclosure leads advisors to avoid conflicts of interest. *Psychological Science, 25*(2), 575-584.

Saunders, C. (2005a). Looking for diamond cutters. *MIS Quarterly*, *29*(1), iii-viii.

Saunders, C. (2005b). From the trenches: Thoughts on developmental reviewing. *MIS Quarterly*, *29*(2), iii-xii

Siebelink, J., van der Putten, P., & Kaptein, M. C. (2016). Do warriors, villagers and scientists decide differently? The impact of role on message framing. *International Conference on Intelligent Technologies for Interactive Entertainment*. Springer, Cham.

Sipior, J. C. (2018). Peer reviewer non-performance: The need to replenish the "commons." *Communications of the Association for Information Systems, 42*(1), 27.

Smith, R. (2006). Peer review: A flawed process at the heart of science and journals. *Journal of the Royal Society of Medicine, 99*(4), 178-182.

Stafford, T. F. (2016). Optimizing the business side of science: Publication review cycles and process management considerations. *Communications of the Association for Information Systems*, *38*(1), 296-306.

Stafford, T. F. (2018a). Reviews, reviewers, and reviewing: The "tragedy of the commons" in the scientific publication process. *Communications of the Association for Information Systems*, *42*(1), Article 25.

Stafford, T. F. (2018b). Rejoined and regenerated: Response to responses to "Reviews, Reviewers, and Reviewing: The 'Tragedy of the Commons' in the Scientific Publication Process." *Communications of the Association for Information Systems*, *42*(1), 663-670.

Starbuck, W. H. (2003). Turning lemons into lemonade: Where is the value in peer reviews? *Journal of Management Inquiry, 12*(4), 344-351.

Straub, D. W. (2009). Diamond mining or coal mining? Which reviewing industry are we in? *MIS Quarterly*, *33*(2), iii-viii.

Susskind, A. & Vicarri, A. (2011). A look at the relationship between service failures, guest satisfaction, and repeat-patronage intentions of casual dining guests. *Cornell Hospitality Quarterly, 52*(4), 438-444.

Squazzoni, F., Bravo, G., & Takacs, K. (2013). Does incentive provision increase the quality of peer review? An experimental study. *Research Policy, 42*(1), 287-294.

Tsui, A. S., & Hollenbeck, J. R. (2009). Successful authors and effective reviewers: Balancing supply and demand in the organizational sciences. *Organizational Researcher Methods*, *12*(2), 259-275.

Tomkins, A., Zhang, M., & Heavlin, W. D. (2017). Reviewer bias in single- versus double-blind peer review. *Proceedings of the National Academy of Sciences, 114*(48), 12708-12713.

Van Lange, P. (1999). Why authors believe that reviewers stress limiting aspects of manuscripts: The SLAM effect in peer review 1. *Journal of Applied Social Psychology, 29*(12), 2550-2566.

Venkatesh, V., Brown, S.A. & Bala, H. (2013). Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. *MIS Quarterly, 37*(1).

Vogt, W. P. (1999). *Dictionary of Statistics and Methodology* (2nd ed.). Thousand Oaks, CA Sage Publications, Inc.

Weber, R. (1999). The journal review process: A manifesto for change. *Communications of the AIS, 2*(12).

Weller, A. C. (2001). Editorial peer review: Its strengths and weaknesses. *Information Today, Inc*.

Yin, R.K. (1994). *Case Study Research: Design and Methods* (5th ed.). Thousand Oaks, CA Sage Publications, Inc.

Zavattaro, S. M., & Orr, S. K. (2017). Developing a research agenda: Your bread and butter (pp. 95-120). *Reflections on Academic Lives*. New York: Palgrave Macmillan.

About the Authors

Chris Street is an Associate Professor of Entrepreneurship at Hill School of Business. His research interests involve commercializing new technologies, corporate entrepreneurship, and technological adaptation processes in established companies.

Kerry W. Ward is an Associate Professor in the Information Systems and Quantitative Analysis department at the University of Nebraska, Omaha. He received his Ph.D. from the Kelly School of Business at Indiana University in Bloomington. He received his MBA from the University of Notre Dame and undergraduate degrees from Indiana University and Wabash College. His research interests include IS strategy, knowledge management, and research methods. Kerry has seven years' experience in consulting and public accounting with Deloitte, Touché, and PriceWaterhouseCoopers.

Appendix A

Table A.1: Researcher Version of Survey

Validity and Reliability in Case Study Research – Researchers											
In your opinion, please indicate how important the following practices are for establishing validity and/or reliability in case study research. Select "not applicable" if you do not believe the choice is applicable to case study research. If you are not familiar with the item, please indicate so by marking "Not Familiar With." – (NFW)											
	not important			somewhat important			very important				
	1	2	3	4	5	6	7	8	9	N/A	NFW
Development of rival hypothesis											
Using explanation-building											
Address rival explanations											
Using triangulation of data											
Coding of data											
Using pattern matching in analyzing the data											
Having multiple cases											
Use theory in single-case studies											
Use replication logic in multiple case studies											
Having a research question											
Having propositions or hypotheses											
Use multiple sources of evidence											
Having key informants review draft case study report											
Establish chain of evidence											
Interviewing multiple informants											
Develop case study database											
Determining inter-coder reliability											
Use case study protocol											
Using an interview guide											
Recording and transcribing interviews											

Table A.2: Reviewer Version of Survey

Validity and Reliability in Case Study Research - Reviewers											
AS A REVIEWER, please indicate how important the following practices are for establishing validity and/or reliability in case study research. Select "not applicable" if you do not believe the choice is applicable to case study research. If you are not familiar with the item, please indicate so by marking "Not Familiar With." – (NFW)											
Remaining survey was identical to the above survey	*not important*			*somewhat important*			*very important*				
	1	2	3	4	5	6	7	8	9	N/A	NFW