

# The Use of Online Panel Data in Management Research: A Review and Recommendations

Christopher O. L. H. Porter

Ryan Outlaw

Jake P. Gale

*Indiana University*

Thomas S. Cho

*Indiana University–Purdue University Indianapolis*

---

*Management scholars have long depended on convenience samples to conduct research involving human participants. However, the past decade has seen an emergence of a new convenience sample: online panels and online panel participants. The data these participants provide—online panel data (OPD)—has been embraced by many management scholars owing to the numerous benefits it provides over “traditional” convenience samples. Despite those advantages, OPD has not been warmly received by all. Currently, there is a divide in the field over the appropriateness of OPD in management scholarship. Our review takes aim at the divide with the goal of providing a common understanding of OPD and its utility and providing recommendations regarding when and how to use OPD and how and where to publish it. To accomplish these goals, we inventoried and reviewed OPD use across 13 management journals spanning 2006 to 2017. Our search resulted in 804 OPD-based studies across 439 articles. Notably, our search also identified 26 online panel platforms (“brokers”) used to connect researchers with online panel participants. Importantly, we offer specific guidance to authors, reviewers, and editors, having implications for both micro and macro management scholars.*

**Keywords:** *online panel data; research methods; sampling; convenience sample*

---

---

*Acknowledgments:* We thank Brittney Amber and Arielle N. Lewis for their help with this project.

*Supplemental material* for this article is available with the manuscript on the JOM website.

*Corresponding author:* Christopher O. L. H. Porter, Kelley School of Business–Indianapolis, Indiana University, 801 W. Michigan Street, BS 4130, Indianapolis, IN 46202, USA.

*E-mail:* colhp@iu.edu

The availability and use of human research participants is vital to management research. In fact, 56% of the empirical articles (not including meta-analyses) published in the *Journal of Management* in 2017 reported data collected directly from human research participants. Researchers who depend on human participants are familiar with the challenges presented by “traditional” convenience samples. Some of these challenges center on issues of validity. For example, statistical conclusion validity, or the degree to which the relationship between variables is accurately represented, is threatened by small sample sizes and range restriction—both of which are often limitations when research is conducted within a single organization (cf. Cohen, 1992; Shadish, Cook, & Campbell, 2002).

Traditional sampling techniques can also present challenges of a more practical nature. Researchers are often required to exert considerable effort before organizational gatekeepers agree to participate (Clark, 2011; Cunliffe & Alcadipani, 2016). As Tracy (2013: 12) notes, this may involve “countless phone calls, follow-up emails, and ‘courtship rituals’ required in order to gain access.” Even when researchers are granted access, many organizations are reluctant to allow them to collect data on “sensitive” topics, such as racial bias, gender inequality, theft, workplace violence, retaliation, incivility, and abusive supervision.

A relatively recent sampling innovation—online panels—appears to have solved a number of these problems by opening the door to a new convenience sample. An *online panel* (OP) is an electronic database of registrants who have indicated a willingness to participate in future web-based research studies (Callegaro, Baker, Bethlehem, Göritz, Krosnick, & Lavrakas, 2014). Two related terms we will refer to are *online panel data* (OPD)—the data derived from an OP—and *online panel platform* (OPP)—the host that provides access to the OP.

The earliest use of an OP in academic journals appears to be the late 1990s (e.g., Li, Kuo, & Russell, 1999, as noted by Göritz, 2007). Since then, the number of OPPs and the use of OPD have steadily increased—a trend reflecting our belief that OPD is one of the most significant sampling developments in modern science. The field of management has not been immune to these trends. As the data we collected for our review indicated, OPD appeared in 6.6%—on average—of the empirical articles in management in just over the last decade. By comparison, OPD appeared in 14.3% of the empirical articles in 2017 (the last year included in our review). But what exactly is OPD? What can we learn from how management scholars have used OPD? What issues, if any, remain unresolved about how to best use OPD? Importantly, what questions should management scholars, including both researchers and evaluators, consider when deciding the degree to which OPD is appropriate? Our review addresses each of these issues. But before delving into them, we first introduce OPD and explain its rising popularity.

## A Brief Introduction to OPD

While our review includes 26 unique OPPs, we begin with, for illustrative purposes, a well-known example—Amazon’s Mechanical Turk (MTurk). Note that a detailed guide on how to set up and administer a survey on MTurk or similar OPPs is beyond the scope of this review (we refer interested readers to Mason and Suri, 2012, as well as Chandler, Mueller, and Paolacci, 2014). MTurk, which launched in 2005, was initially designed to provide “requestors” access to “workers” who were willing to complete simple microtasks (Aguinis & Lawal, 2013) too complicated for computers to perform. Over the past decade or so, those microtasks grew to include participation in scientific surveys and experiments as more

academic researchers took on the role of requestors. Today, OPPs like MTurk, Qualtrics, and StudyResponse provide researchers access to participants from a global online marketplace and are gaining popularity because of several advantages OPPs offer over traditional convenience samples.

One advantage is that OPPs provide researchers with a convenient way to reach a potentially unlimited number of participants while keeping costs to a minimum (Buhrmester, Kwang, & Gosling, 2011). These qualities have made OPPs ideal for scale development or pilot studies where multiple iterations may be required. Additionally, OPPs give researchers access to sample participants from across the globe, facilitating increasingly representative samples (Gleibs, 2017; Goodman & Paolacci, 2017). Conversely, scholars interested in studying specific yet hard-to-reach segments of a population—for example, members of the LGBTQ+ community—can also do so using OPPs and relatively little effort (Smith, Sabat, Martinez, Weaver, & Xu, 2015). The increased anonymity OPPs offer also makes them ideal for researchers to collect data on topics participants might be reluctant to report or admit experiencing (Smith et al., 2015). Some OPPs facilitate intensive research designs, such as those that require temporal separation (e.g., multiwave field studies or experience-sampling methodology), with acceptable retention rates (Chandler et al., 2014). Finally, many institutional review boards consider OPD-based studies “exempt,” potentially saving researchers valuable time (Paolacci, Chandler, & Ipeirotis, 2010).

Of course, OPPs are not without controversy. There are three issues that have received a great deal of attention and deserve mention here given their relevance to management research. First is the issue of non-naïve participants, or “professional survey takers”—participants who frequently engage in surveys and experiments. The concern is that non-naïve participants may systematically respond to surveys and experiments differently than those who rarely take part in research. Evidence indicates that both crosstalk between participants and respondents intentionally attempting to participate more than once in the same study are virtually nonexistent (Chandler et al., 2014). Participant experience may be an issue in terms of attenuating effects sizes (Chandler, Paolacci, Peer, Mueller, & Ratliff, 2015), but this appears to primarily impact researchers employing experiments with common, widely known paradigms (Chandler et al., 2014). Experienced participants are likely less of a problem for researchers conducting novel experiments or survey research; however, more research is needed on this (Cheung, Burns, Sinclair, & Sliter, 2017).

Second, the representativeness of OP participants has been called into question. Yet, there is overwhelming evidence that OPPs are more representative of typical working adults than traditional student samples (Crone & Williams, 2017; Goodman & Paolacci, 2017; Peer, Brandimarte, Samat, & Acquisti, 2017). Moreover, there is evidence that OPD is similar to data collected using traditional samples. In a recent meta-analysis, Walter, Seibert, Goering, and O’Boyle (2018) compared effect sizes of organizational variables collected using OPD to “conventionally sourced” data. The authors concluded that these two approaches yield substantively similar effect sizes, which in turn provides greater confidence in both approaches (Walter et al., 2018).

Third, fears over subpar data quality due to inattentiveness or lack of effort have been expressed (e.g., Chandler et al., 2014), but those fears have largely been refuted. There is evidence that the attention levels of, and psychometric data from, OP participants meet or exceed those from traditional data sources (e.g., Behrend, Sharek, Meade, & Wiebe, 2011;

Buhrmester et al., 2011; Crone & Williams, 2017; Goodman & Paolacci, 2017; Hauser & Schwarz, 2016; Paolacci et al., 2010; Ramsey, Thompson, McKenzie, & Rosenbaum, 2016). Additional evidence that OPD is capable of yielding high-quality data comes from Walter et al.'s (2018) meta-analysis based on 32,121 OP participants across 90 independent samples. Results from their reliability generalization analysis indicated that OP participants provided data that were comparable to conventionally sourced data in terms of psychometric soundness.

## **OPD and Management**

Despite the aforementioned evidence generally supporting the validity of OPD, there remains a deep divide among management scholars over its appropriateness. Evidence of this disagreement can be seen in journals that refuse to publish OPD-based research (Landers & Behrend, 2015) and editorial board members and reviewers who automatically reject such work (for an example, see Walter et al., 2018). This divide is problematic for several reasons, perhaps the most serious of which is the confusion and uncertainty it causes, impeding the ability of our field to mature. For example, the attitudes of editors toward OPD impact the degree to which a journal is seen as a viable outlet for OPD-based research. And for their part, individual reviewers' views on work using OPD may come down to the luck of the draw.

This problematic divide may hit authors the hardest. After all, it is authors who must wrestle with the "Is OPD appropriate?" question throughout the publication process, starting with research design. Authors may question whether editors and reviewers are likely to give a longer leash to OPD use for certain topics or hard-to-reach subpopulations. For example, "Is it acceptable to use OPD to study sexual orientation and potential stigma at work?" Similarly, authors may wonder if OPD is tolerable for certain types of research. For instance, "Is it okay to use OPD for substantive hypothesis testing or am I better off using it only for scale validation?" Related, authors can be left to guess about potential outlets for work containing OPD. Imagine how many authors have asked the question, "Do I even have a chance of publishing this research in a particular journal if I use OPD?"

Management scholars deserve answers to these elusive questions, and our review represents an important step in providing answers. Specifically, our review allows us to offer pointed guidance regarding *when* and *how* OPD should be used by management researchers. That guidance identifies missed opportunities and critical considerations based on a close look at how OPD was employed during approximately its first decade of use by management scholars. Because our review suggests that OPD is an innovation that is likely here to stay, we also provide a comprehensive set of best practices for management scholars as they continue to use OPD in the future. While we are not the first to suggest best practices as it relates to the use of OPD, we uniquely identify areas of (dis)agreement across scholars' recommendations for executing OPD studies. The result highlights the complexities researchers and evaluators must consider as they conduct and evaluate OPD research and should serve as an invaluable resource for making informed decisions about this research.

## **Method**

The first step in conducting our review was identifying journals for inclusion. We began with journal lists from the University of Texas at Dallas Top 100 Business School Research

Rankings (2018) and the Texas A&M/University of Georgia Productivity Rankings (2018). To be comprehensive and given our interest in examining OPD use across a broad range of management topics, we included not only those that covered more micro areas (e.g., *Journal of Applied Psychology*, *Organizational Behavior and Human Decision Processes*) but also those that typically cover mostly macro areas (e.g., *Strategic Management Journal*).

The next step in building our data set was excluding journals that published only theoretical or conceptual articles (i.e., *Academy of Management Review*). Finally, we included several journals that, though not on the aforementioned journal lists, are widely known and sought-after targets for management scholars. This provided the additional benefit of broadening the quality and scope of the work included in our review.<sup>1</sup> These additions largely included specialty journals (e.g., *Strategic Entrepreneurship Journal* and *Leadership Quarterly*). The result was the following 13 journals: *Academy of Management Journal*, *Administrative Science Quarterly*, *Journal of Applied Psychology*, *Journal of International Business Studies*, *Journal of Management*, *Journal of Organizational Behavior*, *Leadership Quarterly*, *Management Science*, *Organizational Behavior and Human Decision Processes*, *Organization Science*, *Personnel Psychology*, *Strategic Entrepreneurship Journal*, and *Strategic Management Journal*.

To ensure we captured all the published articles that used OPD in these journals, we conducted a manual search beginning with 2005—the year MTurk was launched. Although a query-based search (e.g., conducting an electronic, online search for articles that mention the word *MTurk* or *Qualtrics*) would have been faster, conducting a manual search was important for several reasons. First, early in our literature search, we saw evidence that authors were sometimes less than transparent about the source of their data. For example, some references to OPD and OPPs were embedded in footnotes and appendices rather than explicitly identified in Method sections. Second, it was not possible to identify a comprehensive list of the various OPPs to include in a query-based search. The only way we could be confident that we identified a comprehensive list of OPPs was to manually read the Method sections, footnotes, and appendices of every empirical article.

Once our manual search was completed and we had a list of OPPs, we took steps to ensure there were no omissions by conducting a query-based search. In addition to including the list of OPPs generated from our manual search, our query-based search included the terms *online labor market*, *online data*, *online panel*, and *panel data*. We then used Boolean operators to search both ABI EBSCO and Google Scholar databases for the years 2005 through 2017. Our efforts resulted in our identifying 804 studies in 439 articles published between 2006 and 2017.<sup>2</sup> Despite our deliberate starting point, we checked and confirmed that there were no management articles published in 2005 that used OPD. Table 1 presents both the number of studies (*n*) and articles (*k*) using OPD by journal.

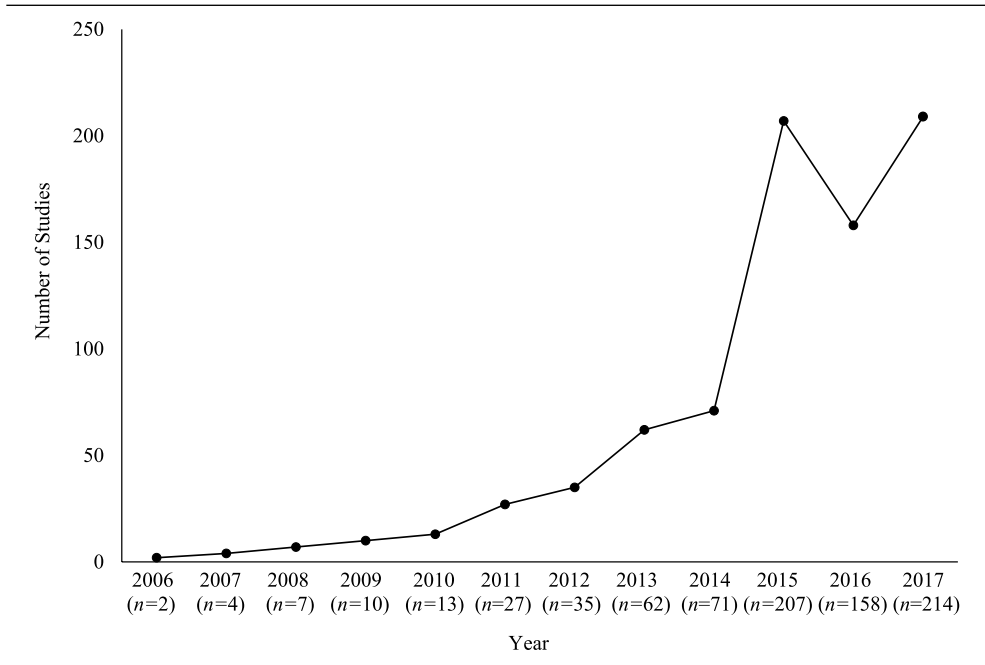
Prior to coding any of the studies and articles, we met as a team to establish the coding criteria, agree upon best practices, and collectively code a subset of studies (*n* = 80) to ensure our independent coding would be consistent. Each author was then assigned three or four journals to code independently. In addition to coding the journal in which the articles and studies appeared, we coded the OPP taking into account whether the OPP was public and openly available to researchers or private, providing access limited to a select few researchers. We also coded the nature of the primary research question. We identified whether OPD was used to address substantive (e.g., hypothesis significance testing for main hypotheses), substantive pilot (e.g., whether an experiment evoked the desired effect), or measurement

**Table 1**  
**Journal by Study Count, Article Count, Question Type, Method Type, and Design Element, 2006–2017**

Journal	Study Count ( <i>n</i> )	Article Count ( <i>k</i> )	Question Type ( <i>k</i> )			Method Type ( <i>k</i> )			Design Element ( <i>k</i> )	
			Measurement	Substantive	Substantive Pilot	Correlational	Experimental	Inductive	Source Separated	Time Separated
<i>AMJ</i>	53	32	4	25	3	13	17	2	2	3
<i>ASQ</i>	12	9	3	6	0	7	2	0	0	1
<i>JAP</i>	131	93	25	62	6	69	23	1	5	25
<i>JIBS</i>	11	8	4	4	0	7	1	0	0	0
<i>JOM</i>	21	17	9	8	0	14	3	0	2	6
<i>JOB</i>	37	32	15	17	0	26	6	0	3	10
<i>LQ</i>	49	32	2	27	3	19	13	0	4	1
<i>MS</i>	61	26	11	12	3	8	18	0	0	1
<i>OBHDP</i>	362	138	13	113	12	40	97	1	2	6
<i>OS</i>	35	24	4	20	0	10	14	0	0	3
<i>PP</i>	25	22	8	13	1	19	3	0	3	6
<i>SEJ</i>	1	1	1	0	0	0	1	0	0	0
<i>SMJ</i>	6	5	0	5	0	2	3	0	1	0
Total	804	439	99	312	28	234	201	4	22	62

*Note:* *n* = number of studies; *k* = number of articles; *AMJ* = *Academy of Management Journal*; *ASQ* = *Administrative Science Quarterly*; *JAP* = *Journal of Applied Psychology*; *JIBS* = *Journal of International Business Studies*; *JOM* = *Journal of Management*; *JOB* = *Journal of Organizational Behavior*; *LQ* = *Leadership Quarterly*; *MS* = *Management Science*; *OBHDP* = *Organizational Behavior and Human Decision Processes*; *OS* = *Organization Science*; *PP* = *Personnel Psychology*; *SEJ* = *Strategic Entrepreneurship Journal*; *SMJ* = *Strategic Management Journal*.

**Figure 1**  
**Online Panel Data Study Count by Year**



(e.g., scale development) questions. We also coded for method type (i.e., correlational, experimental, or inductive) and design elements (i.e., time and source separation). Finally, we coded the primary topic (e.g., leadership or creativity) for each study.<sup>3</sup> Upon completion of each author's independent coding, the team met again to reach agreement where uncertainty was present.

### Findings and Recommendations: Looking Back to Look Ahead

Figure 1 graphically depicts the number of articles published in management journals from 2006 to 2017. Since the first study published in 2006, there has been a fairly steady increase in the publication of OPD-based research. Major shifts along the way included 2010 ( $n = 13$ ) to 2011 ( $n = 27$ ) as well as 2012 ( $n = 35$ ) to 2013 ( $n = 62$ ), representing increases in OPD-based studies of 107.7% and 77.1%, respectively. However, the biggest increase came between 2014 ( $n = 71$ ) and 2015 ( $n = 207$ )—an increase of 191.5% in OPD-based studies. Coupled with the sheer number of articles identified in our review, these findings lend credence to our belief that the field seems beyond the question of whether, at a general level, OPD is appropriate; the acceptance and integration process by management scholars has begun.

These observations reinforce our belief that the time is appropriate for reviewing how OPD has been used in the management literature. These observations also demonstrate the urgency for an informed dialog about how OPD should, and could, best be utilized in

future management research. Now is the time for the field to take a stance and adopt a common language. Accordingly, we develop a set of guidelines for management scholars aimed at using OPD if appropriate, choosing an OPP, reporting the use of OPD, and publishing OPD studies. We refer to these four guidelines collectively as *using, choosing, reporting, and publishing*.

### *On Using OPD*

One of the first, and arguably most important, issues scholars must address is whether OPD is appropriate for answering their research questions. Decisions about appropriateness should be determined primarily based on the (a) topic and (b) nature of the question being addressed. For example, a researcher examining the effects of witnessing abusive supervision—a topic some organizations may not want to acknowledge or address—might be well justified in using OPD. This justification would be especially true if OP participant anonymity reduces fears of retaliation or breaches of confidentiality that might otherwise undermine data collection from traditional convenience samples.

Our data demonstrate that there have been few, if any, topics management researchers have not explored using OPD. Topping the list was leadership ( $k = 49$ ), decision making ( $k = 46$ ), and ethics and morality ( $k = 36$ ), representing 11.2%, 10.5%, and 8.2% of the articles in our data, respectively.<sup>4</sup> Notably, OPD was used extensively to investigate potentially dark and sensitive management topics, such as ethical and moral behavior, abusive supervision, and fairness. The sensitive nature of these topics may, in part, explain the frequency with which they have been explored with OPD. “Conventionally sourced” employees might be hesitant to provide candid, honest responses about these topics, and organizations may have reservations about allowing researchers to collect data on these topics. That said, there is no reason—nor is there evidence to suggest—that OPD should be limited to certain topics.

Among the topics that have not been widely studied with OPD by management scholars, several are noteworthy. In a rare recruitment study, Phillips, Gully, McCarthy, Castellano, and Kim (2014) presented participants with recruiting messages that varied in terms of their reference to the hiring organization’s global presence and travel requirements. The authors wanted to understand the extent to which those messages interacted with participants’ global mind-sets to ultimately influence job pursuit intentions. Indeed, OPD seems capable of facilitating research on recruitment, selection, retirement, turnover, and other processes that occur during or near transitions into, between, or out of traditional jobs. OPD has not been used extensively to explore these sorts of topics, which we found ironic given the possibility that OP participants might be engaged in such e-work while experiencing such transitions. Given what appears to be an increasing acceptance of OPD, it might be only a matter of time before the field observes an increase in the use of OPD across a broader range of topics.

An example of a topic that perhaps could be explored differently in future work using OPD is groups and teams. Although scholars have conducted studies with OP participants who were part of fictitious teams or who were led to believe they were making decisions with others (e.g., Swaab, Phillips, & Schaerer, 2016), the challenges to recruiting real working groups and teams into OPs are obvious. Having said that, we do not see this as beyond the realm of possibility and could envision this being a future reality. One way a researcher might accomplish this would be by building a private panel using participants who were organized in teams and whom the researcher has previously encountered, perhaps in a more traditional



research context. If some meaningful subset of the team is still intact, working interdependently, and willing to participate in future research, these participants could provide useful data obtained in the same fashion as OPD has been obtained.

Although the aforementioned examples are ones likely to be explored by micro and meso scholars, our data also suggest that an excellent opportunity for OPD in future management scholarship is its broader use by macro scholars. Indeed, there is overlap across many of the topics of interest to both micro and macro scholars (e.g., decision making, leadership).

As evidence of the viability of using OPD in macro research, Crilly, Ni, and Jiang (2016) conducted an experiment that replicated their findings from a field study and extended those findings by testing an implied causal mechanism. Specifically, they examined the effects of a firm's type of corporate social responsibility (CSR) activity and foreignness on attributions about why those firms engaged in socially responsible activities. The authors also evaluated the degree to which type of CSR, foreignness, and causal attributions drove overall impressions of the firm.

Turning next to the nature of the research question being asked, we examined the extent to which OPD has been used to address measurement, substantive, and substantive pilot questions (see again Table 1 and also Table 2). Table 1 presents the results of our coding by journal, while Table 2 presents similar information but focuses on  $n$  instead of  $k$  as we found  $n$  more informative for comparisons across OPPs. Together they shed light on how OPD has been used, from whom it has been collected, and where it has been published by management scholars (readers interested in a detailed look at OPD use by topic should refer to Online Supplemental Materials B). As seen in the tables, OPD was used quite extensively to test substantive research questions (e.g., null-hypothesis significance testing). Specifically, 634 (or 78.9%) of the studies in our review tested substantive research questions, which we distinguished from substantive pilot studies ( $n = 46$ , or 5.7%).

Our data further indicate that management scholars used a range of different methodologies when conducting OPD-based research (i.e.,  $n = 477$ , or 59.3%, for experiments and  $n = 321$ , or 39.9%, for correlational research). These findings demonstrate the broad utility of OPD. They also suggest that OPD has been used by management scholars with different backgrounds, training, and expertise. Moreover, these findings suggest that OPD might be especially relevant to a broader set of researchers, including those who have not relied on OPD, including, again, those studying traditional macro topics.

In fact, macro scholars may find that OPD proves superior to traditional samples in some cases. For example, Wowak, Mannor, Arrfelt, and McNamara (2016) had undergraduates code CEO dossiers. Recall that the impetus behind the creation of OPPs was to outsource tasks too difficult for computers—like coding—to e-workers. Perhaps Wowak et al. (2016) could have had OP participants, particularly those with experience working in organizations with CEOs, do that same coding. By way of another example, consider that OPPs operate and exist all over the world. Therefore, studies that require an international sample of working adults, such as the study conducted by Chua, Morris, and Ingram (2009) in which they examined trust in Chinese and American managers' professional networks, may also be good candidates for OPD.

As macro scholars increasingly rely on experiments and other research features historically associated with micro research, they might use OPD to develop and pilot test scales, pretest experimental manipulations, and conduct other research that typically precedes traditional field tests. For instance, Shapira and Shaver (2014) used four waves of MBA students

Table 2  
Online Panel Platform (OPP) by Study Count, Article Count, Question Type, Method Type, and Design Element, 2006–2017

OPP	Study Count ( <i>n</i> )	Article Count ( <i>k</i> )	Question Type ( <i>n</i> )				Method Type ( <i>n</i> )			Design Element ( <i>n</i> )	
			Measurement	Substantive	Substantive Pilot	Correlational	Experimental	Inductive	Source Separated	Time Separated	
MTurk	529	254	81	409	39	149	375	5	1	20	
Qualtrics	44	32	10	34	0	26	17	1	0	11	
StudyResponse	67	58	15	51	1	63	4	0	17	27	
Zoomerang	10	10	4	6	0	8	2	0	1	4	
Other public	52	27	4	47	1	28	24	0	5	6	
Other private	5	1	1	4	0	2	3	0	0	1	
Unspecified	97	57	9	83	5	45	52	0	3	7	
Total	804	439	124	634	46	321	477	6	27	76	

Note: *n* = number of studies; *k* = number of articles.

to pilot test decision-making experiments they later ran with more MBA students. The authors suggested many of their MBA students were also working professionals—a criterion many OPPs allow researchers to include in their selection process.

In these examples, nothing about the pilot or primary study samples precluded the use of OP participants. Moreover, using nonstudents could eliminate or reduce concerns about a potential lack of voluntary participation and coercion. However, we recognize that MBAs may have more direct contact with researchers and the opportunity that contact affords for personal reminders and strengthening personal connections relative to OP participants. Thus, student samples may have an advantage in terms of increased participation and response rates.<sup>5</sup>

To reiterate, we are not suggesting that there was anything inherently wrong with the participants used in any of the aforementioned examples. We are merely emphasizing that OP participants may have been equally appropriate based on our current understanding of OPD validity and its increasing acceptance in the field. Given the commonalities in the work being done by micro and macro scholars alike, it is difficult to think of reasons why macro scholars would not increasingly use OPD. Going forward, we recommend that all scholars in the field at least consider the potential advantages of OPD. The trends we observed in our data (see again Figure 1) lead us to expect a significant increase in the use of OPD “across the board.”

We should also note one other methodological observation based on our review—one that suggests another important consideration and recommendation for management scholars. We found that only 0.7% of OPD studies used inductive methods ( $n = 6$ ). This suggests missed opportunities for management scholars wishing to use richer (e.g., interviews) or more powerful (e.g., longitudinal studies that can shed light on causal processes) designs. Given the capabilities of some of the OPPs included in our data (see again Table 2), we recommend that when possible, management scholars exploit OPPs’ capabilities to accommodate such designs.

In sum, the use of OPD must, first and foremost, be driven by the research question or questions. Although the topic area and the nature of scholars’ research questions are the key determinants of the extent to which OPD is appropriate, we propose four secondary research-driven considerations. Researchers and evaluators should also keep these considerations or “decision points” in mind when deciding whether OPD is appropriate regardless of the type of question being asked (i.e., measurement or substantive). Among the decision points critical enough to be labeled secondary considerations are the representativeness of the participants OPs can offer; the extent to which OPs can offer participants with the necessary knowledge, skills, and abilities (KSAs) required for study participation; the extent to which the study can be influenced by practice effects; and whether OPs can support the study’s technological requirements. We discuss each of these issues below.

*Representativeness.* OPPs and third-party applications have proven to be a tremendous resource for scholars requiring access to specific populations. Tools such as TurkPrime now offer the ability to select samples based on unique participant qualifications ranging from medical conditions to dietary habits, and the list continues to expand. For qualifications not yet available via these services, researchers can administer discreet prescreening surveys (cf. Chandler et al., 2014). Of course, there are certain populations that would be unrealistic to access via OPPs. For example, if research involves studying perceptions of Fortune 100 CEOs, then OPD will probably be inappropriate as these top-level executives are unlikely to belong to most OPs (Stritch, Pedersen, & Taggart, 2017).

*KSAs.* Management research typically requires participants to possess basic KSAs to complete a research task. Indeed, OPPs like MTurk were designed for these types of tasks, making them well suited for many management studies. For example, Tosti-Kharas and Conley (2016) asked OP participants to read a passage and rate that passage for constructs such as emotional tone. This type of study approximates a typical OP study that can be completed by participants with little or no training (Brawley & Pury, 2016). However, there are instances where a lack of knowledge on the part of the participant may serve as a source of error (Fowler, 2014). For example, a study might require the use of expert raters, such as participants who have spent years studying a subject area, to perform a task. In these instances, OPD may be inappropriate.

*Practice effects.* When assessing the appropriateness of OP participants, evidence suggests study experience may be less important for studies involving unique instruments or manipulations (Chandler et al., 2014). Similarly, experience may be less important for studies involving perceptual data, such as personality measures (DeVoe & House, 2016; Miller, Crowe, Weiss, Maples-Keller, & Lynam, 2017). That said, there is evidence that more experienced participants have likely seen and respond differently to studies involving common manipulations or cognitive tests (Chandler et al., 2014, 2015; DeVoe & House, 2016). Therefore, OPD would likely be inappropriate—especially in the case of more experienced participants—for research that involves common manipulations or cognitive measures that cannot be made novel (Paolacci & Chandler, 2014).

*Technological requirements.* As access to technology becomes cheaper and more accessible, researchers have grown increasingly sophisticated with the type of research they can perform remotely. Indeed, Chandler and Shapiro (2016) pointed out that OPPs like MTurk can accommodate technology that requires measuring momentary reaction times, such as Stroop tests (Crump, McDonnell, & Gureckis, 2013) and implicit association tests (Klein et al., 2014). With that said, there are limits in terms of technology that can reasonably be accommodated with OPD. For example, management researchers who use functional magnetic resonance imaging (fMRI) could not reasonably expect to incorporate such a measurement tool in an OP study based on today's technology.

We would be remiss if we did not make two special notes. First, we cannot overemphasize how much we discourage scholars from using OPD solely for the sake of convenience. Convenience—in terms of both speed and cost of data collection—may be the single greatest advantage of OPD. However, as Goodman and Paolacci (2017) warned, that advantage could pose serious threats. If left unchecked, the convenience factor of OPD could inadvertently drive research agendas, resulting in research questions being tossed aside or modified so that they are “OPD-friendly.” We share the concerns raised by those authors and suggest that the research question itself should dictate whether OPD is appropriate—not the other way around.

Second, it is worth noting that the OP landscape is constantly changing such that what seems unrealistic today may very well be a reality tomorrow. To illustrate, consider that Buhrmester et al. (2011) recently suggested physiological measurements with OPD would be “impossible,” yet researchers have already begun using OP participants for studies involving remote eye tracking, facial expressions, and heart rate monitoring (Chandler & Shapiro,

**Table 3**  
**How Research Determines Appropriateness of Online Panel Data (OPD):**  
**Secondary Considerations**

Given Research Question	When OPD Would Have Likely Been (In)appropriate	Justification for (In)appropriateness
Researcher determines whether OPs can provide a representative sample of the target population.	Appropriate: Investigate pregnant women and work (e.g., Little, Major, Hinojosa, & Nelson, 2015). Inappropriate: Explore CEO perceptions of top management team influence (e.g., Elenkov, Judge, & Wright, 2005).	Appropriate: Researchers can select specific attributes to recruit hard-to-reach samples. Inappropriate: Some subpopulations are unlikely to enroll as OP participants.
Researcher determines whether OP participants have KSAs required for meaningful results.	Appropriate: Conduct a content validation study using item-to-definition matching (e.g., Baer, Bundy, Garud, & Kim, 2018). Inappropriate: Code using expert raters (Gamache, McNamara, Mannor, & Johnson, 2015).	Appropriate: Judging whether an item seems to match a definition requires basic verbal ability. Inappropriate: Some tasks require years of experience or highly specific expertise.
Researcher determines whether measures or manipulations are subject to practice effects.	Appropriate: Collect attitude/behavior variables (Boswell, Olson-Buchanan, & Harris, 2014). Inappropriate: Run iterations of the prisoner's dilemma (e.g., Insko, Wildschut, & Cohen, 2013).	Appropriate: Repeated exposure to variables such as personality traits is unlikely to bias results. Inappropriate: Some tasks, manipulations, scenarios, etc. are too common, available, or well-known.
Researcher determines whether OPs can support technological requirements.	Appropriate: Study negotiations involving groups interacting in real time (Yamagishi et al., 2013). Inappropriate: Monitor blood pressure to measure stress (Bono, Glomb, Shen, Kim, & Koch, 2013).	Appropriate: OPs are increasingly able to accommodate more sophisticated study tools. Inappropriate: Some studies require a level of sophistication beyond what OPs can handle.

*Note:* OP = online panel; KSAs = knowledge, skills, and abilities.

2016; Goodman & Paolacci, 2017). Just imagine if OPPs began to specialize in recruiting CEOs from Fortune 100 companies or if technology made it feasible to capture fMRI-type data from OP participants. While those changes may seem a bit of a stretch, we were shocked to discover the number of advancements that have taken place in just the past decade. For that reason, we encourage researchers to constantly be aware of changes that may impact how the research question determines the appropriateness of OPD. Table 3 summarizes our discussion of these secondary considerations and provides current examples of when OPD would and would not be appropriate. Therefore, when coupled with advice about first considering the research topic and the nature of the research question, Table 3 serves as an additional guide for scholars.

### *On Choosing an OPP*

When many scholars think "OPP," they think "MTurk." In fact, MTurk is often used synonymously with OPD. Our review indicates that although MTurk was clearly the most often

used OPP, assuming that an OPD study is an MTurk study is a mistake. Management researchers used as many as 26 different OPPs from 2006 to 2017. These OPPs included MTurk ( $n = 531$ , or 65.8%), StudyResponse ( $n = 67$ , or 8.3%), Qualtrics ( $n = 45$ , or 5.6%), and Zoomerang ( $n = 10$ , or 1.2%). Collectively, those four OPPs appear to be where most (80.9%) of the OPD used by management researchers was derived, as seen in Table 2. To facilitate the interpretation of our results, we labeled the remaining 22 identifiable OPPs as either *other public* ( $n = 52$ , or 6.5%) or *other private* ( $n = 5$ , or 0.6%).<sup>6</sup> Notably, the OPP was unspecified in 101 (12.1%) of our studies. More worrisome, there were multiple unspecified-OPP articles each year from 2010 to 2017, demonstrating a consistent lack of consensus on how to report OPD—a point we cover later.

With so many OPPs to choose from, researchers may be left wondering, “Which OPP is best for me?” Similarly, editors and reviewers may wonder, “Which OPP should researchers use?” Consistent with our previous recommendation that the decision about whether to use OPD should be research driven, we recommend that decisions about which OPP one uses be based on the OPPs’ fit with the research agenda rather than generalizations about those OPPs or their popularity.

Take, for example, a researcher seeking to collect source-separated data. Source separation—a technique used to mitigate common method bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003)—occurs when data are collected from two or more sources. It typically requires a researcher to obtain contact information from participants, which some OPPs do not allow. As seen in Table 2, no MTurk studies used source separation because the OPP has no mechanisms for, and discourages, it (Miller et al., 2017). In contrast, 63.0% of all source-separated studies were conducted using data from StudyResponse. As an aside, it is noteworthy that we found little evidence of the use of either source- or time-separated design elements in our data.

We also recommend that if issues such as the ability to collect source- or time-separated data indeed drive decisions about the OPPs researchers choose, researchers clearly report the impetus behind their decision making. As OPPs continue to evolve, information like this will facilitate research and advance the field. For example, other researchers may better target their own data collection efforts based on that information. Similarly, researchers’ choices might spur competitiveness among OPPs, increasing the quality and range of the services they offer. Moreover, this type of reporting also forces researchers to think beyond surface-level criteria (e.g., cost or ease) when selecting an OPP.

Scholars’ decisions about the appropriateness of OPPs may be due to real or imagined differences about various OPPs’ capabilities handling rigorous research designs. We recognize that the inputs for such decisions are ever changing as OPPs appear, dissolve, and evolve. It follows that scholars’ understanding about what different OPPs can offer must evolve as well. As an example, a research team familiar and comfortable only with MTurk might “choose” to conduct a study utilizing time separation rather than source separation. In doing so, this team has potentially missed an opportunity to utilize a design element that might be more appropriate for the research question. In that vein, that same research team might not even be aware of changes in MTurk’s capabilities since the last time it used MTurk. Regardless, research questions and design elements should drive the choice about OPPs; OPPs should not drive research questions and design elements.

As another example of how the choice of OPP could and should be research driven, it might be that an OPP is chosen because of the type of participants the OPP makes available.

There is preliminary evidence that OPPs vary in their demographic diversity, with MTurk being recognized as particularly diverse relative to other OPPs (Buhrmester et al., 2011; Keith & Harms, 2016). Researchers requiring a diverse sample of participants might therefore choose MTurk or some similar OPP over an alternative OPP. For example, a private OPP that limits its enrollment as participants to students and alumni (e.g., Yale's eLab) might not be able to yield the sort of diversity a researcher needs to explore a particular research question.

Finally, if there is concern that OPP choice could influence results, there is value in using multiple, independent OPP samples (Peterson & Merunka, 2014). While this decision must be approached cautiously—a point we will explain later—it might help assuage concerns about generalizability. Surprisingly, we found little evidence in our data that researchers took advantage of multi-OPP samples (for an exception, see Mochon & Frederick, 2013).

### *On Reporting the Use of OPD*

A holistic, consistent, and transparent approach to reporting basic OPD-related information is critical if we, as a field, are to move beyond unsubstantiated objections to OPD. What might that basic information entail? As a starting point, scholars should report all of the data necessary for future, secondary analyses (e.g., meta-analyses) of their findings. Beyond sample demographic data, researchers should also report means, standard deviations, and effect sizes (for a current review of best reporting practices, see Appelbaum, Cooper, Kline, Mayo-Wilson, Nezu, & Rao, 2018). One potential moderator that might be important for secondary analyses that is unique to OPD research is the specific OPP used. While transparency about the OPP used might seem intuitive, recall that the OPP was unidentified in over 12% of the studies in our review. As research utilizing OPD amasses, comparisons of, for example, effects across OPPs will be facilitated to the extent that researchers report such information.

While there may be utility in combining samples (e.g., a multi-OPP sample) in a single study, we encourage scholars to take special care in reporting such results. Specifically, we recommend that researchers demonstrate and report the appropriateness of combining data from different sources, including different OPPs or OPPs and traditional samples (e.g., Rouse, 2015). Moreover, researchers should ensure that sufficient information is provided to allow interested scholars to understand each individual sample. To illustrate a case of inadequate multisample reporting, the authors of one study in our review combined an OP sample with a traditional sample and reported only an overall, aggregate sample size. This lack of detail makes it impossible to determine to what extent the final sample was composed of OP participants compared to traditional participants. To make matters worse, the authors provided no justification for their decision to combine the samples.

It may seem obvious that authors should be transparent in reporting the fact that they used OPD at all. However, our data reveal cases where information was so ambiguous that it was nearly impossible to determine whether an OP sample had been used. Thus, at a minimum, authors must clearly report that OPD has been used. We also raise this point for a second reason. Recall that we conducted a manual search for this review, in part, because OPD use was occasionally reported in footnotes and appendices, even for primary studies. In the interest of transparency, such information belongs “front and center” in Method sections, and we recommend that authors and evaluators insist on this in future work. Simply put, scholars reporting and evaluating research that includes OPD should expect and demand the same

degree of transparency required when using traditional convenience samples. Perhaps efforts to avoid drawing attention to OPD use and lack of transparency might be attributed to its novelty over the past decade. However, there is no reason for management researchers to be anything less than up-front about OPD use going forward.

Finally, researchers should report study incentives. Although recent work has attempted to highlight the ethical concerns surrounding OP participant compensation (Crone & Williams, 2017; Goodman & Paolacci, 2017), our review suggests the ongoing relevance of this concern in management research. We adopt the view of Aguinis and Lawal (2013), who view OP participants as e-workers; thus, participants should be appropriately compensated for their work. That said, 47% of studies in our data set failed to report any form of payment for OP participants. Additionally, we identified extreme pay discrepancies in which OP participants were paid less than \$1 per hour (readers might compare this to the current federal minimum wage which is \$7.25 per hour). We also identified instances where OP participants were paid significantly less than undergraduates for identical work.

More problematic, and to our point about transparency in reporting, only 45 studies reported both the time required to participate and compensation, making it virtually impossible to determine the extent to which participants were paid equitably. Justice and equity are regularly evoked constructs in management. As such, we could not help but notice and admonish the irony in some researchers' "do as I say, not as I do" behavior. Our hope is that clear and consistent reporting of participant payment and time requirements will help hold researchers accountable for fair and equitable treatment of OP participants.

### *On Publishing OPD Studies*

As we demonstrated, OPD-based research is being published across a broad array of management journals. However, Table 1 only tells part of the story. To develop recommendations about publishing OPD studies, we found it important to go back to the beginning. The first published management study using OPD was published in *Academy of Management Journal* (i.e., Piccolo & Colquitt, 2006) using data obtained from StudyResponse. Later that year, a second study (i.e., Judge, Ilies, & Scott, 2006) appeared in *Personnel Psychology*. Interestingly, although both *Academy of Management Journal* and *Personnel Psychology* were early adopters, neither published OPD-based research for the next several years. Meanwhile, other journals, such as *Organizational Behavior and Human Decision Processes* and *Journal of Applied Psychology*, began to publish OPD-based research regularly.

By 2011, 8 of the 13 journals included in our review had published studies utilizing OPD, the exceptions being *Journal of International Business Studies*, *Journal of Organizational Behavior*, *Organization Science*, *Strategic Entrepreneurship Journal*, and *Strategic Management Journal*. The next milestone was 2016, the year every journal in our review had published OPD-based research and a remarkable observation for two reasons. First, this finding signaled that OPD was no longer limited to any realm of management scholarship. Second, this finding confirmed that OPD can be used and published by both macro and micro scholars.

Thus, our data suggest that scholars have not limited their potential publication outlets, at least among outlets included in our review. We have little reason to expect that, going forward, scholars will limit their potential publication outlets unless specifically advised by



editorial teams that their journal will not publish OPD-based research (we revisit this below). Having said that, we acknowledge that almost all OPD-based studies in our review were published in either micro (e.g., *Journal of Applied Psychology*) or mixed research (i.e., micro and macro “big tent” journals, such as *Journal of Management* and *Academy of Management Journal*). But again, we did find macro-oriented articles featuring OPD published in macro-oriented journals (e.g., Crilly et al., 2016; Harmon, Kim, & Mayer, 2015), a trend we anticipate will increase.

We caution readers not to use our data to draw conclusions about journals’ receptivity to OPD-based research. Instead, authors should turn to evaluators themselves for these answers. As such, we encourage evaluators—in particular, journal editors—to do their part to reduce lingering ambiguity. Now is the time for editors to take a position regarding their receptivity to OPD. There are two reasons why we strongly recommend that evaluators provide prospective authors with clear statements regarding the viability of publishing OPD-based research.

First, those statements could result in a reduction in selection bias when authors choose an outlet for their work. We believe this is an important point to note because of the differences we observed in OPD-based article publication rates across journals. Those differences could be a function of submitting authors’ selection bias as opposed to the journals (i.e., editors and editorial boards) themselves. Second, our data do not suggest clear patterns regarding journals’ preferences in the OPD-based research they publish. However, and as seen in Table 1, it may be the case that journals vary in their expectations of research design strength when OPD-based research is published. For example, more than any other journal included in our review, *Journal of Applied Psychology* published OPD-based research featuring complex design elements. Table 4 summarizes our recommendations for using, choosing, reporting, and publishing OPD.

### *On OPD Best Practices and the Prevalence of Disagreement*

We would be remiss if we had not looked both within and outside of management in an effort to supplement our recommendations regarding using, choosing, reporting, and publishing OPD. Thus, we compiled the most complete set of best practices concerning how to conduct OPD research. After reviewing dozens of articles from far-ranging disciplines (e.g., economics to public administration) and identifying hundreds of recommendations from those articles, a surprising theme of disagreement began to emerge. While it is true that some OPD best practices appear to be universally agreed upon, many appear to contradict each other, at least on the surface. Others directly, and unmistakably, contradict each other. Appendix A (which readers can obtain in Online Supplemental Materials C) shines a light on this disagreement by providing an exhaustive list of best practices, the rationale behind those practices, contradictions, and evidence (or lack thereof) supporting those practices. Importantly, readers should note that not all recommendations were data driven. Table 5 presents an abbreviated compilation of the practices, but readers are encouraged to consult the complete list in the appendix.

In total, we identified 67 unique practices that we then grouped into 10 topics ranging from the recruitment and selection of OP participants to institutional responsibilities. Space does not permit us to detail each best practice and all of the disagreements. Therefore, we highlight

**Table 4**  
**General Recommendations Regarding Using, Choosing, Reporting,**  
**and Publishing Online Panel Data (OPD)**

Guideline	Key Recommendation	Implications
Using OPD	Research topic and the nature of the research question should be the primary factors determining whether OPD is appropriate for a study	<ul style="list-style-type: none"> <li>• OPD should not be limited to any particular type of management scholarship</li> <li>• The use of OPD does not prevent researchers from employing powerful research designs; in some cases, the use of OPD can facilitate powerful research designs</li> </ul>
Choosing an OPP	Research design and needs should drive decisions about what OPP best fits a research question	<ul style="list-style-type: none"> <li>• Researchers should be aware of substantive differences across OPPs</li> <li>• Researchers should consider the appropriateness of using multiple OPPs, even within the same study</li> <li>• Researchers should explain OPP choice if their decision was driven by methodological considerations</li> </ul>
Reporting OPD	Researchers using OPD should be held to the same reporting standards as researchers using traditional convenience samples	<ul style="list-style-type: none"> <li>• In most cases, OPD use should be clearly reported in Method sections, as should the OPPs from which the data was obtained</li> <li>• Efforts to combine data from samples should be justified and reported such that samples could be disaggregated by other researchers</li> <li>• Selecting or utilizing OP participants meeting specific criteria (e.g., approval ratings) should be reported</li> <li>• Researchers should strive to treat OP participants equitably and should demonstrate evidence they have (i.e., report participant compensation and time requirements)</li> </ul>
Publishing OPD	Scholars (in particular, editors and reviewers) should explain their position on publishing OPD	<ul style="list-style-type: none"> <li>• Positions about OPD appropriateness should be made clear to prospective authors and be evidence based</li> <li>• Researchers should be prepared to defend decisions for study execution; defenses should be evidence based</li> </ul>

*Note:* OP = online panel; OPP = online panel platform.

three of the most highly contested topic areas to demonstrate how the information provided in Appendix A (in the online supplemental materials) informs researchers and evaluators in their efforts to publish and critique OPD-based research. Notably, the three examples we focus on here are not entirely unique to OPD-based research. However, each relates to OPD validity—validity which previous work has questioned. Readers should also note that Appendix A identifies several practices unique to OPD (e.g., use of participant reputation information, capturing Internet protocol addresses, awareness of OPP's policies, etc.). Finally, the examples we discuss here point to the need for further research. This need is particularly true for recommendations that lack empirical support, of which we found more than a few (see Online Supplemental Materials D for a complete list of the research cited in Appendix A).

**Table 5**  
**Abbreviated Compilation of Best Practices**

---

Recommendation by Topic

---

*Topic 1: Recruitment and selection*

1. **Post a “HIT” more than once and be sure to spread those HITs out across different times of the day or even days of the week**
2. **Select only workers who have completed relatively few (e.g., 0–100) studies**
3. **When reputation information is available, restrict samples to “high-reputation” workers (e.g., >95% approval) and possibly larger number of completed studies**
4. **Make use of built-in and user-designed qualification features**
5. Avoid qualification requirements not crucial to your research question
6. **Include eligibility requirements clearly in your recruitment advertisement**
7. Design presurveys that do not give away participation requirements
8. **Describe research tasks generically at the outset**
9. **Initially provide some details of the experiment and approximately what participants will be doing**

*Topic 2: Study planning and design*

10. Be aware of the existence of multiple OPPs and make use of those OPPs
11. **Create unique completion codes that participants must submit to get paid**
12. Be aware of and make use of third-party apps (e.g., TurkPrime) to help manage the research process
13. Increase your sample size to offset anticipated decreases in power
14. Avoid common experimental paradigms and psychological measures
15. Ensure study design consistency when combining samples
16. **Temporally separate IVs and DVs when possible and/or appropriate**
17. **Use source separation for surveys when possible and/or appropriate**
18. **Avoid OPD for cross-cultural research in non-English-speaking countries or when unnecessary**
19. **Make use of OPD for cross-cultural research**

*Topic 3: Measures and controls*

20. **Ask participants if they have participated in similar experimental manipulations before**
21. **Track participant IDs to account for non-naivete; asking participants if they have participated in similar experimental manipulations before is not enough**
22. Measure the completion rate and bounce rate when possible
23. Ask workers how they found your study
24. Ask participants why they participated in your study
25. Measure perceived equity for participation
26. Measure sources of “noise” in the participant’s physical environment
27. Control for the number of studies previously completed by the participant

*Topic 4: Informing*

28. Post informed consent
29. **Provide debriefing when appropriate**
30. Specify any physical environment requirements ahead of time
31. Ensure you provide good directions and that your survey formatting is free of error

*Topic 5: Data quality*

32. **Provide warnings that inattentiveness will not result in compensation**
  33. **Pay inattentive workers but consider blocking them from future participation**
  34. **Offer a second chance to participants who fail attention checks**
  35. **Award bonuses for high-quality work and let participants know ahead of time that bonuses are available**
  36. **Set upper and lower rates on survey completion times and reject work exceeding those limits**
  37. **Do not put a time limit on how fast or slow a survey can be completed by participants**
  38. **Create unique attention checks and/or use instructional manipulation checks**
  39. **Use conventional attention checks to identify and potentially remove responses provided by careless respondents**
- 

*(continued)*

**Table 5 (continued)**

## Recommendation by Topic

- 
- 40. Ask participants whether they were attentive and give them option to have data removed**  
**41. Either prescreen for attentiveness or simply avoid using ex-post screening methods to identify careless respondents**

*Topic 6: Comparisons*

- 42. Track participant IDs when available**  
**43. Compare reliability estimates of your OPD sample to relevant comparison samples**  
**44. Capture IP addresses and reject responses from the same IP address**

*Topic 7: Managing relationships*

- 45. Thank workers and embed tasks with “meaning”—explain meaning of tasks they will complete**  
**46. Monitor discussion boards for chatter about your study**  
**47. Avoid experiments involving deception and consider guaranteeing you will not use deception in your studies**  
 48. Review formal OPP-specific guidelines and act ethically by, for example, clearly identifying yourself to participants, providing reasonable time estimates, paying as soon as possible, and maintaining lines of communication  
 49. Read forums to get a sense of OP participants and introduce yourself to the OP community via web forums if possible  
 50. Provide justifiable and concrete reasons to a participant if rejecting that participant’s work

*Topic 8: Compensation*

- 51. Pay a “fair” wage**  
**52. Pay an appealing—but not overly appealing—wage**  
**53. Pay a low wage—or at least avoid enticing monetary incentives**  
**54. Pay at least median reservation wage**  
**55. Pay U.S. federal minimum wage**  
**56. Pay participants whatever going market rate is (e.g., \$2 per hour)**  
 57. Increase compensation when follow-up time frames increase or more effort is required on the part of the participant  
**58. Use a “hook” strategy where difficult, up-front tasks that pay more must be completed before easy tasks are offered (total payment forfeited if entire study is not completed)**

*Topic 9: Reporting*

59. Be transparent with regard to materials used in your study and the methods used to recruit participants  
 60. Report the amount of compensation participants received and the average study completion time  
 61. If using attention checks or similar indicators to screen for quality, report results both before and after screening techniques were applied  
 62. Collect and report the following: demographics, compensation, the participant’s country of residence, and how non-naivete was handled

*Topic 10: Institutional responsibilities*

63. Journals should offer clear instructions to authors on reporting of survey response rates and how to address nonresponse  
**64. Reviewers and editors should create standards for “low-quality” data screening and reporting**  
**65. Journals should require authors to report pay and the average length of the study**  
**66. Universities/departments should provide funding to pay participants at least minimum wage**  
**67. Internal review boards should consider fair pay**
- 

*Note:* Bolded best practices represent those in which there is disagreement. DV = dependent variable; HIT = human intelligence task; ID = identifiers; IP = Internet protocol; IV = independent variable; OP = online panel; OPD = online panel data; OPP = online panel platform.

The best practice for ensuring high data quality (or identifying and addressing “low-quality” responses) is among one of the most debated topics and is our first example. To illustrate, suppose a researcher embeds a conventional attention check item, such as “Please select the circle under ‘neutral,’” in a survey. If a participant selects the wrong circle, the researcher

assumes that the participant put forth little effort and decides to remove that participant's data. Is such action justifiable, ethical, or effective? Arguments against such techniques include evidence that suggests checks do not substantially improve the data (e.g., Downs, Holbrook, & Peel, 2012; Goodman, Cryder, & Cheema, 2013). Others have argued that such checks may create resentment among participants (Peer, Vosgerau, & Acquisti, 2014).

Some researchers support the use of attention checks but only in certain circumstances. For example, Peer et al. (2014) found that attention checks were effective at improving data quality but only when participants with lower "approval ratings" were recruited. How the attention checks are presented is also debatable. For example, some researchers argue that checks should be used only in "screener surveys" and not in post hoc analyses. Participants who pass the checks in the screener are allowed to continue to the substantive survey, while those who fail the checks are not invited. The concern is that researchers who include such checks in the substantive study may be dishonest and abuse screening techniques during data analysis to obtain desired results (e.g., Chandler et al., 2014; Simmons, Nelson, & Simonsohn, 2011).

As a final point to consider regarding data quality, identifying and removing "bad" data need not be limited to these conventional attention checks. Some researchers have advocated for less traditional approaches, such as creating novel checks, using instructional manipulation checks (e.g., Hauser & Scharz, 2016), or simply asking participants if they were attentive (e.g., Aust, Diedenhofen, Ullrich, & Musch, 2013). Alternatively, researchers have used other indicators of poor data quality, such as survey completion times, response set tendencies, or inconsistent responses. Some researchers have even suggested creating a higher-order scale using multiple indicators to gauge data quality (e.g., Huang, Bowling, Liu, & Li, 2015). Meade and Craig (2012) provided an excellent resource for researchers seeking an in-depth look at available options for identifying careless responses.

Our second example concerns best practices surrounding compensation and is another topic rife with controversy. Some argue that participants should be paid a low wage. Some quantitative evidence has emerged suggesting pay—even as low as \$0.04 per hour—does not impact data quality (Buhrmester et al., 2011), while some qualitative evidence suggests it does (Lovett, Bajaba, Lovett, & Simmering, 2018). Of course, regardless of pay's impact on data quality, there are still ethical issues with which to contend (Gleibs, 2017). For that reason, some advocate for relatively attractive wages, even suggesting the current U.S. federal minimum wage (Goodman & Paolacci, 2017). In between the two groups are researchers who argue that attractive wages open the door to problems (Chandler et al., 2014), thus suggesting a "middle-of-the-road" approach. For example, Stritch et al. (2017) suggested paying participants the going market rate (e.g., \$2 per hour).

Our final example of a highly debated best practice concerns the use of OPD to conduct cross-cultural research. Some researchers endorse the use of OPD to conduct cross-cultural research with little reservation (e.g., Goodman & Paolacci, 2017; Woo, Keith, & Thornton, 2015). Yet, others discourage the use of non-U.S. samples for multiple reasons. Concerns arise when English-based OPPs (such as Amazon's MTurk) are used to recruit and select participants in countries where English is not the native language. The fear is that such samples may not be representative of the population (Buhrmester et al., 2011; Cheung et al., 2017). Second, evidence suggests that non-U.S. OP participants may provide inferior-quality data (Feitosa, Joseph, & Newman, 2015; Litman, Robinson, & Rosenzweig, 2015). That said, we were able to locate several studies that recruited foreign participants using an OPP

located in those participants' native country with no reported data quality issues (e.g., Ng & Feldman, 2012, 2015).

These examples illustrate the disagreement that exists regarding how to execute OPD research. Our primary goal is to ensure that researchers and evaluators of OPD-based research are armed with as much data-driven information to guide their decisions as possible. As we noted at the outset and as our review suggests, OPD is likely to continue to be a convenience sample that an even broader group of management scholars utilize. To realize that potential, we must collectively gain a better understanding of OPD, including when to use it, how to use it, how to report it, and where to publish it. We hope our efforts to raise awareness on these issues and promote informed, critical decision making regarding best practices increase the overall quality of the work produced in our field. Where our efforts uncovered disagreement, we hope scholars devote attention to create consensus that can further guide researchers.

### **Limitations and Opportunities for Future Research**

Although we strove to ensure our work was based on a thorough and rigorous review of the literature, there were some limitations that represent additional opportunities for future research. First, our review does not cover an exhaustive set of management journals. Our sample of journals was chosen based on efforts to balance impact and breadth, but future research could use our list of OPPs to electronically search through an even wider range of management journals. Second, although our review was able to show trends of OPD use in the management literature, we were unable to systematically explore why those trends occurred. We are unable to speak directly to the thoughts, aspirations, and decision-making processes of authors, editors, and reviewers. Future research could shine a light on this "black box" to better understand why these trends occurred, perhaps by collecting data from editors and reviewers who have critiqued work based on OPD or from authors who have attempted to publish such work. A third limitation is that it has been just over a decade since management scholars began using OPD. A decade from now, we would expect scholars replicating our work would generate a sample that would dwarf our data set and include a broader range of topics and OPPs.

Aside from addressing our limitations, there are other important opportunities for future research that follow from our review. Settling the many debates about best practices that our review highlighted is a critical direction for future research. Another opportunity involves exploring whether there may be differences in the scholarly impact—as measured by citations—of articles utilizing OPD compared to articles using other convenience samples. We raise the issue of citations given their far-ranging impact—from pay and promotion decisions to enhancing reputations of departments and universities (Judge, Cable, Colbert, & Rynes, 2007). Could the use of OPD influence citation count? Judge and colleagues (2007) explored a similar question by looking at whether nonstudent samples influenced citation count and found no evidence linking the two. However, that study was performed around the time OPD use was just taking off and no attempt was made to identify samples beyond student or non-student. If a study similar to the Judge et al. analysis was conducted now, what might the relationship between OP samples and citation count look like?

A similar question concerns the extent to which the notoriety of an OPP matters when it comes to an article's impact, as measured by citations. For example, could a better-known

OPP, like MTurk, be viewed as a more trustworthy convenience sample than a lesser-known OPP? In turn, could that trustworthiness ultimately result in more citations? While that scenario is possible, another possibility is that lesser-known OPPs “fly under the radar,” avoiding the scrutiny that more widely known OPPs may generate. We hope future research addresses these and similar questions as they relate to distal consequences of using OPD, both for authors and for the field. While we would have welcomed the opportunity to answer these and similar questions in our review, the relative novelty of OPD and the time required for sufficient variance in citations to amass prevented us from doing so.

## Conclusion

We presented a review of just over a decade of OPD use by management scholars—one that suggests that our field has largely embraced OPD. Our findings suggest a growing legitimacy of OPD in the field. We believe the time has come for the field to embrace a sentiment similar to that expressed over 30 years ago by Ilgen (1986) concerning the appropriateness of laboratory research for management scholarship. Rather than objecting to, being unwilling to consider, and underestimating the utility of OPD, management scholars are better served by asking when and how OPD can best be exploited to answer research-driven questions.

## Notes

1. We thank an anonymous reviewer for highlighting this benefit of our methodology.
2. A complete list of all of the articles with studies included in our review can be found in Online Supplemental Materials A.
3. One caveat is that some studies addressed multiple topics. In those cases, we decided which topic best represented the primary study topic based on subjective evaluations of the title, abstract, keywords, and when necessary, a complete reading of the study.
4. For this discussion, we reference articles (*k*) rather than studies (*n*) because of the similarity of topics across studies within articles. As a robustness check, we ran analyses both ways—using articles as well as studies. The results of analyses were largely similar. When we narrowed our focal topics to those with 12 or more published articles, the only difference in results concerned emotion and affect, negotiation, individual differences, and creativity.
5. We thank an anonymous reviewer for bringing this to our attention.
6. For ease of interpretation, we separately identified and included only online panel platforms (OPPs) representing at least 1% of the data in our graphs and tables. OPPs that failed to meet this criterion were collapsed into one of two categories: *other public* or *other private*. The four most frequently used OPPs, when coupled with OPPs that were unspecified by authors (i.e., 12.1%) represented 93.0% of the studies. Thus, although simplified, our graphs and tables accurately and holistically represent our data.

## References

- Aguinis, H., & Lawal, S. O. 2013. eLancing: A review and research agenda for bridging the science–practice gap. *Human Resource Management Review*, 23: 6-17.
- Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. 2018. Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board Task Force report. *American Psychologist*, 73: 3-25.
- Aust, F., Diedenhofen, B., Ullrich, S., & Musch, J. 2013. Seriousness checks are useful to improve data validity in online research. *Behavior Research Methods*, 45: 527-535.

- Baer, M. D., Bundy, J., Garud, N., & Kim, J. K. 2018. The benefits and burdens of organizational reputation for employee well-being: A conservation of resource approach. *Personnel Psychology*, 71: 571-595.
- Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. 2011. The viability of crowdsourcing for survey research. *Behavior Research Methods*, 43: 800-813.
- Bono, J. E., Glomb, T. M., Shen, W., Kim, E., & Koch, A. J. 2013. Building positive resources: Effects of positive events and positive reflection on work stress and health. *Academy of Management Journal*, 56: 1601-1627.
- Boswell, W. R., Olson-Buchanan, J. B., & Harris, T. B. 2014. I cannot afford to have a life: Employee adaptation to feelings of job insecurity. *Personnel Psychology*, 67: 887-915.
- Brawley, A. M., & Pury, C. L. 2016. Work experiences on MTurk: Job satisfaction, turnover, and information sharing. *Computers in Human Behavior*, 54: 531-546.
- Buhrmester, M., Kwang, T., & Gosling, S. D. 2011. Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6: 3-5.
- Callegaro, M., Baker, R. P., Bethlehem, J., Göritz, A. S., Krosnick, J. A., & Lavrakas, P. J. 2014. *Online panel research: A data quality perspective*. New York: Wiley.
- Chandler, J., Mueller, P., & Paolacci, G. 2014. Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods*, 46: 112-130.
- Chandler, J., Paolacci, G., Peer, E., Mueller, P., & Ratliff, K. A. 2015. Using nonnaive participants can reduce effect sizes. *Psychological Science*, 26: 1131-1139.
- Chandler, J., & Shapiro, D. 2016. Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology*, 12: 53-81.
- Cheung, J. H., Burns, D. K., Sinclair, R. R., & Sliter, M. 2017. Amazon Mechanical Turk in organizational psychology: An evaluation and practical recommendations. *Journal of Business Psychology*, 32: 347-361.
- Chua, R. Y. J., Morris, M. W., & Ingram, P. 2009. Guanxi vs. networking: Distinctive configurations of affect- and cognition-based trust in the networks of Chinese vs. American managers. *Journal of International Business Studies*, 40: 490-509.
- Clark, T. 2011. Gaining and maintaining access: Exploring the mechanisms that support and challenge the relationship between gatekeepers and researchers. *Qualitative Social Work*, 10: 485-502.
- Cohen, J. 1992. A power primer. *Psychological Bulletin*, 112: 155-159.
- Crilly, D., Ni, N., & Jiang, Y. 2016. Do-no-harm versus do-good social responsibility: Attributional thinking and the liability of foreignness. *Strategic Management Journal*, 37: 1316-1329.
- Crone, D. L., & Williams, L. A. 2017. Crowdsourcing participants for psychological research in Australia: A test of microworkers. *Australian Journal of Psychology*, 69: 39-47.
- Crump, M. J., McDonnell, J. V., & Gureckis, T. M. 2013. Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS One*, 8: e57410.
- Cunliffe, A. L., & Alcadipani, R. 2016. The politics of access in fieldwork: Immersion, backstage dramas, and deception. *Organizational Research Methods*, 19: 535-561.
- DeVoe, S. E., & House, J. 2016. Replications with MTurkers who are naïve versus experienced with academic studies: A comment on Connors, Khamitov, Moroz, Campbell, and Henderson (2015). *Journal of Experimental Social Psychology*, 67: 65-67.
- Downs, J. S., Holbrook, M. B., & Peel, E. 2012. Screening participants on Mechanical Turk: Techniques and justifications. *Advances in Consumer Research*, 40: 113-114.
- Elenkov, D. S., Judge, W., & Wright, P. 2005. Strategic leadership and executive innovation influence: An international multi-cluster comparative study. *Strategic Management Journal*, 26: 665-682.
- Feitosa, J., Joseph, D. L., & Newman, D. A. 2015. Crowdsourcing and personality measurement equivalence: A warning about countries whose primary language is not English. *Personality and Individual Differences*, 75: 47-52.
- Fowler, F. J. 2014. *Survey research methods*. Thousand Oaks, CA: Sage.
- Gamache, D. L., McNamara, G., Mannor, M. J., & Johnson, R. E. 2015. Motivated to acquire? The impact of CEO regulatory focus on firm acquisitions. *Academy of Management Journal*, 58: 1261-1282.
- Gleibs, I. H. 2017. Are all "research fields" equal? Rethinking practice for the use of data from crowdsourcing market places. *Behavior Research Methods*, 49: 1333-1342.
- Goodman, J. K., Cryder, C. E., & Cheema, A. 2013. Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making*, 26: 213-224.



- Goodman, J. K., & Paolacci, G. 2017. Crowdsourcing consumer research. *Journal of Consumer Research*, 44: 196-210.
- Göriz, A. S. 2007. Using online panels in psychological research. In A. N. Joinson, K. Y. A. McKenna, T. Postmes, & U.-D. Reips (Eds.), *The Oxford handbook of Internet psychology*: 473-485. Oxford, UK: Oxford University Press.
- Harmon, D. J., Kim, P. H., & Mayer, K. J. 2015. Breaking the letter vs. spirit of the law: How the interpretation of contract violations affects trust and the management of relationships. *Strategic Management Journal*, 36: 497-517.
- Hauser, D. J., & Schwarz, N. 2016. Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods*, 48: 400-407.
- Huang, J. L., Bowling, N. A., Liu, M., & Li, Y. 2015. Detecting insufficient effort responding with an infrequency scale: Evaluating validity and participant reactions. *Journal of Business and Psychology*, 30: 299-311.
- Ilgel, D. R. 1986. Laboratory research: A question of when, not if. In E. A. Locke (Ed.), *Generalizing from laboratory to field settings*: 257-267. Lexington, MA: Heath.
- Insko, C. A., Wildschut, T., & Cohen, T. R. 2013. Interindividual-intergroup discontinuity in the prisoner's dilemma game: How common fate, proximity, and similarity affect intergroup competition. *Organizational Behavior and Human Decision Processes*, 120: 168-180.
- Judge, T. A., Cable, D. M., Colbert, A. E., & Rynes, S. L. 2007. What causes a management article to be cited—article, author, or journal? *Academy of Management Journal*, 50: 491-506.
- Judge, T. A., Ilies, R., & Scott, B. A. 2006. Work-family conflict and emotions: Effects at work and at home. *Personnel Psychology*, 59: 779-814.
- Keith, M. G., & Harms, P. D. 2016. Is Mechanical Turk the answer to our sampling woes? *Industrial and Organizational Psychology*, 9: 162-167.
- Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Jr., Bahník, S., Bernstein, M. J., . . . Nosek, B. A. 2014. Investigating variation in replicability: A "many labs" replication project. *Social Psychology*, 45: 142-152.
- Landers, R. N., & Behrend, T. S. 2015. An inconvenient truth: Arbitrary distinctions between organizational, Mechanical Turk, and other convenience samples. *Industrial and Organizational Psychology*, 8: 142-164.
- Li, H., Kuo, C., & Russell, M. G. 1999. The impact of perceived channel utilities, shopping orientations, and demographics on the consumer's online buying behavior. *Journal of Computer-Mediated Communication*, 5: 1-23.
- Litman, L., Robinson, J., & Rosenzweig, C. 2015. The relationship between motivation, monetary compensation, and data quality among US- and India-based workers on Mechanical Turk. *Behavior Research Methods*, 47: 519-528.
- Little, L. M., Major, V. S., Hinojosa, A. S., & Nelson, D. L. 2015. Professional image maintenance: How women navigate pregnancy in the workplace. *Academy of Management Journal*, 58: 8-37.
- Lovett, M., Bajaba, S., Lovett, M., & Simmering, M. J. 2018. Data quality from crowdsourced surveys: A mixed method inquiry into perceptions of Amazon's Mechanical Turk masters. *Applied Psychology: An International Review*, 67: 339-366.
- Mason, W., & Suri, S. 2012. Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, 44: 1-23.
- Meade, A. W., & Craig, S. B. 2012. Identifying careless responses in survey data. *Psychological Methods*, 17: 437-455.
- Miller, J. D., Crowe, M., Weiss, B., Maples-Keller, J. L., & Lynam, D. R. 2017. Using online, crowdsourcing platforms for data collection in personality disorder research: The example of Amazon's Mechanical Turk. *Personality Disorders: Theory, Research, and Treatment*, 8: 26-34.
- Mochon, D., & Frederick, S. 2013. Anchoring in sequential judgments. *Organizational Behavior and Human Decision Processes*, 122: 69-73.
- Ng, T. W., & Feldman, D. C. 2012. The effects of organizational and community embeddedness on work-to-family and family-to-work conflict. *Journal of Applied Psychology*, 97: 1233-1251.
- Ng, T. W., & Feldman, D. C. 2015. Idiosyncratic deals and voice behavior. *Journal of Management*, 41: 893-928.
- Paolacci, G., & Chandler, J. 2014. Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, 23: 184-188.
- Paolacci, G., Chandler, J., & Ipeirotis, P. G. 2010. Running experiments on Amazon Mechanical Turk. *Judgement and Decision Making*, 5: 411-419.
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. 2017. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70: 153-163.

- Peer, E., Vosgerau, J., & Acquisti, A. 2014. Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods*, 46: 1023-1031.
- Peterson, R. A., & Merunka, D. R. 2014. Convenience samples of college students and research reproducibility. *Journal of Business Research*, 67: 1035-1041.
- Phillips, J. M., Gully, S. M., McCarthy, J. E., Castellano, W. G., & Kim, M. S. 2014. Recruiting global travelers: The role of global travel recruitment messages and individual differences in perceived fit, attraction, and job pursuit intentions. *Personnel Psychology*, 67, 153-201.
- Piccolo, R. F., & Colquitt, J. A. 2006. Transformational leadership and job behaviors: The mediating role of core job characteristics. *Academy of Management Journal*, 49: 327-340.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88: 879-903.
- Ramsey, S. R., Thompson, K. L., McKenzie, M., & Rosenbaum, A. 2016. Psychological research in the internet age: The quality of web-based data. *Computers in Human Behavior*, 58: 354-360.
- Rouse, S. V. 2015. A reliability analysis of Mechanical Turk data. *Computers in Human Behavior*, 43: 304-307.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.
- Shapira, Z., & Shaver, J. M. 2014. Confounding changes in averages with marginal effects: How anchoring can destroy economic value in strategic investment assessments. *Strategic Management Journal*, 35: 1414-1426.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. 2011. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22: 1359-1366.
- Smith, N. A., Sabat, I. E., Martinez, L. R., Weaver, K., & Xu, S. 2015. A convenient solution: Using MTurk to sample from hard-to-reach populations. *Industrial and Organizational Psychology*, 8: 220-228.
- Stritch, J. M., Pedersen, M. J., & Taggart, G. 2017. The opportunities and limitations of using Mechanical Turk (Mturk) in public administration and management scholarship. *International Public Management Journal*, 20: 489-511.
- Swaab, R. I., Phillips, K. W., & Schaerer, M. 2016. Secret conversation opportunities facilitate minority influence in virtual groups: The influence on majority power information processing, and decision quality. *Organizational Behavior and Human Decision Processes*, 133, 17-22.
- Texas A&M University & University of Georgia. 2018. *Texas A&M/University of Georgia rankings of management department research productivity*. Retrieved from <http://www.tamugarankings.com/rankings/2017-2/>
- Tosti-Kharas, J., & Conley, C. 2016. Coding psychological constructs in text using Mechanical Turk: A reliable, accurate, and efficient alternative. *Frontiers in Psychology*, 7: 741.
- Tracy, S. J. 2013. *Qualitative research methods: Collecting evidence, crafting analysis, communicating impact*. New York: Wiley-Blackwell.
- University of Texas at Dallas. 2018. *The UTD top 100 business school research rankings*. Retrieved from: <http://jindal.utdallas.edu/the-utd-top-100-business-school-research-rankings/index.php>
- Walter, S., Seibert, S., Goering, D. D., & O'Boyle, E. H. 2018. A tale of two sample sources: Do results from online panel data and conventional data converge? *Journal of Business and Psychology*. Advance online publication. doi:10.1007/s10869-018-9552-y
- Woo, S. E., Keith, M., & Thornton, M. A. 2015. Amazon Mechanical Turk for industrial and organizational psychology: Advantages, challenges, and practical recommendations. *Industrial and Organizational Psychology*, 8: 171-179.
- Wowak, A. J., Mannor, M. J., Arrfelt, M., & McNamara, G. 2016. Earthquake or glacier? How CEO charisma manifests in firm strategy over time. *Strategic Management Journal*, 37: 586-603.
- Yamagishi, T., Mifune, N., Li, Y., Shinada, M., Hashimoto, H., Horita, Y., . . . Simunovic, D. 2013. Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality. *Organizational Behavior and Human Decision Processes*, 120: 260-271.