This version of the referenced work is the **post-print** version of the article—it is NOT the final published version nor the corrected proofs. If you would like to receive the final published version please send a request to any of the authors and we will be happy to send you the latest version. Moreover, you can contact the publisher's website and order the final version there, as well.

The current unpublished reference for this work is as follows:

Paul Benjamin Lowry, John D'Arcy, Bryan Hammer, and Gregory D. Moody (2016). "'Cargo Cult' science in traditional organization and information systems survey research: A case for using nontraditional methods of data collection, including Mechanical Turk and online panels," ***Journal of Strategic Information Systems (JSIS)*** (accepted 25-June-2016)

* = corresponding author

If you have any questions, would like a copy of the final version of the article, or would like copies of other articles we've published, please contact any of us directly, as follows:

- **Prof. Paul Benjamin Lowry:**
  - Email: Paul.Lowry.PhD@gmail.com
  - Website: http://www.cb.cityu.edu.hk/staff/pblowry
  - System to request Paul's articles: https://seanacademic.qualtrics.com/SE/?SID=SV_7WCaP0V7FA0GWWx
- **Dr. John D'Arcy:**
  - Email: jdarcy@udel.edu
  - Website: http://www.lerner.udel.edu/faculty-staff/john-darcy
- **Dr. Bryan Hammer:**
  - Email: bryan.hammer@okstate.edu
  - Website: http://spears.okstate.edu/profiles/?id=489
- **Dr. Gregory D. Moody:**
  - Email: Gregory.Moody@unlv.edu
  - Website: http://faculty.unlv.edu/wpmu/gmoody/

* = corresponding author

"Cargo Cult" Science in Traditional Organization and Information Systems Survey Research: A Case for Using Nontraditional Methods of Data Collection, Including Mechanical Turk and Online Panels

INTRODUCTION

The goal of our study was not to replicate Jia and Reich's information technology (IT) climate study (2008, 2011, 2013; 2008) but to build on and depart from some of their ideas as well as to try a different methodological approach that expanded on the traditions of IT climate research. That is, rather than using a matched approach to represent "shared" perceptions of IT service climate, we evaluated IT employees' perceptions of IT service. For this reason, we named our reconceptualized IT service construct "internal IT service perceptions." We also used Amazon.com's Mechanical Turk (MTurk), a unique source of online panel data, to gain access to IT employees at hundreds rather than only a handful of organizations. On these points we stand guilty as charged—and are unrepentant. Accordingly, we agree with the respondent authors' claim that our study suffered from certain general limitations, which we in fact noted in our paper.

However, the respondent authors set forth three major contentions about our study that lie beyond the contradistinction between shared and individual cognition. The first contention is that MTurk samples have different demographics/backgrounds than those found within a single or limited set of organizations. The second is that these differences systematically bias the data and invalidate its applicability to organizational research. The third is that due to this broader variability in the MTurk population, the results of our study are generalizable but not specific (i.e., contextual), like those of the majority of published organizational research.

Before addressing these critiques, we wish to discuss a more important topic, on which we disagree strongly with the respondents: the assumption that traditional theory and paper-based survey methods in IT climate research, and organizational research in general,

are sacrosanct. A pattern has taken hold in which traditional organizational researchers, reviewers, and editors are quick to misconstrue and reject new methods while defending the "best practices" of paper surveys, which have been the methodology of choice for several decades. Although organizations themselves have implemented significant innovations, the published research on organizations has not undertaken innovation to the same degree. Traditionalists and the researchers who make up the reviewing system in the organization science and information systems (IS) fields are quick to downplay the legitimacy of new theories and methods, but they fail to apply the same level of scrutiny to their own traditions. This thwarts scientific progress.

Richard Feynman, in his 1974 Caltech commencement address, used the apt metaphor of a *cargo cult* to describe scientists who cling to past thinking and methods:

> *In the South Seas there is a cargo cult of people. During the war they saw airplanes land with lots of good materials, and they want the same thing to happen now. So they've arranged to make things like runways, to put fires along the sides of the runways, to make a wooden hut for a man to sit in, with two wooden pieces on his head like headphones and bars of bamboo sticking out like antennas—he's the controller—and they wait for the airplanes to land. They're doing everything right. The form is perfect. It looks exactly the way it looked before. But it doesn't work. No airplanes land. So I call these things cargo cult science, because they follow all the apparent precepts and forms of scientific investigation, but they're missing something essential, because the planes don't land.* (Wilson, 2002, p. 155)

We similarly argue that "planes will not land" in the organization sciences and IS as long as old theories, methods, and measures are treated as sacrosanct and established as "best practices." Just because an approach worked or was "publishable" in the past (e.g., administering paper-based surveys in organizations) does not mean that is the best path forward. This kind of thinking protects tradition but likewise guarantees "bland" (simple) replications, subject to the same limitations, that will not meaningfully advance theory or IS research. This is same kind of problem recently lamented by Grover and Lyytinen (2015) who call for more openness and innovativeness and a "push to the edges" in both approaches to theory and method to get IS research out of its doldrums. We hope to contribute insights

into this topic that will help our discipline avoid the practice of cargo-cult science. This does not mean we propose a complete abandonment of current methodologies, but we do maintain that IS researchers should cultivate scientific attitudes toward newer methods.

MTurk represents a potential revolution in crowd-sourced organization data collection that should not be glibly dismissed. Traditional organization and IS researchers would also greatly benefit from noting that virtually every objection that they may have to MTurk applies to survey research in general, and especially to traditional paper-based organizational surveys. Thus, exploring the use of MTurk for organizational studies may actually improve traditional methods, because the strategies for ensuring MTurk data quality could easily be applied to traditional methods.

We agree that certain questions and challenges apply to gathering organizational data via MTurk (or other forms of online data panels), and certainly researchers need to be cautious about any method or data source they use, especially new ones. The respondent authors presented critiques of MTurk; however, their presentation was unbalanced and failed to reflect the positive ways MTurk is being used by researchers. Indeed, their paper largely followed the argumentation and layout of Harms and DeSimone (2015), but in citing contrary results from the literature, they failed to cite the many positive aspects of MTurk mentioned in exactly the same citations. Moreover, they failed to mention the many ways MTurk's limitations have been addressed with concrete solutions. They also painted an inaccurate picture of the demographics and sampling issues associated with MTurk, assuming the worst possible motives and characteristics in our study's sample—despite evidence we provided to the contrary.

Interestingly, many of their arguments are similar to those first posed for Internet-based samples when these were seen as a threat to traditional research (Birnbaum, 2004; Buchanan & Smith, 1999b; Mathy et al., 2002). Those doubts were largely put to rest by

empirical examinations that compared Internet-based samples to those from traditional sources (Buchanan & Smith, 1999a; Meyerson & Tryon, 2003; Stanton, 1998), with many guidelines and recommendations provided for the latter (Birnbaum, 2004; Buchanan & Smith, 1999b; Reips, 2002; Skitka & Sargis, 2006). Much like these researchers before us, we hope to assuage the doubts and refute the criticisms regarding MTurk use by providing science-based guidance and recommendations.

In the remainder of this response, we proceed by summarizing some of the possible advantages of using MTurk. We then discuss some of the major limitations of MTurk, including those outlined by the respondent authors, and the ways in which these limitations can be addressed. As a result of juxtaposing the issues associated with MTurk studies and those associated with traditional organizational paper-based survey research, we find that their limitations are strikingly similar. In fact, a case can be made that paper-based surveys suffer from significant limitations. Most importantly, we conclude with concrete solutions that can be implemented to improve any survey-based research in IS—whether collecting responses from the diverse subject pools of MTurk and other online data panels or doing so from one or several organizations with paper-based surveys.

CRITIQUES OF MTURK IN ORGANIZATIONAL AND IS RESEARCH

A compelling new source of data, MTurk leverages the crowdsourcing model that enables new means of accessing and filtering that were previously impossible. Researchers outside the organization sciences and IS have long been using MTurk as a valid source of data. “Psychological researchers have long been open-minded in adopting new technologies to aid in the process of conducting research, while staying committed to protecting the integrity of that process,” (p. 304) and as a result, they have led the way in the use of MTurk for data collection. Their adoption of MTurk has included a balanced, critical examination of its weaknesses and strengths (Rouse, 2015).

Within the IS field, MTurk has already been used as a valid, reputable source for data collection. Steelman et al.'s (2014) study identified 20 quality IS publications that used MTurk, and there are many more to date. Using the technology acceptance model (TAM, Davis, 1989) and expectation-disconfirmation theory (EDT, Oliver, 1980), they statistically compared multiple samples (e.g., students, consumer panel) against several different MTurk samples (e.g., US-only, worldwide, international-only) and found a single difference within the TAM, which they attributed to cultural differences. These authors concluded that MTurk is a viable sampling alternative provided researchers follow certain guidelines in collecting data and reporting results. In fact, many of their guidelines, which we followed in our paper, address the respondent authors' concerns.

Certainly, no one should expect all the potential benefits of MTurk to apply to a given study, and certainly not without systematically dealing with the limitations and validity threats that arise in using this innovative form of data collection. In this section, we respond to the common critiques of using MTurk for organizational and IS research, made not just by the respondent authors but in the literature more generally. We also describe how MTurk (and to an extent, online panels in general) can address these critiques, as summarized in Table 1.

The Convenience Critique: The Myth of the Organizational Study as the Gold Standard for Sampling

One of the ironies of traditional organizational studies conducted with paper-based surveys is that they rely on one of the most biased forms of convenience sampling. That is, their sampling is typically based on the relationships of the researchers with the leaders of a very small number of non-randomly-sampled organizations. Certainly this is the case with the one organizational study by Jia et al. (2008, 2011, 2013; 2008) on which we build. (Note that, of these four papers, one was a conceptual piece, whereas the remaining three used a single

Table 1. Comparing Common Critiques of MTurk to Traditional Organizational Surveys and Ways to Address Them (Part 1 of 3)

Critique of MTurk	Does Critique Also Apply to Traditional Organizational Survey Research?	How Can MTurk Address These Issues?	Potential MTurk Advantage
It relies on convenience sampling.	Yes, and this problem is worse with traditional studies. Traditional studies typically use only one or a select few organizations for their samples, which are often based on researcher contacts in one geographic location.	The sampling diversity in MTurk affords researchers the ability to include participants who may not be amendable to traditional organizational research (e.g., non-WEIRD samples). Further, researchers can conduct targeting sampling to address omitted variables, such as organizational culture, country, etc.	MTurk provides quicker data collection around the clock regardless of researchers' and participants' locations (Buhrmester et al., 2011; Goodman et al., 2013; Kittur et al., 2008). Pilot tests and exploratory studies can be conducted more quickly and with higher generalizability than before, so that final data collections (including those done via traditional means) are of much higher quality and impact (Scott et al., 2015).
We cannot be sure of the respondents' identities.	Yes, with the exception of non-anonymous surveys, which create strong social desirability bias and observer effects. Short of having researchers or an independent party actually hand out the surveys and verify the respondents, traditional approaches require trust that those who complete the surveys are in fact who they say they are (e.g., employed by a particular organization).	Researchers can use MTurk's identity data to screen participants based on (i) their level of expertise, (ii) the number of previous HITs they have completed and the percentage completed successfully, (iii) their verified language, (iv) their verified location, based on IP address, and (v) relevant demographic questions identified by the researchers. For example, additional screening questions could be introduced that only target participants could successfully answer (e.g., a question regarding the URL of the organization's website).	The MTurk population is huge and increasingly studied and known; thus, for the first time, organizational researchers can perform random and stratified sampling and with large numbers of organizations (Landers & Behrend, 2015). Studies can be conducted simultaneously in a large number of languages and cultures, a strategy not scalable in other settings. MTurk enables researchers to apply multiple levels of screeners in order to target specific respondents, including on geography, demographics, and language—an approach that is simply not possible in other research settings (Rouse, 2015; Steelman et al., 2014; Woo et al., 2015).

Table 2. Comparing Common Critiques of MTurk to Traditional Organizational Surveys and Ways to Address Them (Part 2 of 3)

Critique of MTurk	Does Critique Also Apply to Traditional Organizational Survey Research?	How Can MTurk Address These Issues?	Potential MTurk Advantage
Super users/Turkers may distort results.	Generally, no. We note, however, that organizational researchers never ask respondents if they are active participants on MTurk or online panels, or if they commonly respond to all survey requests. It is possible that all organizational research is currently based on 'super survey takers'—however, it would be impossible to determine or control for these effects.	Super users can be readily identified based on the number of HITs completed. "Ballot stuffing" prevention tools can also be used to prevent respondents from taking a survey repeatedly. And, because MTurk super users generally put less time and effort into a task, completion time can be used as an elimination criterion.	Experienced survey takers should be less likely to be confused about survey questions or to become fatigued while completing surveys. Thus, the ability to identify such respondents with MTurk samples could enable higher quality responses.
Respondents may lie and cheat.	Yes, and this problem may be worse with traditional studies. Deviant or angry employees may intentionally undermine a study, especially if they are coerced to participate without compensation or are too busy with important work. On the other hand, some employees may attempt to bolster their relationships by engaging in hypothesis guessing in an effort to ingratiate themselves with the researchers or management.	Researchers can utilize the structural mechanisms in MTurk, such as those that afford the screening mechanisms described above, to verify respondents and communicate this verification process in the HIT description. Additionally, although not MTurk specific, electronic tools such as randomly appearing attention-trap questions, reverse-coded questions, and questions that ask respondents to positively affirm the accuracy of their responses are useful.	MTurk uses a built-in system in which dishonest or deviant Workers are punished. Filters allow Workers with a long-standing history of completed HITs to participate in requests (Steelman et al., 2014).
Respondents may not give the tasks their full attention.	Yes. Employees are notorious for attention lapses at work, facing frequent and multiple interruptions.	Although not MTurk specific, electronic tools such as randomly appearing attention-trap questions, reverse-coded questions, and questions that ask respondents to positively affirm the accuracy of their responses are useful.	When researchers take the proper precautions, MTurk data is highly reliable and in most cases provides data that is just as good or better than that provided by traditional approaches (Goodman et al., 2013; Mason & Suri, 2012; Steelman et al., 2014).

Table 3. Comparing Common Critiques of MTurk to Traditional Organizational Surveys and Ways to Address Them (Part 3 of 3)

Critique of MTurk	Does Critique Also Apply to Traditional Organizational Survey Research?	How Can MTurk Address These Issues?	Potential MTurk Advantage
People who complete surveys for pay may not be "normal."	Yes. How "normal" are the minority of respondents who are willing to participate in a traditional survey without pay?	An approach for garnering "normal" respondents is to offer pay roughly equivalent to the federal minimum wage, at least for US samples.	Very unique data can now be gathered that in the past required travel and invasive contact. It is dramatically less expensive and time consuming, and it enables a wider range of scientists and students to access research participants (Rouse, 2015)
There could be data quality issues.	Yes. These problems are worse with paper-based surveys, because many advanced technological techniques cannot be used, including a basic requirement to prevent common method bias by randomizing the order of questions for each respondent.	Researchers can use the general techniques for electronic surveys described above, in addition to the MTurk-specific screening techniques, to improve data quality.	MTurk has much greater potential for cross-validation and generalization testing of data (Rouse et al., 2015; Steelman et al., 2014). It also offers the ability to implement highly sophisticated methods—beyond Likert-type survey studies—that are problematic or invasive in other settings: truly anonymous longitudinal studies, product and systems-use studies, and studies about controversial topics in which social desirability bias is high (e.g., computer abuse, employee deviance, employee fraud) (Berinsky et al., 2012).
Inability to contextualize research.	Traditionally, this has not been a problem in organizational research because it is applies in a couple of organizations in very specific contextual circumstances. However, most the sampling ironically is WEIRD.	The ability to filter Workers by limiting a HIT based on specified criteria can provide researchers with highly context-specific participants. Researchers, for example can choose location, culture, organization size, language, industry type, and so on. Can break through the WEIRD sample barriers more easily because of greater reach of participant types.	MTurk allows for stronger contextualization; a common flaw of organizational theory is that it is not contextualized. Targeting highly unique respondents and organizations allows for contextualization (Smith et al., 2015).

dataset from the same four organizations and the same measurement; thus, we report their IT service climate research as one study, not three). To wit, Landers and Behrend emphasized this inconvenient truth: "Organizational samples are not gold standard research sources; instead, they are merely a specific type of convenience sample with their own positive and negative implications for validity" (2015, p. 142). They went on to chastise traditional researchers for promoting the notion that "more difficult to collect data is better than easy to collect data" in order to defend the use of organization samples. Zhu et al. (2015) pointedly took up this issue (p. 231):

> *Default admonishment (and apologizing) for not using organizational samples is inappropriate and hinders the growth of the field. Indeed, carefully considering what samples tell us about an effect can help researchers more fully understand the phenomena we study. Although feasibility and appropriateness of using a particular sample are important to consider, they should not be confused with convenience.*

Landers and Behrend went on to offer the following argument (p. 158):

> *Organizational samples are not probability samples and should not be treated as such. In fact, organizational samples are often quite limited in ambiguous and difficult-to-measure ways. Not only are employees within an organization range restricted on whatever selection devices were used to hire them, but there are also a host of omitted variables at higher levels of analysis that may influence the results of a particular study (e.g., organizational culture, industry, country).*

Of particular interest regarding this limitation of organizational samples, MTurk is a viable alternative with which researchers can conduct targeted sampling in an effort to address many of the omitted variables that it was previously impossible to examine—including organizational culture, ethic culture, industry, and country. Moreover, MTurk allows for dramatically greater generalizability, because an organizational researcher can easily study hundreds or thousands of organizations instead of the small handful commonly accessed using traditional means.

The Identity Critique: How Do We Know Respondents Are Who They Say They Are?

Verifying that a study's participants are legitimate targets for that study should be of concern to any organizational researcher, and there are no perfect solutions to this dilemma,

regardless of the survey method employed. Distributing a paper survey to a target organization does not absolve researchers of their responsibility to address this problem. If such a survey is truly anonymous, then how can the researchers know it has reached its target population and that the responses are not faked, coerced, ballot stuffed, or manipulated by management? How can they be sure that disgruntled employees will not use it as an opportunity to pretend to speak as management or simply to submit noxious data? In the one study by Jia and Reich (2008, 2011, 2013; 2008) mentioned above, paper-based surveys were distributed to client contacts, who then distributed and recollected the surveys. Aside from providing little identity assurance, such a process introduces perceived-anonymity and social-desirability bias issues, the dangers of which we discuss below.

MTurk enables several solutions to this dilemma, which provide certain assurances that are not possible in traditional organizational studies. First, all MTurk Workers must provide their identity and apply to become a Worker. MTurk actively data mines and tracks its Workers for evidence of fraud, falsified identities, and use of false locations. Those who do not comply with its terms of use are banned. Workers are also evaluated by those who hire them, such that those who provide poor data or responses are openly evaluated and reported. To become a master Worker, one must have a clean record of doing many human intelligence tasks (HITs)[1] without incident and without assuming a false identity. MTurk also provides a mechanism through which Workers can be evaluated and certified for skills such as language proficiency. Thus, aside from pay, there are strong structural mechanisms in place to ensure not only Workers' identities but also their meaningful responses.

Landers and Behrend (2015) said it best: "The use of MTurk may solve a problem that has vexed the research community for decades, namely, the severe oversampling of

[1] A HIT is a single task created by Requesters (in our context, organizational researchers) and then performed by MTurk Workers in exchange for compensation.

participants from Western, educated, industrialized, rich, and democratic (WEIRD) backgrounds" (p. 152). They further quipped that "because most researchers are WEIRD and consult for WEIRD organizations, most of their research is WEIRD too" (Landers & Behrend, 2015, p. 159). For example, the TAM did not hold up in the context of an international sample from MTurk (Steelman et al., 2014). Although MTurk critics may view this as a limitation of MTurk, we argue this result underlines the point that the majority of IS theories and models are based on WEIRD samples and have not received proper vetting using more diverse populations (Buhrmester et al., 2011).

In our study, we addressed the identity critique by using several of the techniques listed in Table 1. For example, we asked participants to openly describe their organizational role and what role they played in IT; we retained only those participants who provided insightful answers that used American English. Moreover, we coded these respondents' responses to further determine their internal roles and reported on this in the paper.

The Super Turkers Critique: Highly Experienced Respondents May Distort Results

This critique maintains that the use of "professionals" or "super users" can result in a sample that distorts the results of a study. However, this is true only to the extent that a particular kind of knowledge or learning effect threatens the validity of a given study. We were frankly unsure whether such effects could compromise our study, and thus we sidestepped the issue by only allowing non–super users who had successfully completed a maximum of one to three HITs to participate in our study. This ensured that they were not "professional" respondents.[2]

[2] We did not discuss this methodological point in our paper but probably should have, because future organization studies need to determine whether "super Turkers" respond differently than inexperienced Turkers. Although reviewers and editors often ask authors to omit certain methodological details from their papers because of space limitations, we believe that as many such details as possible should be reported in MTurk studies to improve the ability of researchers to compare results.

Ironically, to our knowledge, no survey administered in a traditional organizational study has asked respondents if they are actively engaged in MTurk or online panels. If learning effects or "professional respondents" are a serious concern for an organizational study, it should filter out such people, because research suggests that substantial numbers of the people who participate in such panels are employed part-time or full-time by real organizations (Miliaikeala et al., 2014). With respect to MTurk, this issue can be addressed by eliminating respondents with unreasonably fast response times, because research indicates that super users put less time and effort into tasks (Mason & Suri, 2012).

The Lying Critique: Respondents May Lie and Cheat

The lying critique of MTurk is one of the weakest, because respondents can lie and cheat in virtually any study for a number of reasons. In traditional organizational paper surveys, respondents have an incentive to lie when, for example, they are disgruntled, intentionally want to undermine their organization or the researchers who are imposing on their time, perceive they can skew the results to steer the organization in a certain direction, see the opportunity to make themselves or desired entities look good or bad, want to rush through the survey to get on to more important work, are concerned that the person who distributed the survey to them will read it, and so on. These possibilities are usually ignored in traditional organizational studies. Some of the motivation for lying stems from a lack of anonymity in paper-based surveys; this is where online-based systems excel, because they can provide true anonymity (Gosling et al., 2004).

Moreover, MTurk provides many structural disincentives against general dishonesty, and those who conduct research with MTurk typically introduce additional electronic tools into their online surveys to prevent lying, a precaution we took in our study. For example, insofar as dishonesty produces a cognitive burden that diminishes one's capacity to recognize such safeguards, researchers using MTurk can (1) add randomly appearing attention-trap

questions that only participants who are paying attention and taking the study seriously will notice, (2) use reverse-coded questions to detect answering anomalies, (3) ask questions that only those with insider knowledge would know, and (4) ask respondents to positively affirm the accuracy of their responses as a means of increasing accountability and accuracy of results (Rouse, 2015; Steelman et al., 2014). We employed the first three techniques. Traditional organizational researchers typically ignore all of these techniques and assume honesty and good will in all of their respondents.

The Attention Critique: Respondents Might Not Be Paying Attention Fully

Like the previous critique, this one is especially flimsy because it applies to all behavioral research. There is certainly no scientific reason to believe an unpaid volunteer (or coerced) participant is going to pay more attention than one who is paid and has structural incentives to pay attention. Aside from using advanced techniques like eye tracking, there are several simple solutions to this problem that MTurk researchers have successfully implemented and that may be used to improve traditional organizational research (Steelman et al., 2014). (Note, however, that the first two techniques are more difficult and less effective with a paper-based survey method.[i]) These strategies include (1) adding attention-trap questions that require thought to answer (Goodman et al., 2013; Rouse, 2015), (2) track the amount of time spent on each question and filter out people who are unusually fast, (3) explain the scientific importance of the study (Goodman et al., 2013), (4) convey to the participants the importance of paying attention, because although some questions will be similar, many others will be unrelated (Goodman et al., 2013), and (5) ask respondents to positively affirm the accuracy of their responses, which increases accountability and attention (Rouse, 2015). We implemented the first four techniques in our study.

The Abnormality Critique: People Who Complete Surveys for Pay May Not Be Normal

Issues with selection and participation bias plague every form of behavioral and organizational research (Landers & Behrend, 2015). There is no reason researchers should not assume that the population of Workers represents a balanced random sample of the world population of organizational employees. Likewise, the results of a study that uses "voluntary" respondents with low response rates in a small number of organizations with which a researcher has a relationship cannot be generalized to any particular known population. Neither sample should be labeled as "normal." In fact, "normal" may never be an appropriate label for a behavior research sample.

Nevertheless, there is a common belief among traditionalists that there *must* be something wrong about people who are willing to participate in surveys for money via online panels (e.g., Qualtrics, SurveyMonkey) or crowdsourcing markets such as MTurk (Goodman et al., 2013). However, extrinsic motivations such as payment are a concern only when they modify the effect size of the results by changing the nature of the task, which is possible but uncommon in most studies (Landers & Behrend, 2015). Yet, the same people who voice this concern see no issue with relying on voluntary respondents in low-response-rate samples. What about the typically 80% or more of target participants in an organization who are too busy, self-important, or uninterested to respond to a traditional survey? Both such groups indeed could be represented by different balances of extrinsic and intrinsic motivations, and there could certainly be other relevant demographic differences.

Aside from using larger populations of participants and organizations, the primary solution to this issue is to be very clear about the sampling and screening methods employed in the research. Steelman et al. (2014) provided a comprehensive list of such information that should be collected and reported in MTurk studies. Reporting such information allows researchers to compare results across many samples in order to better understand

commonalities and differences. Nonetheless, researchers agree that when pay is used, a reasonable level of pay should be offered for the time involved, such that it is not exploitative or overly generous for the target group (Rouse, 2015). A recommended approach for US samples is to offer pay that is roughly equivalent to the federal minimum wage based on the average minutes spent on a task (Downs et al., 2010).

The Quality Critique: There Could Be Data Quality Issues

A common rhetorical technique used by MTurk opponents is to adduce a "big list" of citations of every possible bad thing that could happen with MTurk studies and then call into question the general integrity and quality of MTurk data. Our respondent authors followed this pattern. To wit, Landers and Behrend (2015, p. 153) noted that

> *the most substantial barrier to greater adoption of MTurk as a potential method for convenience sampling is reviewer unfamiliarity with and subsequent dismissal of the approach because of a variety of untested assumptions. In our experience, such reviewers do raise concerns worth considering, but the reviewers assume these issues are unique to MTurk as a data source or overestimate the impact of the issue on data quality.*

Not only is the big-list tactic unscientific, but it also ignores two wide-ranging approaches that can address these concerns: (1) the many techniques, listed above, that can be used to deal with the limitations of these data, and (2) the many techniques that can be used in any behavioral study to improve and ensure data quality.

As we have argued, most problems with MTurk data quality involve issues that are common to all surveys, including scaling, measurement, monomethod bias, and the like. In our study, we used partial least squares and most of techniques in this area (e.g., Gefen & Straub, 2005; Lowry & Gaskin, 2014), and the results of the factorial validity and data-quality checks were very strong, with one exception mentioned in the paper.[3]

[3] As mentioned in the original paper, one exception was the first-order reflective constructs that compose IT agility, which had high loadings on each other, indicating lack of discrimination within these normally related subconstructs. Thankfully, these discriminated against IT service quality and worked well with the model, as IT agility was formatively created by them. For these kinds of second-order constructs,

CONCLUSION

In our collective program of research, we have performed most forms of organizational research, and our anecdotal experience can illuminate the mindset of researchers in this community. On the traditional side, we have conducted organizational studies based on widely accepted theory in which Likert-type-scale paper or electronic surveys were distributed to a small number of organizations, and such studies have consequently been warmly embraced and offered a straightforward path to publication (Boss et al., 2015; D'Arcy et al., 2009; Hsu et al., 2015; Moody & Galletta, 2015). However, we have also conducted research to improve organization outcomes with less conventional methodological approaches (and associated theory), such as with MTurk (Chatterjee et al., 2015; Lowry & Wilson, 2016; Steelman et al., 2014) and online panel services (D'Arcy et al., 2014; Lowry et al., 2015; Posey et al., 2015), qualitative interviews with grounded theory (Moody et al., 2016b; Parks et al., 2016), polynomial and curvilinear data relationships (Moody et al., 2016a), real-time eye tracking (Twyman et al., 2014), case data (Wall et al., 2016), interactions with embodied agents in decision-making (Burgoon et al., 2016), mouse tracking (Jenkins et al., 2014), data mining of secondary organization data (Hansen et al., 2007), and so on.

However, our experience is that although these nontraditional approaches can create breakthrough ideas and very exciting research—and have resulted in what is arguably our best work—they make the publishing process a much longer and more arduous path. That is, such approaches create more resistance in editors and reviewers who are traditionalists, because such people are often unreflectively dismissive of methods they have not used themselves and of new ways to conceptualize old ideas. Worse, although new theories,

multicollinearity is a more significant threat (Cenfetelli & Bassellier, 2009), and these constructs passed that test well.

conceptualizations, and methods could actually lead organizational and IS research out of its doldrums of incremental staid research, such researchers often do everything in their power to block nontraditional studies in the review process. . This is a troubling state of organizational research that calls for a rethinking of tradition. Peer review should not result in the pressuring of researchers to follow traditional methods or conform to a reviewer's way of thinking; it should open up discussion on topics of mutual interest, from different perspectives.

However, this is clearly the case with the biases and myths that are used to dismiss the use of MTurk or other paid online panel services to gather data. It is especially surprising to see researchers claim that certain theories must be paired with certain methodological practices, which could not be further from the truth in scientific inquiry. Theory and method should never be conflated. No "proven" method without inherent flaws exists, and organizational research would benefit from a clearer articulation and understanding of the range of methods and data sources available, along with their limitations and advantages.

The issues surrounding the use of MTurk and other nontraditional sampling approaches is reminiscent of the trap into which the discipline of physics fell once the theory of general relativity was widely accepted: the theory and its methods were treated as an unassailable tradition rather than debatable points that could be improved by new theory and methods. Not surprisingly, traditionalists in physics undermined scientific progress as a result. It turns out the theory of relativity had flaws and was not the last word on physics, but many physicists were mired in hero worship of Einstein. The American theoretical physicist John Archibald Wheeler (1911–2008), however, flouted tradition by reopening the dialogue on the theory of relativity and thus helped to identify its flaws and move physics forward. On this point, he said that in science, "there is no law except the law that there is no law." We argue that the law of no law also applies to organization science and IS. Basing organizational research on sample sizes of one or a few organizations goes against standard

statistical practices. Expanding our horizons on this matter may transform our understanding of organization science.

REFERENCES

Berinsky, A. J., Huber, G. A., and Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk. *Political Analysis, 20*(3), 351-368.

Birnbaum, M. H. (2004). Human research and data collection via the Internet. *Annual Review of Psychology, 55*, 803-832.

Boss, S. R., Galletta, D. F., Lowry, P. B., Moody, G. D., and Polak, P. (2015). What do users have to fear? Using fear appeals to engender threats and fear that motivate protective security behaviors. *MIS Quarterly, 39*(4), 837-864.

Buchanan, T. and Smith, J. L. (1999a). Research on the Internet: Validation of a World-Wide Web mediated personality scale. *Behavior Research Methods, 31*(4), 565-571.

Buchanan, T. and Smith, J. L. (1999b). Using the Internet for psychological research: personality testing on the World Wide Web. *British Journal of Psychology, 90*(Part 1), 125-144.

Buhrmester, M. D., Kwang, T., and Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*, 3-5.

Burgoon, J. K., Bonito, J. A., Lowry, P. B., Humpherys, S. L., Moody, G. D., Gaskin, J. E. et al. (2016). Application of expectancy violations theory to communication with and judgments about embodied agents during a decision-making task. *International Journal of Human-Computer Studies, 91*(July), 24-36.

Cenfetelli, R. T. and Bassellier, G. (2009). Interpretation of formative measurement in information systems research. *MIS Quarterly, 33*(4), 689-708.

Chatterjee, S., Moody, G. D., Lowry, P. B., Chakraborty, S., and Hardin, A. (2015). Strategic relevance of organizational virtues enabled by information technology in organizational innovation. *Journal of Management Information Systems, 32*(3), 158-196.

D'Arcy, J., Herath, T., and Shoss, M. K. (2014). Understanding employee responses to stressful information security requirements: A coping perspective. *Journal of Management Information Systems, 31*(2), 285-318.

D'Arcy, J., Hovav, A., and Galletta, D. (2009). User awareness of security countermeasures and its impact on information systems misuse: A deterrence approach. *Information Systems Research, 20*(1), 79-98.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340.

Downs, J. S., Holbrook, M. B., Sheng, S., and Cranor, L. F. (2010, April 10-15). *Are your participants gaming the system?: Screening mechanical turk workers.* Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA.

Gefen, D. and Straub, D. W. (2005). A practical guide to factorial validity using PLS-Graph: Tutorial and annotated example. *Communications of the Association for Information Systems, 16*(5), 91-109.

Goodman, J. K., Cryder, C. E., and Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making, 26*(3), 213-224.

Gosling, S. D., Vazire, S., Srivastava, S., and John, O. P. (2004). Should we trust Web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. *American Psychologist, 59*(2), 93-104.

Grover, V. and Lyytinen, K. (2015). New state of play in information systems research: The push to the edges. *MIS Quarterly, 39*(2), 271-296.

Hansen, J. V., Lowry, P. B., Meservy, R., and McDonald, D. (2007). Genetic programming for prevention of cyberterrorism through dynamic and evolving intrusion detection. *Decision Support Systems, 43*(4), 1362-1374.
Harms, P. D. and DeSimone, J. A. (2015). Caution! MTurk workers ahead—Fines doubled. *Industrial and Organizational Psychology, 8*(2), 183-190.
Hsu, J., Shih, S.-P., Hung, Y. W., and Lowry, P. B. (2015). The role of extra-role behaviors and social controls in information security policy effectiveness. *Information Systems Research, 26*(2), 282-300.
Jenkins, J. L., Grimes, M., Proudfoot, J., and Lowry, P. B. (2014). Improving password cybersecurity through inexpensive and minimally invasive means: Detecting and deterring password reuse through keystroke-dynamics monitoring and just-in-time warnings. *Information Technology for Development, 20*(2), 196-213.
Jia, R. and Reich, B. H. (2008, December 14-17). *IT service climate: The validation of an antecedent construct for IT service quality*. Paper presented at the 29th International Conference on Information Systems (ICIS 2008), Paris, France.
Jia, R. and Reich, B. H. (2011). IT service climate—An essential managerial tool to improve client satisfaction with IT service quality. *Information Systems Management, 28*(2), 174-179.
Jia, R. and Reich, B. H. (2013). IT service climate, antecedents and IT service quality outcomes: Some initial evidence. *Journal of Strategic Information Systems, 22*(1), 51-69.
Jia, R., Reich, B. H., and Pearson, J. M. (2008). IT service climate: An extension to IT service quality research. *Journal of the Association for Information Systems, 9*(5), 294-320.
Kittur, A., Chi, E. H., and Suh, B. (2008, April 5-10). *Crowdsourcing user studies with Mechanical Turk*. Paper presented at the 26th Conference on Human Factors in Computing Systems, Florence, Italy.
Landers, R. N. and Behrend, T. S. (2015). An inconvenient truth: Arbitrary distinctions between organizational, Mechanical Turk, and other convenience samples. *Industrial and Organizational Psychology, 8*(2), 142-164.
Lowry, P. B. and Gaskin, J. (2014). Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. *IEEE Transactions on Professional Communication, 57*(2), 123-146.
Lowry, P. B., Posey, C., Bennett, R. J., and Roberts, T. L. (2015). Leveraging fairness and reactance theories to deter reactive computer abuse following enhanced organisational information security policies: An empirical study of the influence of counterfactual reasoning and organisational trust. *Information Systems Journal, 25*(3), 193-230.
Lowry, P. B. and Wilson, D. W. (2016). Creating agile organizations through IT: The influence of internal IT service perceptions on IT service quality and IT agility. *Journal of Strategic Information Systems, forthcoming*.
Mason, W. and Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44*(1), 1-23.
Mathy, R. M., Schillace, M., Coleman, S. M., and Berquist, B. E. (2002). Methodological rigor with Internet samples: New ways to reach underrepresented populations. *Cyberpsychology & Behavior, 5*(3), 253-267.
Meyerson, P. and Tryon, W. W. (2003). Validating Internet research: A test of the psychometric equivalence of Internet and in-person samples. *Behavior Research Methods, Instruments, & Computers, 35*(4), 614-620.
Miliaikeala, S. J., Heen, M. A., Liberman, J. D., and Miethe, T. D. (2014). A comparison of different online sampling approaches for generating national samples. *UNLV Center for Crime and Justice Policy: Research Brief*. Retrieved from https://www.unlv.edu/sites/default/files/page_files/27/ComparisonDifferentOnlineSampling.pdf
Moody, G. D. and Galletta, D. (2015). Lost in cyberspace: The impact of information scent and time constraints on stress, performance, and attitudes online. *Journal of Management Information Systems, 32*(1), 192-224.
Moody, G. D., Galletta, D., and Lowry, P. B. (2016a). It's complicated: Explaining the relationship between trust, distrust, and ambivalence in online transaction relationships using polynomial

regression analysis and response surface analysis. *European Journal of Information Systems, forthcoming*.

Moody, G. D., Slaughter, S. A., Kirsch, L. J., Weng, Q., and Dunn, B. K. (2016b). Facilitating the transformational: An exploration of control in cyberinfrastructure projects. *Information Systems Research, forthcoming*.

Oliver, R. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research, 17*(4), 460-469.

Parks, R., Xu, H., Chu, C.-H., and Lowry, P. B. (2016). Examining the intended and unintended consequences of organisational privacy safeguards enactment in healthcare: A grounded theory investigation. *European Journal of Information Systems, forthcoming*.

Posey, C., Roberts, T. L., and Lowry, P. B. (2015). The impact of organizational commitment on insiders' motivation to protect organizational information assets. *Journal of Management Information Systems, 32*(4), 179-214.

Reips, U.-D. (2002). Standards for Internet-based experimenting. *Experimental Psychology, 49*(4), 243-256.

Rouse, S. V. (2015). A reliability analysis of Mechanical Turk data. *Computers in Human Behavior, 43*(2015), 304-307.

Scott, B. A., Garza, A. S., Conlon, D. E., and Kim, Y. J. (2015). Why do managers act fairly in the first place? A daily investigation of "hot" and "cold" motives and discretion. *Academy of Management Journal, 57*(6), 37-57.

Skitka, L. J. and Sargis, E. G. (2006). The Internet as psychological laboratory. *Annual Review of Psychology, 57*(January), 529-555.

Smith, N. A., Sabat, I. E., Martinez, L. R., Weaver, K., and Xu, S. (2015). A convenient solution: Using MTurk to sample from hard-to-reach populations. *Industrial and Organizational Psychology, 8*(2), 220-228.

Stanton, J. M. (1998). An empirical assessment of data collection using the Internet. *Personnel, 51*(3), 709-726.

Steelman, Z. R., Hammer, B. I., and Limayem, M. (2014). Data collection in the digital age: Innovative alterantives to student samples. *MIS Quarterly, 38*(2), 355-378.

Twyman, N. W., Lowry, P. B., Burgoon, J. K., and Jay F. Nunamaker, J. (2014). Autonomous scientifically controlled screening systems for detecting information purposely concealed by individuals. *Journal of Management Information Systems, 31*(3), 106-137.

Wall, J. D., Lowry, P. B., and Barlow, J. (2016). Organizational violations of externally governed privacy and security rules: Explaining and predicting selective violations under conditions of strain and excess. *Journal of the Association for Information Systems, 17*(1), 39-76.

Wilson, J. (2002). Responsible authorship and peer review. *Science and Engineering Ethics, 8*(2), 155-174.

Woo, S. E., Keith, M., and Thornton, M. A. (2015). Amazon Mechanical Turk for industrial and organizational psychology: Advantages, challenges, and practical recommendations. *Industrial and Organizational Psychology, 8*(2), 171-179.

Zhu, X., Barnes-Farrell, J. L., and Dalal, D. K. (2015). Stop apologizing for your samples, start embracing them. *Industrial and Organizational Psychology, 8*(2), 228-232.