Check for updates

Original Manuscript

Organizational Psychology Review

Organizational Psychology Review
1–38
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20413866221097571
journals.sagepub.com/home/opr
SAGE

Conducting organizational survey and experimental research online: From convenient to ambitious in study designs, recruiting, and data quality

Niels Van Quaquebeke
Kühne Logistics University, Germany
University of Exeter, UK

Mojtaba Salem
Technical University Munich, Germany

Marius van Dijke
Erasmus University, Netherlands

Ramon Wenzel
Curtin University, Australia

Abstract

Conducting organizational research via online surveys and experiments offers a host of advantages over traditional forms of data collection when it comes to sampling for more advanced study designs, while also ensuring data quality. To draw attention to these advantages and encourage researchers to fully leverage them, the present paper is structured into two parts. First, along a structure of commonly used research designs, we showcase select organizational psychology (OP) and organizational behavior (OB) research and explain how the Internet makes it feasible to conduct research not only with larger and more representative samples, but also with more complex research designs than circumstances usually allow in offline settings. Subsequently,

Paper received 23 April 2019. Received revised January 5, 2022. Accepted February 26, 2022

Corresponding author:
Niels Van Quaquebeke, Kühne Logistics University, Management Department, Grosser Grasbrook 17, 22767 Hamburg, Germany.
Email: niels.quaquebeke@the-klu.org

because online data collections often also come with some data quality concerns, in the second section, we synthesize the methodological literature to outline three improvement areas and several accompanying strategies for bolstering data quality.

Plain Language Summary: These days, many theories from the fields of organizational psychology and organizational behavior are tested online simply because it is easier. The point of this paper is to illustrate the unique advantages of the Internet beyond mere convenience—specifically, how the related technologies offer more than simply the ability to mirror offline studies. Accordingly, our paper first guides readers through examples of more ambitious online survey and experimental research designs within the organizational domain. Second, we address the potential data quality drawbacks of these approaches by outlining three concrete areas of improvement. Each comes with specific recommendations that can ensure higher data quality when conducting organizational survey or experimental research online.

Keywords
Statistics/methods, online research, online survey, online experiment, online studies

Introduction

These days, most research in the fields of organizational psychology (OP) and organizational behavior (OB) utilizes the Internet in one way or another. To give just two obvious examples: OP/OB scholars routinely use online crowdsourcing platforms to recruit organization members for their surveys (Porter et al., 2019) and they can also draw on a global population that uses phones, tablets, or laptops to respond to much more advanced experimental stimuli and measures than would be possible with paper-and-pencil tests (Askitas & Zimmermann, 2015; Clement, 2020; Gosling & Mason, 2015).

Nevertheless, there is still serious skepticism over online recruited samples and the reliability of the data that were collected online—to the extent that some journal editors and reviewers are more likely to reject studies that only use online surveys and experiments (Landers & Behrend, 2015; Walter et al., 2019; Zack et al., 2019). Some of the criticism is warranted when considering the segments of the general population that participate in online panels like those on the Amazon Mechanical Turk platform (Shaw & Hargittai, 2021), as well as how professional some of the panelists have become in filling in surveys (Hargittai & Shaw, 2020). In fact, some people form dedicated communities to exchange advice on how to complete studies (Sheehan, 2018) while others have even programmed machine bots to do their work for them (Storozuk et al., 2020). So, it seems that the time is ripe to shift from simply regarding online research as a way to conveniently duplicate offline research, and instead become more ambitious about such data collections. Here, we do not mean that completely different methods of data generation should be employed (for that, see, e.g., overviews on Big Data approaches to organizational research, Wenzel & Van Quaquebeke, 2018) – though that is certainly a valuable avenue – but rather that the "bread-and-butter" methods of OB/OP scholars (experimental and survey research) need to be seriously updated in the way how they are deployed online.

Needless to say, other researchers have begun to review the literature. However, despite a large body of methodological studies on different aspects of online surveys and experiments, the recent reviews are fairly limited in the scope of topics they cover. For example, Porter and colleagues (2019) and Aguinis and colleagues (2021) specifically considered issues of

recruiting samples online using Amazon's MTurk and similar online panels. Another example is the review by Göritz (2006), which only covers the topic of incentives in online surveys and experiments. While highly valuable, the focus on such specific subtopics limits researchers' understanding of the broader implications of conducting organizational surveys and experiments online. Meanwhile, broader reviews of the online aspects of surveys and experiments are older than fifteen years (i.e., the latest relevant reviews are Eaton & Struthers, 2002; Stanton & Rogelberg, 2001). Many of these cover issues around programming or specifications for hard- and software—issues that are usually addressed automatically by any querying platform these days or that have largely been rendered moot by increases in processing power across both computers and smartphones. Lastly, there are more contemporary reviews that focus on novel ways of collecting data online (cf. Wenzel & Van Quaquebeke, 2018, on Big Data); however, these do not speak as strongly to the main research methodologies used in OP/OB research (experimental and survey research). It therefore seems timely to provide an updated overview on how to fully leverage the methodological opportunities that the online environment offers for conducting organizational research, including issues of recruiting and data quality assurance.

In short, the current paper generally strives to improve OP/OB scholars' understanding of how they can leverage online surveys and experiments to their best effect. To achieve this, we first briefly sketch the core challenges faced by the fields of OP/OB: issue that have arguably always been around, but that may have become more pressing due to more complex theorizing lacking samples that are appropriately sized and representative. Against this background, we showcase examples of common and more ambitious research designs and illustrate for each how the Internet enables scholars to overcome traditional sample feasibility and representativeness limits. Subsequently, we synthesize extant online methodological research and best practices into three areas (1. improve response rate, 2. improve response quality, 3. respect the participant) to improve data quality. For each area, we offer concrete recommendations for successfully exploiting the unique online methodological opportunities. In sum, with the present paper, we hope to enable more informed decisions on the Internet-related aspects of organizational research among future researchers, Internal Review Boards, reviewers, and editors.

General Challenges Faced by OP/OB researchers

OP/OB scholars study the thoughts, feelings, and behaviors of individual members (e.g., employees or supervisors) and groups of members (e.g., work teams) of work-organizations. Even in traditional work arrangements, the underlying dynamics can be fairly complex. Additionally, as work becomes globally more intertwined and diverse, organizational contexts more dynamic, and concepts of work more fluid, OP/OB theories are called upon to account for these new realities (and populations).

As such, convenient (i.e., accessible) samples and methods of query may not suffice to appropriately test OP/OB theorizing (Bell et al., 2010; Molina-Azorín et al., 2019; Spector, 2019). For instance, it makes little sense to blindly apply insights about autocratic leadership obtained in a Western context to sub-Saharan Africa; to assume that all stakeholders within an organization view a topic similarly; to think that innovation in S&P500 companies mirrors innovation in start-ups; to believe that regular employees' motivational structures can be mapped onto gig-economy workers, or that the work-family interface operates similarly for local and expatriate workers.

In effect, then, OP/OB scholars find themselves challenged on three fronts: First, they need to account for the real-world complexity in their designs; second, to simultaneously collect samples that are more representative of various (harder-to-reach) populations in said realities; and third, to reduce sampling bias

and the related response biases in order to alleviate doubts about the robustness of their results and interpretations (Zack et al., 2019). In response to these challenges, we provide an updated overview of how the Internet allows for a) more complex survey and experimental study designs while b) simultaneously improving sample representativeness and c) ensuring higher data quality.

Designing and Recruiting for More Ambitious Surveys and Experiments in Organizational Online Research

While simpler (e.g., single-source, cross-sectional) research designs may suffice for some research questions (Joshi, 2016), studying theoretically important organizational problems often demands more complex designs. And whether study designs are simple or not, it is pivotal to obtain samples that reflect the populations under investigation (Kulik, 2011). Fortunately, the ubiquity of the Internet across various devices makes it feasible to conduct more complex research designs, which allow researchers to push these studies to target population and recruit participants based on desired qualifications and inclusive profiles, e.g., individuals with specific work habits, medical conditions, or cultural backgrounds (Hewson et al., 2003; Hewson, 2017). Collecting larger samples for more complex study designs from wider populations additionally allows researchers to mitigate power concerns that are prevalent in much of OP/OB research (Tenney et al., 2021) and may also help reducing non-representative bias (Chandler et al., 2014).

In what follows, we will discuss online study designs first for various forms of surveys and then for experiments. In each subsection, we illustrate examples that we identified via Google Scholar using the following keywords: "online survey," "web-based survey," "Internet-based survey," "online experiment," "web-based experiment," "Internet-based experiment," and "online data collection." We filtered our query-based search based on two inclusion criteria: First, we only relay examples that recruited participants on the Internet AND collected their responses to test hypotheses using online surveys or experiments. Second, we focus on examples that investigated organizational topics.

Cross-Sectional, Single Source Online Surveys

Despite the many drawbacks of cross-sectional, single-source data, such simple online surveys still represent the majority of conducted organizational research (Spector, 2019). While not overcoming all methodological and conceptual concerns, there are ways to make such data collection more compelling for theory testing with more representative samples—namely by increasing the sample size and heterogeneity, or by targeting specific, niche samples.

Recruit for Large and Heterogenous Samples. Naturally, one big advantage of online surveys is their recruiting convenience, especially for cross-sectional, single-source study designs. Unlike traditional recruitment, commercial vendors or crowdsourcing platforms such as Qualtrics, Survey Monkey, Prolific Academic, StudyResponse, and Witmart give researchers access to large sample sizes from a global population of employees with relatively little cost and effort (Buhrmester et al., 2011; Callegaro et al., 2014; Porter et al., 2019), thereby significantly increasing their studies' statistical power (Brock et al., 2012; Casler et al., 2013; Gosling et al., 2004; Gosling & Mason, 2015; Luce et al., 2007). These samples are usually quite heterogeneous, but can also be limited to highly specific groups (for a recent review of different panels, see Porter et al., 2019).

While getting access to senior executives or other highly paid professions is unlikely on any of these crowdsourced platforms, researchers can still quickly recruit participants who work in actual jobs and some managerial positions (Schyns et al., 2018; Scott et al., 2014; Taylor et al., 2019). Such individuals can answer

questions regarding their jobs and the perceptions of upper echelon superiors. Moreover, the respondents may work in diverse jobs, including blue collar work, which is often inaccessible for regular paper-and-pencil research despite representing a large portion of the workforce. Indeed, some organizational theories may specifically speak about these working populations (Sessions et al., 2020) or develop theory about them (e.g., gig-workers, Ashford et al., 2018; Petriglieri et al., 2019) and thus need research access to them.

Self-selection may be an issue, of course. For instance, well-paid or very busy people are unlikely to volunteer for work on crowdsourced platforms (Baker et al., 2010). Likewise, while it is becoming less common, some people may still be offline, whether intentionally or due to circumstance (Gosling et al., 2004). Indeed, it needs to be recognized that the size of a sample does not protect against systematic selection bias (Wenzel & Van Quaquebeke, 2018). That said, there are a variety of online channels beyond crowdsourced panels (e.g., Facebook ads, Instagram ads, Google ads, YouTube ads) that still allow one to reach massive and more varied populations than any mono channel recruiting would. Notably, with the help of different URL-parameters attached to the study link, researchers can track participants' entrance points to the study to account (or control) for the different targeted channels.

Employing all of the above strategies can not only significantly increase sample size (and thus statistical power), but also allow researchers to move beyond WEIRD populations (Henrich et al., 2010; Peer et al., 2017) and recruit from vocational and cultural populations that are otherwise harder to reach from within the ivory towers of a metropolitan university. In this way, scholars can improve the external validity and generalizability of their findings (Griffiths et al., 2014; Miner et al., 2012).

Recruit for Niche and Underrepresented Samples and Sensitive Topics. Conducting one's surveys or experiments online not only allows for recruiting larger sample sizes, but also for identifying and targeting niche (e.g., employee/customer) populations and topics. Indeed, one can just think of groups such as humanitarian aid workers, activist groups, or independent remote workers - all of whom flock to their special interest online forums. As such, researchers can collect unique samples by posting adverts in online classifieds and forums (e.g., Craigslist, Backpage, Quora, Adoos, Reddit) (Antoun et al., 2015; Vogel et al., 2016) or specific pages/groups on social networking sites (e.g., Facebook, LinkedIn, Twitter) to recruit niche populations in different cultures (Autio et al., 2013; Sessions et al., 2020). Additionally, snowball sampling often works well when, for instance, attempting to capture niche, but highly networked samples (e.g., entrepreneurs, expats – or their spouses, or creative people) or communities of practice (e.g., cleaners, coders, or HR professionals) (Baltar & Brunet, 2012; Bhutta, 2012; Browne, 2005; Dusek et al., 2015). For example, Arafa and colleagues (2021) used an online snowball approach to recruit a single-source sample of 1,629 Egyptian government staff, particularly in the public health sector and across four governorates, in order to investigate COVID-19's psychological impact. Additionally, researchers may look for and use publicly available contact lists. For example, for their study on intergroup leadership across expats and locals, Salem et al. (2018) used contact details by the United Nations and conferences on humaniatarian aid operations to identify 524 humanitaria aid field workers - a very niche population. Likewise, Ipsen and colleagues (2021) compiled the email contact information of participants from disability-related conferences, as well as relevant organizations, groups, and/or service providers, to send direct email requests with survey links. After additional screening, this approach resulted in a sample of 1,374 adults with disabilities in rural areas of the USA. Reaching such underrepresented or hard-to-reach populations helps researchers refine and extend

organizational theories because they can test the applicability of extant theories in rarely explored contexts that are qualitatively different from the most common ones (Fisher & Aguinis, 2017).

Reaching such specific populations can likewise be achieved via specifications on the above-mentioned crowdsourced platforms. For example, DeCelles and colleagues (2020) invited a large number of MTurk users, who were then reduced to a specific panel of 300 social movement activists after participants answered two filtering questions. And for their study of how anger frames affect social movement mobilization, DeCelles and colleagues (2020) used the Qualtrics panel to recruit a sample of 94 New York finance industry professionals who supported the Occupy Wall Street movement. (Interestingly, before using Qualtrics, the researchers first approached many finance professionals in-person, but could only recruit nine people for their paper-and-pencil survey.) Meanwhile, Heretick and Learn (2020) recruited 180 adults through Prolific Academic for an online survey about bystander responses to coercive sexual harassment in interactions between professors and students. It is the sheer size of these crowdsourced platforms that allows scholars to filter out the sub-populations needed for theorizing.

A more targeted sampling approach is also well suited for studies on sensitive topics, which often report great challenges in recruiting research participants (Ramo & Prochaska, 2012). For instance, Brewer and colleagues (2020) turned to social media as a recruitment platform and found 242 registered nurses in seven days, using their responses to examine how organizations respond to instances of bullying in the context of hospitals. Another example is burned-out employees. While organizational scholarship does theorize about burned-out employees (Halbesleben & Buckley, 2004; Macik-Frey et al., 2007), the actual samples used in much organizational research are subclinical. This means that respective participants still partake in regular work life even if they are more stressed. Truly burned-out samples are missing, and thus, theories on employee burnout may be critically misspecified. Yet, with a bit of ambition, such people may again be found in certain forums or subscribe to special interest mailing lists.

Multi-Source and Multilevel Online Surveys

A multi-source survey is a research design where the ratings for one or multiple variables in the study are matched from different stakeholders (for instance, a focal employee with ratings from superiors, colleagues, or even clients/beneficiaries in the organization) (Magalhães et al., 2019). By a slightly different logic, multilevel surveys inquire individuals who are nested within higher levels or clusters (e.g., teams, organizations, and so on), with the aim of examining theoretical relationships between constructs across levels (Molina-Azorín et al., 2019). Collecting perspectives from multiple sources can triangulate perceptions; as such, they not only address concerns about bias in responding, but more importantly, can also lead to new theoretical endeavors (Schyns & Day, 2010; van Gils et al., 2010). Meanwhile, a multilevel structure allows one to examine how higher-order (aggregates) may exhibit dynamics that are independent from, or even interact with, the individual level (González-Romá & Hernández, 2017).

Unfortunately, both research designs remain less practiced (relative to simple cross-sectional studies) in the organizational research domain because of the various practical hurdles for their operationalization (Bell et al., 2010; Molina-Azorín et al., 2019; see also, Aguinis et al., 2013; Mathieu et al., 2012). Particularly for multilevel designs, a widely accepted rule of thumb suggests that at least 30 higher-level entities and at least 30 individuals in each higher-level entity are required to have appropriate statistical power (Hox, 2002; Scherbaum & Ferreter, 2009). And an even more challenging 50/20 rule applies for testing cross-level interaction effects, meaning that data should encompass at least 50 higher-level entities with 20 individuals per entity (Hox, 2002).

Yet, the proper use of the Internet's opportunities can significantly reduce feasibility constraints, particularly concerning the size and matching of respondents. Consider, for instance, a study by Audenaert and colleagues (2019), who investigated employment relationships across job positions in a public organization with various autonomous agencies. The authors used two online surveys that were linked via an embedded code (see later section on "matching via URL-parameter"), but with a time-lag of three months to minimize common source bias. In this way, they could easily match the participants across both surveys and it allowed them to collect data on 936 employees within 82 job functions (i.e., each job included 2 to 39 employees) (Audenaert et al., 2019). And while it is generally more advantageous to know one's population at the start of sampling (and have their details), these matching codes can also be auto generated within the study and attached to the URL—useful when applying a snowball sampling technique where the survey will be forwarded to other matched respondents. This makes the whole process not only very feasible, but also less prone to error because responses are automatically linked via code. For example, Taylor and colleagues (2019) examined abusive supervision by collecting a matched sample from 500 full-time employees and their supervisors. Likewise, Martin and colleagues (2016) recruited a multi-source (i.e., matched leader-follower) sample of 229 active-duty soldiers in the US Army after they asked their personal contacts at West Point U.S. Military Academy to forward their online survey to other soldiers in leader and non-leader positions. A similar approach was chosen by Korman and colleagues (2021), who targeted managers for a study on burnout. In their study, managers not only supplied their own responses, but also forwarded an automated personalized survey link to their significant others who then reported their perceptions about the manager's health. Note that even language barriers are now surmountable given that survey wording can dynamically shift based on a language selection question (Delice et al., 2019). A case in point is a study by Lu and colleagues (2018), who collected multi-source, multi-wave data from Chinese employees and their expatriate managers in 48 teams across China (Lu et al., 2018). That said, any snowballling smapling for mulit-source data shoud be taken with a grain of salt as likely only others with whom the participants have a better relationship are approached pro (Marcus et al., 2017).

Longitudinal Online Surveys

A longitudinal survey refers to a data collection design that requires the same participants to offer their observations over a meaningful span of time, i.e., at least more than two repeated measures (Joshi, 2016). The theoretical implications of longitudinal data collection are that researchers can study the change in and role of temporal characteristics in constructs, based on findings that have better external validity and a better indication of causality (Wang et al., 2017).

Online surveys render longitudinal data collection more affordable than they used to be because they eliminate most of the costs related to logistics, preparation, and administration of the questionnaire (Lučić et al., 2018; Spitz et al., 2013). Indeed, online surveys allow one to recruit a sufficiently large sample at the outset, thus ensuring that the study retains enough respondents across multiple waves (Reips, 2013). This is a critical advantage because attrition is the main challenge in longitudinal surveys (Lučić et al., 2018). For example, Spurk and colleagues (2020) used an online longitudinal survey in three waves to collect data from 1,477 participants over nine months, finding diverse developmental relationships between career adaptability and proactive career behaviors (Spurk et al., 2020). To test for attrition bias, the authors reported dropout rates for career adaptability (i.e., 31.72% from T1 to T2 and 23.65% from T2 to T3) and proactive career behaviors (i.e., 28.71% from T1 to T2

and 23.93% from T2 to T3). They subsequently performed a multinomial logistic regression analysis, with variables at T1 as predictors of a self-constructed multi-categorical variable reflecting (non)participation across waves (1 = participated only at T1, 2 = participated only at T1 and T2, and 3 = participated at all waves). Their analysis showed that attrition bias was only a minor concern for their online longitudinal survey.

As longitudinal designs entail going beyond a one-shot participation, the perceived anonymity of online surveys becomes an additionally beneficial feature for respondents (Porter et al., 2019). Indeed, with offline methods generally seen as more infringing upon and jeopardizing of anonymity (Singh et al., 2009), online surveys are instrumental for studying sensitive topics (e.g., abusive supervision) and collecting data from people with unconventional sexual orientations, disabilities, or simply remote geographic locations. For example, Lindsey and colleagues (2017) examined why and for whom the ethnic representation of managers affects interpersonal mistreatment in the workplace. Notwithstanding the sensitivity of measuring ethnicity-related topics, they managed to recruit 330 employees (with the qualification of at least 30h/w) at T1 and still retained 284 participants at T2 (i.e., 75% retention rate) in the online survey (i.e., T1 measured ethnic representation and dissimilarity; T2, interpersonal mistreatment).

Online ESM Survey

The experience sampling method (ESM) is a variant of a longitudinal survey design. It captures the momentary behaviors, experiences, thoughts, and attitudes of a person in his or her natural everyday life at several designated time intervals, ranging from a few hours to a day (Niemann & Schenk, 2014; Riediger, 2009). Instant access to emails, web browsers, and user-friendly survey apps such as Expimetrics—all through the convenience of smartphones—has reduced the costs and simplified the logistics of ESM (Beal, 2015; Arnold & Rohn, 2019). Hence, online ESM surveys can now easily achieve larger sample sizes and greater geographic dispersions without causing excessive discomfort (Niemann & Schenk, 2014; Zirkel et al., 2015). And that data can often be matched with additional data coming from smart devices, such as movements, location, or even health markers (Wenzel & Van Quaquebeke, 2018) without the need for extra hardware that would be both costly and burdensome to respondents. Theoretically, online real-time reports are considered more accurate than memory-based reports and are therefore referred to as the "gold standard" for measuring constructs such as affect (Lucas et al., 2021; Schwarz, 2011). Therefore, the rich data obtained via online ESM surveys can open up new perspectives to the interplay among experiences, behaviors, events, and contextual characteristics within and between individuals. A case in point is the investigation of abusive supervision and self-blame by Troester and Van Quaquebeke (2020), who recruited an ESM sample of 275 participants via MTurk for ten working days, twice a day. With this data, the authors could theorize on the link between employees' perception of abusive supervision, the cognitive interpretation of the same, subsequent emotional appraisal, and the resulting behavior response. Moreover, according to Fisher and To (2012), the study of "process" constructs (such as job satisfaction, goals, recovery, moods, emotions, and intrinsic motivation) can also benefit greatly from using online ESM data collection. For example, Kim and colleagues (2021) recruited full-time employees by posting study's invitation and web link on several online social network and community websites (e.g., professional group pages in Facebook, volunteer sections in Craigslist, intranets) in the United States and South Korea. Their final U.S. sample consisted of 779 day-level data points in 10 workdays, i.e., compliance rate of 79%. Similarly,

the final South Korean sample offered 1,024 day-level data points in 5 workdays, i.e., a compliance rate of 92%. The authors first emailed an online survey link at 8:30 a.m. (T1) to assess morning fatigue and sleep quality, and they followed up with a second online study link at 6 p.m. (T2) to measure daily microbreaks, work engagement, fatigue. This method allowed them to confirm when and how microbreaks serve as an effective energy management strategy in the workplace (Kim et al., 2021).

Online Experiments

The experimental design is the gold standard of organizational research methods, as it allows one to test for causal arguments (Antonakis et al., 2010). However, despite calls for the use of experimental designs (e.g., studies that triangulate survey findings with experiments to ensure internal validity), organizational research has largely relied on cross-sectional or longitudinal data (Antonakis et al., 2019; Thau et al., 2014). Some have argued that experiments are underused for several reasons: the potentially low statistical power of offline experimental designs, experimenter effects, the difficulty of enabling/scheduling actual interaction, and, particularly important for organizational researchers, non-random student samples (Reips, 2002; Reips & Krantz, 2010). All these issues can, however, be addressed by running experiments online.

Thanks to the Internet's enormous recruiting capabilities, as already outlined above, online experiments can reach demographically diverse samples well beyond the scope of laboratory environments (Finley & Penningroth, 2015; Paolacci et al., 2010). For example, van Dijke and colleagues (2019) recruited an ESM sample of 338 full-time workers—reflecting regular working hours and a variety of companies—via a professional Dutch research agency, which collected five daily measures per participant. In doing so, the authors found that experimentally induced nostalgia in the treatment group increased intrinsic motivation throughout the working day, especially among participants who experienced low interactional justice. Moreover, because it is much easier to recruit larger samples online even for experiments (Finley & Penningroth, 2015; Jun et al., 2018), researchers can significantly increase the number of treatment groups without reducing statistical power and thus come closer to real-world complexities (Awad et al., 2018; Dimmery et al., 2019).

Another advantage of online experiments is the ability to run experiments around the clock with a high degree of standardization and even provide automated responses, which relieves researchers from dealing with scheduling difficulties and other operational time constraints (Reips, 2002). In that sense, online setups also limit experimenter effects due to the absence of any personal contact (Hewson, 2017).

For cases that require group interactions, online studies also allow researchers to randomly match participants, thus circumventing non-random scheduling (e.g., where friends come together to the lab). Real interactions can, for instance, be facilitated via platforms such as oTree or Lioness. If the paradigm requires more control, then scholars can preprogram dynamic fake interactions that render experimental interactions more realistic. To this end, implementing a waiting room among the actual participants can further increase realism (Reh et al., 2018). Researchers can also easily incorporate precise timings in stimulus displays, track trajectories of participants' mouse movements (Kochari, 2019), and use graphics and animations as online experimental materials (Hewson et al., 2003; Horton et al., 2011).

Additionally, online experiments are especially suited for investigating sensitive topics (including those that may present ethical dilemmas) because particular populations are only more likely to respond to an intervention under the protection of anonymity (Reips, 2010; Reips & Krantz, 2010). Consider, for instance, research on abusive supervision

(Schyns et al., 2018; Taylor et al., 2019; Troester & Van Quaquebeke, 2020) or workplace discrimination. Bailey and colleagues (2013), for instance, conducted an online field experiment to examine potential hiring discrimination based on sexual orientation. The authors reviewed over 4,600 resumes, which varied randomly with regard to gender and implied sexual orientation, finding no evidence for discrimination against gay men or lesbians by employers across four U.S. cities (Bailey et al., 2013). In a different domain, Gross and colleagues (2014) advertised their depression intervention study on Google AdWords and, by doing so, were able to recruit a multilingual sample of 26,194 participants with depressive symptoms, of whom 3,828 were English speakers, 7,477 were Spanish speakers, 5,395 were Chinese speakers, and 9,494 were Russian speakers, which allowed them to test the efficacy of their Mood Screener intervention across cultures.

Being More Ambitious about Data Quality in Online OP/OB Surveys and Experiments

Most research comparing the integrity of data from offline and online approaches suggests that there is little difference between them in terms of psychometric properties and respondents' demographic characteristics (Dillman et al., 2014; Griffiths et al., 2014; Luce et al., 2007; Porter & Whitcomb, 2007; Reips, 2013), often producing comparable responses for behavioral research such as is conducted in the fields of OP/OB (Berinsky et al., 2012; Mason & Suri, 2012).

Nevertheless, some scholars have expressed concerns and skepticisms regarding online data quality. The main points of critique mostly pertain to a wide variety of response biases; that is, when participants respond inaccurately or falsely to questions. For example, respondents' study involvement is usually rather low in online settings because of inattention (Fleming & Bowden, 2009): They sometimes answer on issues for which they have no real insight (Brühlmann et al., 2020), and some have even developed scripts or bots to automatically fill in surveys (Griffin et al., 2021; Prince et al., 2012). In fairness, these are reasonable concerns that researchers need to address (Brewer et al., 2018).

The core question—which we will answer below—is thus how the unique opportunities of online data collection can be leveraged for not only more ambitious study designs and more adequate samples, but also to ensure better data quality. To this end, we first studied several reviews of online surveys and experiments that highlighted a wide range of challenges that researchers face in conducting online data collection. While Stanton and Rogelberg (2001) and Thompson and Surface (2007) were the last to provide the methodological overview for collecting data online with the specific focus on organizational research, other studies reviewed a plethora of challenges and solutions related to the online data quality, e.g., response rate issues such as incentives, data quality issues such as careless responding, and online research ethics issues such as consent and debriefing (Aguinis et al., 2021; Benfield & Szlemko, 2006; Birnbaum, 2004; Fan & Yan, 2010; Gosling & Mason, 2015; Hewson et al., 2003; Kraut et al., 2004; Porter et al., 2019).

Building on the approach used in prior reviews (e.g., Stanton & Rogelberg, 2001), we used a query-based search strategy with a Boolean phrase (i.e., Ti: Web survey/experiment* OR online survey/experiment* OR internet survey/experiment*) AND (Ti: response rate* OR nonresponse* OR response quality* OR online research ethical issues*) to identify relevant papers. We subsequently conducted a review of the reference lists, which supplemented the initial literature search and resulted in a large number of papers (i.e., more than 300 research articles, book chapters, conference proceedings, and reports). Subsequently, we filtered our query-based search based on two criteria. First, to be eligible for inclusion in the review,

we focused on the studies that either showed empirically validated findings or reported widely practiced strategies: for example, the papers that used experiments or similar techniques to evaluate the effectiveness of one or more strategies that improved response rate or the quality of online surveys and experiments. Second, we only considered studies if they offered a clear conclusion about a given strategy. For example, when we found generic solutions to a challenge (i.e., solutions that would equally apply to offline research) or when the underlying support for a given strategy was too vague or inconclusive to formulate a clear recommendation, we refrained from including them in the literature review and best recommendations.

In total, the search procedure resulted in a sample of 151 papers. These were published between 2002 and 2021 in 75 journals that spanned from method journals to content journals in OP/OB, as well as the fields of business, management, social and natural sciences (see Appendix 1 for the list of research method and OP/OB journals). Moreover, we identified and reviewed 20 book chapters and 8 conference proceedings. Taken together, we qualitatively synthesized the findings into three broader areas of improvement that researchers must consider when designing their online surveys and experiments in order to ensure higher online data quality.

Improve Response Rate …

The first area of improvement pertains to taking several precautions to boost response rate in online surveys and experiments. The response rate is a key indicator of data quality (Ball, 2019; Coşkun & Dirlik, 2022) and is defined as "the ratio of people who have accessed a study divided by the total number of people solicited to take part in this study" (Göritz, 2005, p. 2). Notably, a good response rate is key to being able to interpret the obtained results. It ensures that a sound recruitment strategy will also translate into sound results that are less riddled by various forms of self-selection bias that put serious bounds on the interpretations. Such efforts to improve response rate are ideally combined with efforts to simultaneously reduce dropout rate, which refers to the number of respondents who opened the online survey or experiment but did not finish it (Nayak & Narayan, 2019).

Naturally, some of the strategies that have long been used in offline studies have also been shown to work in online studies. For example, researchers generally need to frame the study title, introduction, and questions in a way that is appealing to the audience, but also does not simultaneously create a demand-effect (Fan & Yan, 2010; Shapiro-Luft & Cappella, 2013; Shropshire et al., 2009). Moreover, one can garner a higher response rate by providing academic contact information and including institutional logos in the email invitation and the online study's first page (Göritz & Luthe, 2013; Walston et al., 2006). Unsurprisingly, long online questionnaires (i.e., 30–45 min) generally deter people from partaking in a study – this is also true online. Further, study length negatively affects completion rate (Deutskens et al., 2004), with longer studies generating more "don't know" answers (Galesic & Bosnjak, 2009) and shorter answers to open-ended questions (Krosnick et al., 2002). So, researchers are well advised to keep their studies reasonably short and then provide assurances about the study length in their invitations (i.e., max 10–20 min). At the same time, they may want to refrain from divulging a longer length to avoid negatively affecting dropouts and completion rates (Marcus et al., 2007).

Beyond such general themes, other issues are even more pronounced online than offline. The following strategies are specifically applicable to improving response rates and mitigating dropout in organizational online surveys and experiments involving otherwise serious participants.

By Optimizing Invitations via SMS and Email. Generally, researchers have to find ways for their study invitations to actually reach the

participants. While we mentioned some general recruitment strategies above (i.e., via platforms, Google AdWords, etc.), we also want to highlight that targeted studies—those directed at specific respondents by way of their email—require some consideration.

Here, research has shown, for instance, that SMS pre-notifications to inform respondents of an upcoming online study increase study response (Sammut et al., 2021). Such pre-notifications also help to inform participants that they may have to double check their spam folders because the study invitations were flagged as mass email by spam-blocking tools (de Bruijne & Wijnant, 2014; Mavletova & Couper, 2014). Alternatively, the inviting email address may need to be white-listed via the IT department of the participating organization, or invitation emails need to be sent in smaller batches across longer periods of time.

By Using Online Reminders. Following the initial invitation, researchers should send additional reminders in a timely manner (Lewis & Hess, 2017). The most optimal timing seems two days after the initial invitation (Van Mol, 2017; Bosnjak et al., 2008; Fan & Yan, 2010). Moreover, if the initial invitation was personalized, then online data collection suites have an option to only remind those that have not yet completed the study, thus reducing annoyance for those who already did. A variant of this could be that an additional calendar-file (e.g., .ics) is offered for download so that, in particular for ESM studies, participants are always reminded via the calendar of their own device.

By Personalizing Study (Invitations). It also helps the response rate if researchers personally address respondents in the invitation letter (Sammut et al., 2021), even more so when one can highlight a prior relationship or involve some well-known employees from the same organization (Joinson & Reips, 2007; Sánchez-Fernández et al., 2012) – at least when anonymity is not a bigger issue. Here, online studies offer much more flexibility and speed to personalize invitations because emails can be sent using mail merge functions (e.g., Outlook) that employ an underlying database (e.g., .csv, Salesforce) or by feeding the information into the study platform, most of which then act as the mailer (Sauermann & Roach, 2013). That very information can subsequently be used within the study to directly address the respondent (e.g., "Thank you, Claire, you have now finished mor than 50% of the study") to ensure continued personal involvement.

By Using Forced Answering at the Start or with Non-response Options. Forced answering is an effective high-hurdle technique (Faran & Zanbar, 2019), whereby respondents with no interest—who have a subsequent negative impact on data quality—are encouraged to drop out at the start of the online study. This means respondents can only proceed to the next question when having responded to all or some of the prior items (de Leeuw et al., 2015; Questback, 2013). Note that when researchers believe that respondents can have a real non-answer or prefer to avoid personally sensitive questions, they should still use forced answering but provide non-response options (e.g., "prefer not to answer", "no opinion", "don't know") to minimize both missing responses and potential dropouts (Valentijn et al., 2015).

By Using Simple Language. People exhibit an even lower tolerance for complicated texts when online versus offline. Thus, offering writing instructions and response options in simple language to minimize non-response rate is even more important in online studies (Christian et al., 2007; Smyth et al., 2009; Toepoel & Couper, 2011).

By Using Multi Device (Responsive) Layouts. Surveys and experiments should generally be designed for the users of various devices, e.g., smartphones, tablets, etc., with a so-called

"responsive layout" (de Leeuw & Toepoel, 2018). Such device representation ensures that more people can access the study (e.g., many populations in Africa or Asia have better access to a smartphone than a workstation computer; Toepoel & Lugtig, 2015) and view the material in a sensible format. Having to scroll too much might spur frustration that leads people to drop out, especially when having to scroll horizontally (de Leeuw et al., 2018).

By Avoiding Offering Final Study Results and Instead Give Personalize Feedback. Sharing the final study results with respondents has no effect on completion rate for online studies and, oddly, seems to lower the response rate (Recklitis et al., 2009; Göritz, 2010). Thus, researchers should not try to incentivize respondents with *general* study results and other non-material incentives (Kalleitner et al., 2020).

That said, when topic salience is high, *personalized* feedback at the end of the study can increase response and completion rate (Marcus et al., 2007; Ye et al., 2017; Kühne & Kroh, 2018). Fortunately, online studies can provide this feedback automatically. Indeed, any answer that lends itself to computational functions (e.g., sum, mean, max) can be used to derive a new value within the study. For instance, typical Likert-scale values may be aggregated into scores for feedback (e.g., personality type, number of correct answers, or benchmark comparisons). Moreover, computed values (e.g., "TransformativeLeadership = MEAN(A,B,C)") in combination with Display Logics (e.g., "IF TransformativeLeadership > 5") can serve to provide respondents with personalized, albeit pre-formulated feedback (e.g., "Your leadership style is …").

By Choosing Unconditional Prepaid Personal Cash Incentives over Other Forms of Rewards. Material (monetary) incentives increase response rates by 19% and completion rates by 27% compared to offering no material incentive at all (Göritz, 2010). Notably, donation incentives—which promise to make a monetary donation to a cause in return for participation—have no significant effect on either (Knowles & Stahlmann-Brown, 2021; Göritz, 2005) and, if used alone, can actually lower the response rate (Göritz & Neumann, 2016; Pedersen & Nielsen, 2014). Hence, researchers are advised to go for personal incentives.

When doing so, researchers need to decide whether they will offer material incentives a) unconditionally prior to taking the study, b) conditionally only upon completion, or c) based upon attention performance. Cash incentives provide the strongest direct and positive effect on the response rate and completion rate (Becker et al., 2019; Spreen et al., 2020). Some research suggest that these should be prepaid/unconditional, meaning that all potential respondents receive the cash incentive regardless of actual completion (Becker & Glauser, 2018). This is supposedly because "beyond their monetary value, they are perceived as unconditional tokens of appreciation, which increases trust" (Veen et al., 2016, p. 3; see also Meuleman et al., 2018). Normally, delivering prepaid cash is not a major problem for online organizational studies, but researchers can also hand out electronic gift certificates (e.g., Amazon gift certificates) via emails when a sampling with mailing addresses is not feasible (Kraut et al., 2004).

By Using Lotteries Only for Specific Populations. Another common form of incentivizing is to offer lotteries (also known as prize draws) to respondents after they complete the online study (Coryn et al., 2020; Göritz & Luthe, 2013). These incentives are frequently used because they are easy to implement for a few winners, and costs can be considerably lowered when recruiting a larger sample of respondents. However, their effect on response and completion is generally small (Becker et al., 2019; Göritz, 2010; Sánchez-Fernández et al., 2010). Oddly enough, when respondents are female, have a lower income, lack topic saliency, or prior survey experience, then lottery incentives (e.g., ten 25-euro or five 20-euro

gift certificates) seem to improve response and retention rates in online longitudinal studies (Heerwegh, 2006; Sammut et al., 2021; Su et al., 2008).

Improve Response Quality …

Response quality refers to whether the study responses are a good representation of what participants felt or perceived in the moment of query regarding the subject matter at hand (Cornesse & Blom, 2020). This includes, among others, not cheating or being careless about responding, inputting data in the required format, and adequately responding to the provided stimulus or target referent. In the following, we list the various measures that researchers can implement to ensure higher response quality.

By Detecting Cheating and Careless Responding. Respondents do not always pay attention to instructions and the content of response options (Ward & Meade, 2018). Experience shows that such careless responding behavior is fairly common in samples from online crowdsourcing platforms (Chmielewski & Kucker, 2020; Brühlmann et al., 2020), which can undermine the quality of online data (Jones et al., 2015). Additionally, cheating in online surveys and experiments may occur, thereby introducing a systematic response bias. For example, incentives may motivate respondents to participate multiple times, click as quickly as possible through a study, or encourage them to uncover the "right" responses from professional panel participant communities on the Internet. Thus, it is important to take actions that mitigate such threats to data quality.

By distributing screeners throughout the online study. Screeners ask respondents to act in a specific way—e.g., choosing a particular response option (Berinsky et al., 2014)—which researchers can then use to filter out those who failed to comply. For example, a frequently used screener is, "Please select 'disagree' for this line. Thank you for reading carefully" (Jones et al., 2015). Other common examples in online studies include, "Please verify where you are in the survey by marking a '2' for this item" (Miller & Baker-Prewitt, 2009), and "As a validation check, please answer 'strongly agree' for this question" (Gao et al., 2015). In doing so, researchers should refrain from using screeners with only a few answer options (e.g., "While watching the television, have you ever had a fatal heart attack?") because such yes/no questions have a higher chance of being answered 'correctly' by a speeder (Danilova et al., 2021). Moreover, because respondents' level of attentiveness differs during the completion period, multiple screeners should be distributed throughout the online study (Berinsky et al., 2014). The most effective strategy is to place an initial screener in the beginning as a "warning shot" and then employ other screeners immediately before or within the most critical scales (Jones et al., 2015). The best practice is one screener for every 50–100 items (Meade & Craig, 2012). However, to not irritate honest respondents, researchers should give participants a clear word of warning in the beginning of a study (Breitsohl & Steidelmüller, 2018; Kung et al., 2018). Note that screeners are also referred to as "red herring questions" (Miller & Baker-Prewitt, 2009), "trap questions" (Dillman et al., 2014), "validation questions" (Gao et al., 2015), and "instructional manipulation checks" (Oppenheimer et al., 2009). More recent studies indicate that qualitive open-ended questions (e.g., with minimum three-character responses) work as effective screeners, especially against bots (Griffin et al., 2021). A variant of the above is to ask participants something in the beginning of the survey in a free text field (e.g., "what is your favorite city to visit?") and then ask for the same thing at the end. If the two do not match, the respondent either has not paid attention, or the survey was actually completed by a bot.

In this vein, we advise that researchers code careless responses in the database without taking explicit action against the respective careless respondents. Doing so allows for a later

analysis between those who were inattentive vs. those who were not, which could potentially reveal, e.g., a systematic link between careless response behavior and respondent demographics or experimental conditions (Newman et al., 2021). At the same time, it also avoids the risk that reminders evoke false response behavior (Ward & Meade, 2018). As such, researchers should not immediately correct invalid responses because this would give respondents the confidence that any critical plausibility breaches or omissions will be flagged with a chance for correction. Note that email verifications and the like are still recommended.

By preventing multiple submissions. When researchers are not barred from collecting IP addresses (for ethical and sometimes legal reasons), they should collect IP addresses in order to glean information about the machine used to access the study link. This enables researchers to identify response fraud (e.g., when self-report and other reports are filled in from the same computer, or when a participant tries to partake more than once in a study) and restrict response submission per IP address (Birnbaum, 2021; for an example see Mayer et al., 2012). Additionally, researchers may want to use tools such as IP Hub or IPVOID, which trace back the information on the IPs to identify and remove suspicious or blacklisted IPs (Kennedy et al., 2019). As an alternative to IP address checking, researchers can set cookie session IDs (i.e., small text files that are saved through the browser on respondents' device) that last for a few hours (or even days) so that screened-out respondents may not reenter the online survey or experiment (Nayak & Narayan, 2019; Richter et al., 2018). Note, however, that these can be deleted on a device, which would then enable someone to participate in the study again from the same device.

Another route to take would be personalized studies, i.e. where participants receive a unique login that is mapped to their email address (Sammut et al., 2021). Panel providers such as MTurk provide a variant of this, giving panelists their own unique IDs that are linked to the incentive payout (although it cannot be ruled out that people are registered with multiple emails within the panel). Should researchers opt for this approach, our recommendation is to use automatic login to avoid the errors and hassle that accompany manual logins. An additional reason to employ such recruiting is that a researcher may have a database, for instance, with performance data that needs to be matched onto responses, but preferably in a depersonalized way. An automatic login is a very convenient method because the personalization is not immediately obvious (and can also be hidden from the researcher and not written into the final dataset) when the individual login credentials (or token) are embedded directly in the unique URL parameters (Nestler et al., 2015; Thompson & Surface, 2007).

By co-opting respondents to self-report. A surprisingly effective strategy to detect cheating or careless responding is to simply ask participants at the end of the study to self-report whether they were cheating or did not pay full attention to the study (Clifford & Jerit, 2016). Indeed, the authors of the present paper and affiliated researchers have had good experiences with simply asking respondents whether their data should be used for the later analysis (Aust et al., 2013). When doing so, we always briefly explain that the scientific insights that people rely on need sound data. We also highlight that their self-reporting has no effect on their outcomes (e.g., credit in panel or lottery chances) or produces other repercussions. In other words, it is an appeal to respondents' integrity and their wish to further science (cf. Fazekas et al., 2014), and it often works.

By analyzing response patterns. Online studies allow some data analysis to occur in a real-time manner (Donohoe et al., 2012), which empowers researchers to immediately begin identifying carelessness in two ways (Brühlmann et al., 2020): First, via responses that are consistently far from the mean of scale items for each question (i.e., also called multivariate outlier analysis) (Ehlers et al., 2009); second, via the longest continuous string of identical responses for a given

respondent (e.g., selecting "strongly agree" for 15 consecutive items) (Johnson, 2005). The latter is known as the Long String and can be calculated by taking the absolute frequency of the most commonly chosen response option in each grid, summing them up throughout the online study, and dividing the sum by the total number of grid questions (Meade & Craig, 2012). This index theoretically ranges from 0 to 1, with lower values indicating less straight-lining (i.e., higher data quality). A real-time calculation of such indexes can help researchers make real-time decisions about whether they need to sample more trustworthy respondents to meet their sampling target or whether they can stop with the recruitment (Reyes, 2020).

By analyzing completion time based on defined criteria. With any study that is conducted online, it is possible to log completion times (for sections or whole surveys) to the millisecond. This allows researchers to check respondents' completion time and remove quick responses from the dataset. Again, if done in real-time, this helps researchers to determine whether they need to continue recruiting or can close the study for participation. However, in doing so, researchers should be cautious about the criteria used to identify speeding and refrain from applying random cut-off values. Relaxed criteria recognize 'false positives' as quick responses whereas stiffer criteria yield 'false negatives' (Greszki et al., 2014). Therefore, we recommend that researchers base the definition of speeding on their best estimation of the minimum time it takes to read instructions and provide answers in their online studies. While some studies recommend that it should be at least two seconds per item (Bowling et al., 2021; Huang et al., 2012), it seems the sounder strategy is to pretest the survey with serious participants who are a) native to the study language, b) technologically versed, and c) match the envisioned participants' profile in terms of background in order to determine the lower bound for participation time.

By Providing Immediate Answer Validation. Some items require a certain format that participants may not always adhere to when quickly answering the items. A typical validation relates to the recording of dates: The provision of a defined answer format by means of a smaller box for the month and a larger one for the year, as well as placing symbolic instructions (e.g., MM/YYYY) within the answer boxes, elicits more consistent responses (e.g., years in four digits) (Christian et al., 2007). Online, researchers can also design their studies to immediately validate responses against a specified format (Conrad et al., 2017). For instance, they can immediately check if inserted values represent integers or letters, if certain symbols are present (e.g., checking the @-sign for an email), and if a minimum (or maximum) threshold of character count is met (or exceeded). Additionally, researchers can use more advanced tools such as RegExr (i.e., a logical syntax that can be plugged into webforms) to validate input against predefined restrictions. Implementing such measures not only assists participants in quickly identifying the right input format, but it also helps to reduce errors in later analyses when certain values are not recognized or considered missing by the software.

By Making the Interaction Feel More Real and Personal. Dynamic pre-filling can infuse the study with more realistic-seeming interactions and increase data quality in terms of time and accuracy (Dolnicar et al., 2013). Any available information (e.g., previous answer, computed value, information from a database, such as from a different respondent) can serve as a dynamic insertion to a given item or question, and thus help to personalize or manipulate the study experience for the respondent. Researchers should explore the functionality of their chosen software platform to use pre-filling in several ways.

Moreover, in interactive online studies such as repeated public goods experiments, which involve behaviors like cooperation and punishment (Arechar et al., 2018), respondents can actually chat with each other, with a research confederate, or even with a well-programmed

chat bot that can take over various conversational tasks at scale using software platforms such as SMARTRIQS (Kim et al., 2019; Molnar, 2019). If a true chat functionality cannot be programmed as an applet within one web page, then a chat can simulated through dispersed responses over multiple pages. Whatever participants enter into a text-field will then be displayed on the next page, thereby lending credibility that whatever they read from the other (bogus) respondents on the next page is also real. Obviously, if such a design is implemented, researchers should consider a forced delay between the two pages so that responses by the other (bogus) respondents seem more realistic (e.g., Reh et al., 2018).

By Minimizing Matching Error via URL-parameter. Survey designs that require data matching, such as multi-source or longitudinal designs, are especially prone to errors in the matching process. Online studies can take some of the headache out of this by linking studies via URL-parameters.

URL parameters are values added after the base URL. They start with a "?" followed by the name of the parameter, an equal sign, and the parameter's value, e.g., ?department = 23 (all subsequent parameters can be added with a leading ampersand "&"). Any information that is known before a URL is sent out can be used as a URL parameter (for more information, see Litman et al., 2017). For example, "https://www.mysurvey.com?department=23&level=3" would pass on the parameter "department" with the value "23" and "level" with the value "3". Via the study's landing page, the parameter values can then be retrieved and written into the database, and later inserted into a placeholder within the study. Furthermore, if the respective information was retrieved, an item within the study could then say "dear [first-name], what impact did your leader, [boss-name], have on the development of the relationship with your client, [clientname]."

Using URL parameters offers helpful capabilities beyond the personalization described above. For instance, to match responses across longitudinal study waves, researchers can generate a unique identifier code that is passed as a URL parameter along with each invite link. Likewise, such parameters can easily be used for multi-source snowball samples. For example, to match leaders and followers without error, leaders may be targeted as focal respondents (via whatever recruiting technique) and encouraged to send a corresponding link to at least three of their subordinates or significant other (e.g., Korman et al., 2021). The link to this follower-survey can comprise a referral token such as the response ID that was generated within the leader-survey, indicating which leader the link came from (e.g., https://www.mysurvey.com?leader=xyz).

By Lowering Response Frustration with Display Logics. Researchers should always aim to reduce burden by avoiding irrelevant requests. Thus, they should design online studies to dynamically display sets of questions based on certain criteria, such as a respondent profile, an intended manipulation, or prior responses that specify distinct survey routes (Molnar, 2019; Nayak & Narayan, 2019). For instance, questions about staffing numbers should only be shown to respondents with reliable organizational knowledge (e.g., work in HR), while items about supervising teams should be hidden from frontline staff.

Additionally, a planned missing data design can reduce participant burden and thus result in higher validity and reduced rates of unplanned missing data. For instance, a three-form design facilitates data for 33% more survey questions than can be answered by any one respondent, or reduces completion time by the same fraction (Graham, 2012). In such a design, the items are divided into four sets: X, A, B, and C. Set X is centrally administered to everyone, and each respondent will additionally answer two out of the remaining three item sets AB, CA, or BC. Note that the implementation of such a design requires a carefully crafted item logic and distribution so that responding feels natural (Jia et al., 2014; Jorgensen et al., 2014). Such planned

missing data design results in data that are characterized by missing completely at random (MCAR), which is subsequently handled by using contemporary imputation methods (e.g., FIML for SEM in Mplus; Flatau Harrison et al., 2018; Lüdtke et al., 2017).

By Using Pictures, Audios, and Videos Only Very Selectively. Online surveys and experiments offer unique opportunities to use multimedia features such as pictures, audio, and video for garnering interesting data. However, if designed carelessly, multimedia features negatively affect the quality of data collection by introducing response bias (Chen & Tseng, 2017; Mahon-Haft & Dillman, 2010). As a result, we recommend the below strategies as a check against the careless use of multimedia features.

By refraining from using pictures for mere aesthetic reasons. Contrary to common belief, the inclusion of pictures does not help with keeping respondents interested and reducing dropout (Barbulescu & Cernat, 2012). Moreover, researchers should be aware of the systematic influence of pictures on respondents' behavior: Because pictures prime specific memories, they can affect how respondents construct their answers based on the retrieval of those memories (Couper et al., 2007; Trübner, 2020). For these reasons, researchers should refrain from using pictures. However, the option to provide pictures either as a stimulus or as a response can be a rich source of data. For example, a study design could require participants to produce and transmit images of their work environment or office desk (e.g., online ESM surveys, Arnold & Rohn, 2019), take part in picture-rating exercises (e.g., Wilson et al., 2018), test a construct via pictures (e.g., Hou et al., 2020), or reflect on a topic specifically revolving around pictures (e.g., Rennung et al., 2016).

By using audio and video only with bolded instructions and selectively. When audio is included in online studies, researchers should highlight instructions and symbols to nudge respondents into turning on their audio to a sufficient volume level (Beier & Schulz, 2015). Given the concerns about sound presentation—ranging from background noise to poor speaker quality—an effective way to increase control over sound delivery online is to ask participants to wear headphones (Woods et al., 2017). Audio recordings are, for instance, useful for examining the impact of vocal characteristics (e.g., gender, tone, etc.) on respondents' behavior (Couper, 2007)—an aspect that is often overlooked, for example, in leadership research. Case in point: Lavan and colleagues (2019) used audio recordings in their online experiments to investigate how familiar and unfamiliar voices affect identification perception. Likewise, it is imaginable that audio recordings can be used on the respondent side. For instance, in cases where respondents are not comfortable writing (potentially because of a high analphabetism ratio in the sample), then app-based studies or a kiosk-type tablet (with symbolic instructions) may use voice recordings to nevertheless generate rich insights (Revilla et al., 2020).

With respect to videos, issues such as slow Internet speed, data caps, computer configuration, or workplace policies may prevent or demotivate respondents from watching video clips completely and thus cause nonresponse bias (Dixon & Tucker, 2010). Therefore, researchers are generally advised to refrain from using videos to minimize indiscernible distortions in the sample (Shapiro-Luft & Cappella, 2013). Yet, if videos are considered essential for the online data collection, forced watching should be activated so that respondents cannot proceed to subsequent screens unless the entire video is watched. For example, video data collection may be necessary in online experiments that involve preferential looking tasks, whereby two stimuli are presented side-by-side and the experimenter usually records fixation durations to each stimulus (Ismail et al., 2021; Semmelmann et al., 2017). A more recent development includes the option of checking body language, head movements, and eye gazes via the built-in webcams on participants' computers (Aghajanzadeh et al., 2020). With

the popularity of HTML5 technology, Internet browsers now support the activation of webcams without the need to install additional webcam plug-ins.

By Using JavaScript (or Software Solutions) to Tackle Experimental Challenges of Similarity in Stimulus and Assessment of Time. Online studies pose challenges for particular types of experiments such as reaction-time and positioning designs (for an overview, see Kohavi et al., 2020; Woods et al., 2015). For example, screen positions within a browser can be hard to interpret as researchers cannot usually know whether a browser window has been rescaled and to what size. Likewise, reaction times cannot be reliably assessed across digital devices by normal means (Sauter et al., 2020). To ensure accuracy and test-retest reliability in presentation and response recording, researchers should use specialty scripts that require a working knowledge of programming languages such as JavaScript (Anwyl-Irvine et al., 2020; Garaizar & Reips, 2019). However, if researchers are not able to program and code, they should turn to dedicated online research software providers. Currently available software solutions include Gorrilla.sc (Anwyl-Irvine et al., 2020), jsPsych (de Leeuw, 2015), Millisecond (i.e., its product Inquisit or LIONESS; Giamattei et al., 2020), Psychstudio, ScriptingRT (Schubert et al., 2013), Tatool (von Bastian et al., 2013), QRTEngine, WebExp (Keller et al., 2009), and SMARTRIQS (Molnar, 2019), which have all been developed around offering precise timing and a full interface for task design and experiment administration, with no to minimal programming skills.

Respect the Participant…

Naturally, an integral part of doing research is to minimize any negative effects on people and organizations. Doing research online entails taking participants' digital security, privacy, and dignity seriously so that they feel respected and safe to participate truthfully in the study, as well as remain motivated to help researchers out on future studies (James & Busher, 2015). Generally, tools such as the Data Ethics Canvas—which help researchers identify and manage respective issues (Data Ethics Canvas, 2021)—can be helpful support resources to this end. More specifically, we recommend that scholars consider the following strategies.

By Following the Established Legal Framework. Needless to say, researchers must adhere to the legal frameworks of the countries in which their research is being undertaken to ensure privacy beyond data security. This is especially critical for online research, which is often undertaken across different countries. As it stands, the current EU general data protection (GDPR) regulation is one of the strictest in the world and offers guidance as to what is expected (more information can be obtained here: https://www.eugdpr.org). It is too comprehensive to cover here, but adhering to it is imperative for doing research with organizations and individuals within Europe, and it represents a useful guide about what other jurisdictions may mandate in the future.

By Securing Data Transmission and Storage. Researchers should always ensure that research data transmission and storage involve several layers of security. First, encrypt data during transmission by ensuring that the survey system uses a secure transmission protocol (as indicated by "https://") (Benfield & Szlemko, 2006; McInroy, 2016). Second, protect access to the research database, related file directories, and other storage entities. For instance, using longer (but easy to remember) passphrases alongside 2-factor authentication is superior to using an 8-character password (Keith et al., 2007). Third, manage file-privileges (e.g., open, edit, share) differentially across collaborators (e.g., study lead, co-authors, research assistants) so they only have the absolute necessary level of control (Denissen et al., 2010; Reips, 2002). Naturally, all data that are

exported from the database into a datasheet on a local hard drive should be protected at all times (e.g., via password encryption).

As part of the above, researchers should take necessary precautions so that the server is secure against hacking, short-circuit fault, or theft (Barchard & Williams, 2008; Wilt et al., 2012). This includes, among other things, ensuring that the server is self-maintained, frequently installing security updates, doing remote backups (preferably on other media or servers in different locations), and securing the building where the server is stored. Luckily, most of these issues are handled by the major software providers themselves (Callegaro et al., 2015).

By Minimizing Privacy Risks. Although online studies can offer respondents anonymity and confidentiality, the threat of respondent identifiability remains a serious ethical concern (Roberts & Allen, 2015; Whelan, 2007). For example, combinations of variables (e.g., sex, age, and ethnicity) can expose individual respondents' identity relatively easily if the collected data comes from employees of a single company and/or contains identifying information such as their IP address (Barchard & Williams, 2008). Researchers should mitigate these immediate risks by capturing identifying information via a separate survey or web page that is technically separate from the main study. The overt identifiable information and the main study response can then be linked through a meaningless but unique identifier that connects the two data sets, if warranted (e.g., when required for statistical analysis of controls) (e.g., Leung & Unal, 2013). In addition, researchers may want to obfuscate the data by only allowing certain researchers to have access to, for instance, the meaning behind otherwise meaningless variable names and values. Benfield and Szlemko (2006) also suggest that the IP addresses should either be deleted early in data analysis or not recorded at all if possible. Ultimately, researchers who want to record IP addresses should explicitly mention so in the consent request (Granello & Wheaton, 2004; Saari & Scherbaum, 2020).

By Enabling Informed Consent Opportunity. Obtaining informed consent and doing debriefings can be a challenge online. Yet, similar to issues of privacy and data security, we recommend that researchers take these steps seriously, particularly before engaging participants in online experiments.

Researchers should prepare a consent request in understandable language that clearly asks respondents for their digital permission upfront (e.g., using animated videos when soliciting informed consent; McInroy, 2017). In this light, researchers should brief respondents upfront about a) the content of the research, b) the identity and contact options of the researchers, c) the extent of confidentiality and anonymity of the responses, d) the average length of time to complete the study, and e) the risks involved in participating, such as being asked to disclose uncomfortable or potentially embarrassing information (Kunz et al., 2020; Sue & Ritter, 2012). Given that people are increasingly accustomed to using computer software, online accounts (e.g., Facebook, Gmail), and some forms of online payment without signatures, it has become customary for consent to be given via a button that indicates "I consent" after reading the introductory information (Rowbotham et al., 2013).

By Allowing a "Quit the Study" Option. Some also argue that researchers should equip each screen with a "Quit the study" button that takes participants to a page with researchers' contact information and other debriefing information, so that respondents who drop out of the study still receive the debriefing (Barchard & Williams, 2008; Gupta, 2017). Alternatively, researchers should program a window with the debriefing information that will pop up when respondents close the study's browser window. We want to caution, however, that the latter may not always work because of the pop-up blockers that exist in modern browsers.

By Debriefing Respondents. Debriefing is an essential research ethics procedure wherein respondents are informed about their participation in research and provided with controls over their data privacy (Wang & Kitsis, 2013). Yet it remains a major ethical concern in online studies, especially as some respondents drop out prematurely (Benfield & Szlemko, 2006; Hilbig & Thielmann, 2021; Pittenger, 2003; Vitak et al., 2016). The most optimal strategy is that debriefing information should be given via a link and not included in the body of an email (Benbunan-Fich, 2017; McCambridge et al., 2012). This increases the likelihood that recipients will actually read the debriefing information. Technically, researchers should program a debriefing so that respondents are emailed automatically by way of a trigger that executes as soon as the online study's window is closed, assuming that participants were asked to provide a valid email address (e.g., Zong et al., 2018). Otherwise, a text at the end of the survey or experiment must suffice.

Limitation of the Covered Literature

We must make one limiting remark regarding the presented synthesis. Conducting studies online is naturally interwoven with the state and adoption of technology within the sampled population. The last decade alone has seen a tremendous expansion of Internet technology and connected devices. Some studies on screen resolutions (Callegaro, 2010; Erens et al., 2019), for instance, might already be outdated despite only being published a few years ago. The possibilities of tomorrow are already on our doorstep: Early studies are, for example, experimenting with online remote eye-tracking, facial expressions, and heart rate monitoring (Chandler & Shapiro, 2016; Goodman & Paolacci, 2017). It may only be a few years before we see a revival of the classic interview format, led not by humans, but by artificial intelligence (AI)-enabled assistants such as Alexa, Siri, Google Assistant, or Cortana in virtual reality (VR). Thus, the present paper can and should only be considered a snapshot in time.

Conclusion

As noted at the onset of this paper, much research in the fields of OP and OB takes the Internet for granted – often more or less duplicating offline research strategies – without a clear understanding of the unique opportunities and challenges for conducting surveys and experiments online. In the present paper, we showed that the development and spread of the Internet across the world and into organizations has increased the feasibility of conducting and recruiting for more ambitious research designs. That said, certain data quality challenges remain that should be actively tackled. To this end, we presented some guidelines that are to assist researchers, reviewers, and editors.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Niels Van Quaquebeke https://orcid.org/0000-0002-6141-4659

References

Aghajanzadeh, S., Jebb, A. T., Li, Y., Lu, Y.-H. & Thiruvathukal, G. K. (2020). Observing human behavior through worldwide network cameras. In Woo, S. E., Tay, L. & Proctor, R. W. (Eds.), *Big data in psychological research* (pp. 109–123). American Psychological Association.

Aguinis, H., Gottfredson, R. K. & Culpepper, S. A. (2013). Best-practice recommendations for

estimating cross-level interaction effects using multilevel modeling. *Journal of Management*, *39*(6), 1490–1528. https://doi.org/10.1177/0149206313478188

Aguinis, H., Villamor, I. & Ramani, R. S. (2021). MTurk research: Review and recommendations. *Journal of Management*, *47*(4), 823–837. https://doi.org/10.1177/0149206320969787

Antonakis, J., Banks, G. C., Bastardoz, N., Cole, M. S., Day, D. V., Eagly, A. H. & Weber, R. (2019). The leadership quarterly: State of the journal. *Leadership Quarterly*, *30*(1), 1–9. https://doi.org/10.1016/j.leaqua.2019.01.001

Antonakis, J., Bendahan, S., Jacquart, P. & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, *21*(6), 1086–1120. https://doi.org/10.1016/j.leaqua.2010.10.010

Antoun, C., Zhang, C., Conrad, F. G. & Schober, M. F. (2015). Comparisons of online recruitment strategies for convenience samples: Craigslist, Google AdWords, Facebook and Amazon's Mechanical Turk. *Field Methods*, *28*(3), 231–246. https://doi.org/10.1177/1525822X15603149

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N. & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Arafa, A., Mohamed, A., Saleh, L. & Senosy, S. (2021). Psychological impacts of the COVID-19 pandemic on the public in Egypt. *Community Mental Health Journal*, *57*(1), 64–69. https://doi.org/10.1007/s10597-020-00701-9

Arechar, A. A., Gächter, S. & Molleman, L. (2018). Conducting interactive experiments online. *Experimental Economics*, *21*(1), 99–131. https://doi.org/10.1007/s10683-017-9527-2

Arnold, K. D. & Rohn, K. C. (2019). Internet-enabled experience sampling: Innovations in signaling and data sources. In *Advancing educational research with emerging technology* (pp. 124–151). IGI Global.

Ashford, S. J., Caza, B. B. & Reid, E. M. (2018). From surviving to thriving in the gig economy: A research agenda for individuals in the new world of work. *Research in Organizational Behavior*, *38*, 23–41. https://doi.org/10.1016/j.riob.2018.11.001

Askitas, N. & Zimmermann, K. F. (2015). The internet as a data source for advancement in social sciences. *International Journal of Manpower*, *36*(1), 2–12. https://doi.org/10.1108/IJM-02-2015-0029

Audenaert, M., George, B. & Decramer, A. (2019). How a demanding employment relationship relates to affective commitment in public organizations: A multilevel analysis. *Public Administration*, *97*(1), 11–27. https://doi.org/10.1111/padm.12378

Aust, F., Diedenhofen, B., Ullrich, S. & Musch, J. (2013). Seriousness checks are useful to improve data validity in online research. *Behavior Research Methods*, *45*(2), 527–535. https://doi.org/10.3758/s13428-012-0265-2

Autio, E., Dahlander, L. & Frederiksen, L. (2013). Information exposure, opportunity evaluation, and entrepreneurial action: An investigation of an online user community. *Academy of Management Journal*, *56*(5), 1348–1371. https://doi.org/10.5465/amj.2010.0328

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A. & Rahwan, I. (2018). The moral machine experiment. *Nature*, *563*(7729), 59–64. https://doi.org/10.1038/s41586-018-0637-6

Bailey, J., Wallace, M. & Wright, B. (2013). Are gay men and lesbians discriminated against when applying for jobs? A four-city, internet-based field experiment. *Journal of Homosexuality*, *60*(6), 873–894. https://doi.org/10.1080/00918369.2013.774860

Baker, R., Blumberg, S. J., Brick, J. M., Couper, M. P., Courtright, M., Dennis, J. M., Dillman, D. A., Frankel, M. R., Garland, C. P., Groves, R. M., Kennedy, C., Krosnick, J. A., Lavrakas, P. J., Lee, S., Link, M. W., Piekarski, L., Rao, K., Thomas, R. K. & Zahs, D. (2010). Research synthesis: AAPOR report on online panels. *Public Opinion Quarterly*, *74*(4), 1–71. https://doi.org/10.1093/poq/nfq048

Ball, H. L. (2019). Conducting online surveys. *Journal of Human Lactation*, *35*(3), 413–417. https://doi.org/10.1177/0890334419848734

Baltar, F. & Brunet, I. (2012). Social research 2.0: Virtual snowball sampling method using Facebook. *Internet Research*, *22*(1), 57–74. https://doi.org/10.1108/10662241211199960

Barbulescu, M. & Cernat, A. (2012). The impact of pictures on best-worst scaling in web surveys. *International Review of Social Research*, *2*(3), 79–93. https://doi.org/10.1515/irsr-2012-0028

Barchard, K. A. & Williams, J. (2008). Practical advice for conducting ethical online experiments and questionnaires for United States psychologists. *Behavior Research Methods*, *40*(4), 1111–1128. https://doi.org/10.3758/BRM.40.4.1111

Beal, D. J. (2015). ESM 2.0: State of the art and future potential of experience sampling methods in organizational research. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(April), 383–407. https://doi.org/10.1146/annurev-orgpsych-032414-111335

Becker, R. & Glauser, D. (2018). Are prepaid monetary incentives sufficient for reducing panel attrition and optimizing the response rate? An experiment in the context of a multi-wave panel with a sequential mixed-mode design. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, *139*(1), 74–95. https://doi.org/10.1177/0759106318762456

Becker, R., Möser, S. & Glauser, D. (2019). Cash vs. Vouchers vs. Gifts in web surveys of a mature panel study—main effects in a long-term incentives experiment across three panel waves. *Social Science Research*, *81*, 221–234. https://doi.org/10.1016/j.ssresearch.2019.02.008

Beier, H. & Schulz, S. (2015). A free audio-CASI module for LimeSurvey. *Survey Methods: Insights from the Field (SMIF)*, *2015*, 1–10. https://doi.org/10.13094/SMIF-2015-00011

Bell, B. A., Morgan, G. B., Kromrey, J. D. & Ferron, J. M. (2010). The impact of small cluster size on multilevel models: A Monte Carlo examination of two-level models with binary and continuous predictors. *JSM Proceedings, Survey Research Methods Section*, *1*(1), 4057–4067.

Benbunan-Fich, R. (2017). The ethics of online research with unsuspecting users: From A/B testing to C/D experimentation. *Research Ethics*, *13*(3-4), 200–218. https://doi.org/10.1177/1747016116680664

Benfield, J. A. & Szlemko, W. J. (2006). Internet-based data collection: Promises and realities. *Journal of Research Practice*, *2*(2), 1–15.

Berinsky, A. J., Huber, G. A. & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon. Com's mechanical Turk. *Political Analysis*, *20*(3), 351–368. https://doi.org/10.1093/pan/mpr057

Berinsky, A. J., Margolis, M. F. & Sances, M. W. (2014). Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *American Journal of Political Science*, *58*(3), 739–753. https://doi.org/10.1111/ajps.12081

Bhutta, C. B. (2012). Not by the book: Facebook as a sampling frame. *Sociological Methods Research*, *41*(1), 57–88. https://doi.org/10.1177/0049124112440795

Birnbaum, M. H. (2004). Human research and data collection via the internet. *Annual Review of Psychology*, *55*(1), 803–832. https://doi.org/10.1146/annurev.psych.55.090902.141601

Birnbaum, M. H. (2021). Advanced training in web-based psychology research. *Zeitschrift für Psychologie*, *229*(4), 260–265. https://doi.org/10.1027/2151-2604/a000473

Bosnjak, M., Neubarth, W., Couper, M. P., Bandilla, W. & Kaczmirek, L. (2008). Prenotification in web-based access panel surveys: The influence of mobile text messaging versus E-Mail on response rates and sample composition. *Social Science Computer Review*, *26*(2), 213–223. https://doi.org/10.1177/0894439307305895

Bowling, N. A., Huang, J. L., Brower, C. K. & Bragg, C. B. (2021). The quick and the careless: The construct validity of page time as a measure of insufficient effort responding to surveys. *Organizational Research Methods*, Advance online publication. https://doi.org/10.1177/10944281211056520

Breitsohl, H. & Steidelmüller, C. (2018). The impact of insufficient effort responding detection methods on substantive responses: Results from an experiment testing parameter invariance. *Applied Psychology*, *67*(2), 284–308. https://doi.org/10.1111/apps.12121

Brewer, E. W., Torrisi-Steele, G. & Wang, V. (2018). Addressing survey research. In *Handbook of research on innovative techniques, trends, and analysis for optimized research methods* (pp. 341–359). IGI Global.

Brewer, K. C., Oh, K. M., Kitsantas, P. & Zhao, X. (2020). Workplace bullying among nurses and organizational response: An online cross–sectional study. *Journal of Nursing Management*, *28*(1), 148–156. https://doi.org/10.1111/jonm.12908

Brock, R. L., Barry, R. A., Lawrence, E., Dey, J. & Rolffs, J. (2012). Internet administration of paper-and-pencil questionnaires used in couple research: Assessing psychometric equivalence. *Assessment*, *19*(2), 226–242. https://doi.org/10.1177/1073191110382850

Browne, K. (2005). Snowball sampling: Using social networks to research non–heterosexual women. *International Journal of Social Research Methodology*, *8*(1), 47–60. https://doi.org/10.1080/1364557032000081663

Brühlmann, F., Petralito, S., Aeschbach, L. F. & Opwis, K. (2020). The quality of data collected online: An investigation of careless responding in a crowdsourced sample. *Methods in Psychology*, *2*, 100022. https://doi.org/10.1016/j.metip.2020.100022

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*(1), 3–5. https://doi.org/10.1177/1745691610393980

Callegaro, M. (2010). Do you know which device your respondent has used to take your online survey. *Survey Practice*, *3*(6), 1–12. https://doi.org/10.29115/SP-2010-0028

Callegaro, M., Manfreda, K. L. & Vehovar, V. (2015). *Web survey methodology*. Sage.

Callegaro, M., Villar, A., Yeager, D. & Krosnick, J. A. (2014). A critical review of studies investigating the quality of data obtained with online panels based on probability and nonprobability samples. In Callegaro, M., Baker, R., Bethlehem, J., Göritz, A. S., Krosnick, J. A. & Lavrakas, P. J. (Eds.), *Online panel research: A data quality perspective* (pp. 23–53). John Wiley & Sons, Ltd.

Casler, K., Bickel, L. & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, *29*(6), 2156–2160. https://doi.org/10.1016/j.chb.2013.05.009

Chandler, J., Mueller, P. & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods*, *46*(1), 112–130. https://doi.org/10.3758/s13428-013-0365-7

Chandler, J. & Shapiro, D. (2016). Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology*, *12*(1), 53–81. https://doi.org/10.1146/annurev-clinpsy-021815-093623

Chen, H. & Tseng, C. H. (2017). Effects of applying multimedia and dialogue boxes to web survey design. *International Journal of Business and Information*, *12*(1), 65–82. https://doi.org/10.6702/ijbi.2017.12.1.3

Chmielewski, M. & Kucker, S. C. (2020). An MTurk crisis? Shifts in data quality and the impact on study results. *Social Psychological and Personality Science*, *11*(4), 464–473. https://doi.org/10.1177/1948550619875149

Christian, L. M., Dillman, D. A. & Smyth, J. D. (2007). Helping respondents get it right the first time: The influence of words, symbols, and graphics in web surveys. *Public Opinion Quarterly*, *71*(1), 113–125. https://doi.org/10.1093/poq/nfl039

Clement, J. (2020). Mobile internet usage worldwide. Annual Report. Online Publication.

Clifford, S. & Jerit, J. (2016). Cheating on political knowledge questions in online surveys. *Public Opinion Quarterly*, *80*(4), 858–887. https://doi.org/10.1093/poq/nfw030

Conrad, F. G., Couper, M. P., Tourangeau, R. & Zhang, C. (2017, April). Reducing speeding in web surveys by providing immediate feedback. *In Survey Research Methods*, *11*(1), 45–61. https://doi.org/10.18148/srm/2017.v11i1.6304

Cornesse, C. & Blom, A. G. (2020). Response quality in nonprobability and probability-based online panels. *Sociological Methods & Research*, Advance online publication. https://doi.org/10.1177/0049124120914940

Coryn, C. L., Becho, L. W., Westine, C. D., Mateu, P. F., Abu-Obaid, R. N., Hobson, K. A. & Ramlow, M. (2020). Material incentives and other potential factors associated with response rates to internet surveys of American evaluation association members: Findings from a

randomized experiment. *American Journal of Evaluation*, *41*(2), 277–296. https://doi.org/10.1177/1098214018818371

Coşkun, B. K. & Dirlik, E. M. (2022). What prevents you to plug in to online surveys? In *Handbook of research on advanced research methodologies for a digital society* (pp. 131–148). IGI Global.

Couper, M. P. (2007). Technology and the survey interview/questionnaire. In Conrad, F. G. & Schober, M. F. (Eds.), *Envisioning the survey interview of the future* (pp. 58–76). Wiley Series in Survey Methodology.

Couper, M. P., Conrad, F. G. & Tourangeau, R. (2007). Visual context effects in web surveys. *Public Opinion Quarterly*, *71*(4), 623–634. https://doi.org/10.1093/poq/nfm044

Danilova, A., Naiakshina, A., Horstmann, S. & Smith, M. (2021, May). Do you really code? Designing and evaluating screening questions for online surveys with programmers. In *2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE)* (pp. 537–548).

Data Ethics Canvas (2021). Open Data Institute. Retrieved January 5, 2022, from https://theodi.org/article/the-data-ethics-canvas-2021/.

de Bruijne, M. & Wijnant, A. (2014). Improving response rates and questionnaire design for mobile web surveys. *Public Opinion Quarterly*, *78*(4), 951–962. https://doi.org/10.1093/poq/nfu046

DeCelles, K. A., Sonenshein, S. & King, B. G. (2020). Examining anger's immobilizing effect on institutional Insiders' action intentions in social movements. *Administrative Science Quarterly*, *65*(4), 847–886. https://doi.org/10.1177/0001839219879646

de Leeuw, E. D., Hox, J. J. & Boeve, A. (2015). Handling do-not-know answers: Exploring new approaches in online and mixed-mode surveys. *Social Science Computer Review*, *34*(1), 116–132. https://doi.org/10.1177/0894439315573744

de Leeuw, E. D. & Toepoel, V. (2018). Mixed-mode and mixed-device surveys. In *The palgrave handbook of survey research* (pp. 51–61). Palgrave Macmillan.

de Leeuw, J. R. (2015). Jspsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Delice, F., Rousseau, M. & Feitosa, J. (2019). Advancing teams research: What, when, and how to measure team dynamics over time. *Frontiers in Psychology*, *10*, 1–20. https://doi.org/10.3389/fpsyg.2019.01324

Denissen, J. J., Neumann, L. & Van Zalk, M. (2010). How the internet is changing the implementation of traditional research methods, people's daily lives, and the way in which developmental scientists conduct research. *International Journal of Behavioral Development*, *34*(6), 564–575. https://doi.org/10.1177/0165025410383746

Deutskens, E., Ruyter, K. d., Wetzels, M. & Oosterveld, P. (2004). Response rate and response quality of surveys: An experimental study. *Marketing Letters*, *15*(1), 21–36. https://doi.org/10.1023/B:MARK.0000021968.86465.00

Dillman, D. A., Smyth, J. D. & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. John Wiley & Sons.

Dimmery, D., Bakshy, E. & Sekhon, J. (2019). Shrinkage estimators in online experiments. *25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, *KDD '19*, 2914–2922. https://doi.org/10.1145/3292500.3330771

Dixon, J. & Tucker, C. (2010). Survey nonresponse. In Marsden, P. V. & Wright, J. D. (Eds.), *Handbook of survey research* (pp. 593–631). Emerald Group Publishing Limited.

Dolnicar, S., Grün, B. & Yanamandram, V. (2013). Dynamic, interactive survey questions can increase survey data quality. *Journal of Travel & Tourism Marketing*, *30*(7), 690–699. https://doi.org/10.1080/10548408.2013.827546

Donohoe, H., Stellefson, M. & Tennant, B. (2012). Advantages and limitations of the e-Delphi technique: Implications for health education researchers. *American Journal of Health Education*, *43*(1), 38–46. https://doi.org/10.1080/19325037.2012.10599216

Dusek, G. A., Yurova, Y. V. & Ruppel, C. P. (2015). Using social media and targeted snowball sampling to survey a hard-to-reach population: A case study. *International Journal of Doctoral Studies*, *10*, 279–299. https://doi.org/10.28945/2296

Eaton, J. & Struthers, C. W. (2002). Using the internet for organizational research: A study of cynicism in the workplace. *CyberPsychology & Behavior*, *5*(4), 305–313. https://doi.org/10.1089/109493102760275563

Ehlers, C., Greene-Shortridge, T. M., Weekley, J. A. & Zajack, M. D. (2009). The exploration of statistical methods in detecting random responding. *Paper Presented at the Annual Meeting of the Society for Industrial/Organizational Psychology*. Atlanta, GA.

Erens, B., Collins, D., Manacorda, T., Gosling, J., Mays, N., Reid, D. & Taylor, W. (2019). Comparing data quality from personal computers and mobile devices in an online survey among professionals. *Social Research Practice*, *7*, 15–26.

Fan, W. & Yan, Z. (2010). Factors affecting response rates of the web survey: A systematic review. *Computers in Human Behavior*, *26*(2), 132–139. https://doi.org/10.1016/j.chb.2009.10.015

Faran, Y. & Zanbar, L. (2019). Do required fields in online surveys in the social sciences impair reliability? *International Journal of Social Research Methodology*, *22*(6), 637–649. https://doi.org/10.1080/13645579.2019.1630899

Fazekas, Z., Wall, M. T. & Krouwel, A. (2014). Is it what you say, or how you say it? An experimental analysis of the effects of invitation wording for online panel surveys. *International Journal of Public Opinion Research*, *26*(2), 235–244. https://doi.org/10.1093/ijpor/edt022

Finley, A. & Penningroth, S. (2015). Online versus in-lab: Pros and cons of an online prospective memory experiment. In *Advances in psychology research,* (Vol. 113, pp. 135–161). Nova Science Publishers, Inc.

Fisher, C. D. & To, M. L. (2012). Using experience sampling methodology in organizational behavior. *Journal of Organizational Behavior*, *33*(7), 865–877. https://doi.org/10.1002/job.1803

Fisher, G. & Aguinis, H. (2017). Using theory elaboration to make theoretical advancements. *Organizational Research Methods*, *20*(3), 438–464. https://doi.org/10.1177/1094428116689707

Flatau Harrison, H., Griffin, M. A., Gagne, M. & Andrei, D. (2018). Assessing shortened safety climate measures: Simulating a planned missing data design in a field setting. *Safety Science*, *104*, 189–201. https://doi.org/10.1016/j.ssci.2017.11.004

Fleming, C. M. & Bowden, M. (2009). Web-based surveys as an alternative to traditional mail methods. *Journal of Environmental Management*, *90*(1), 284–292. https://doi.org/10.1016/j.jenvman.2007.09.011

Galesic, M. & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, *73*(2), 349–360. https://doi.org/10.1093/poq/nfp031

Gao, Z., House, L. A. & Xie, J. (2015). Online survey data quality and its implication for willingness-to-pay: A cross-country comparison. *Canadian Journal of Agricultural Economics/Revue Canadienne D'agroeconomie*, *64*(2), 199–221. https://doi.org/10.1111/cjag.12069

Garaizar, P. & Reips, U. D. (2019). Best practices: Two web-browser-based methods for stimulus presentation in behavioral experiments with high-resolution timing requirements. *Behavior Research Methods*, *51*(3), 1441–1453. https://doi.org/10.3758/s13428-018-1126-4

Giamattei, M., Yahosseini, K. S., Gächter, S. & Molleman, L. (2020). LIONESS Lab: A free web-based platform for conducting interactive experiments online. *Journal of the Economic Science Association*, *6*(1), 95–111. https://doi.org/10.1007/s40881-020-00087-0

González-Romá, V. & Hernández, A. (2017). Multilevel modeling: Research-based lessons for substantive researchers. *Annual Review of Organizational Psychology and Organizational Behavior*, *4*(1), 183–210. https://doi.org/10.1146/annurev-orgpsych-041015-062407

Goodman, J. K. & Paolacci, G. (2017). Crowdsourcing consumer research. *Journal of Consumer Research*, *44*(1), 196–210. https://doi.org/10.1093/jcr/ucx047

Göritz, A. S. (2005). Contingent versus unconditional incentives in WWW-studies. *Advances in Methodology & Statistics*, *2*(1), 1–14. https://doi.org/10.51936/kljg4387

Göritz, A. S. (2006). Incentives in web studies: Methodological issues and a review. *International Journal of Internet Science*, *1*(1), 58–70.

Göritz, A. S. (2010). Using lotteries, loyalty points, and other incentives to increase participant response and completion. In S. D. Gosling & J. A . Johnson (Eds.), *Advanced methods for conducting online behavioural research* (pp. 219–233). American Psychological Association.

Göritz, A. S. & Luthe, S. C. (2013). Effects of lotteries on response behavior in online panels. *Field Methods*, *25*(3), 219–237. https://doi.org/10.1177/1525822X12472876

Göritz, A. S. & Neumann, B. P. (2016). The longitudinal effects of incentives on response quantity in online panels. *Translational Issues in Psychological Science*, *2*(2), 163–173. https://doi.org/10.1037/tps0000071

Gosling, S. D. & Mason, W. (2015). Internet research in psychology. *Annual Review of Psychology*, *66*(1), 877–902. https://doi.org/10.1146/annurev-psych-010814-015321

Gosling, S. D., Vazire, S., Srivastava, S. & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *The American Psychologist*, *59*(2), 93–104. https://doi.org/10.1037/0003-066X.59.2.93

Graham, J. W. (2012). Planned missing data designs I: The 3-form design. In *Missing data* (pp. 279–294). Springer.

Granello, D. H. & Wheaton, J. E. (2004). Online data collection: Strategies for research. *Journal of Counseling & Development*, *82*, 387–339. https://doi.org/10.1002/j.1556-6678.2004.tb00325.x

Greszki, R., Meyer, M. & Schoen, H. (2014). The impact of speeding on data quality in nonprobability and freshly recruited probability-based online panels. In Callegaro, M., Baker, R., Bethlehem, J., Göritz, A. S., Krosnick, J. A. & Lavrakas, P. J. (Eds.), *Online panel research: A data quality perspective* (pp. 238–262). John Wiley & Sons, Ltd.

Griffin, M., Martino, R. J., LoSchiavo, C., Comer-Carruthers, C., Krause, K. D., Stults, C. B. & Halkitis, P. N. (2021). Ensuring survey research data integrity in the era of internet bots. *Quality & Quantity*, Advance online publication, 1–12. https://doi.org/10.1007/s11135-021-01252-1

Griffiths, M. D., Lewis, A. M., Ortiz de Gortari, A. B. & Kuss, D. J. (2014). Online forums and solicited blogs: Innovative methodologies for online gaming data collection. *Studia Psychologica*, *14*(3), 5–24. https://doi.org/10.21697/sp.2015.14.2.07

Gross, M. S., Liu, N. H., Contreras, O., Muñoz, R. F. & Leykin, Y. (2014). Using google AdWords for international multilingual recruitment to health research websites. *Journal of Medical Internet Research*, *16*(1), e18. https://doi.org/10.2196/jmir.2986

Gupta, S. (2017). Ethical issues in designing internet-based research: Recommendations for good practice. *Journal of Research Practice*, *13*(2), D1–D1.

Halbesleben, J. R. & Buckley, M. R. (2004). Burnout in organizational life. *Journal of Management*, *30*(6), 859–879. https://doi.org/10.1016/j.jm.2004.06.004

Hargittai, E. & Shaw, A. (2020). Comparing internet experiences and prosociality in Amazon Mechanical Turk and population-based survey samples. *Socius*, Advance online publication, https://doi.org/10.1177/2378023119889834

Heerwegh, D. (2006). An investigation of the effect of lotteries on web survey response rates. *Field Methods*, *18*(2), 205–220. https://doi.org/10.1177/1525822X05285781

Henrich, J., Heine, S. J. & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Heretick, D. M. & Learn, I. (2020). Severity of coercive sexual harassment in professor–student interaction and peer bystander responses. *Journal of Social Behavioral, and Health Sciences*, *14*(1), 87–105. https://doi.org/10.5590/JSBHS.2020.14.1.07

Hewson, C. (2017). Research design and tools for internet research. In *The SAGE handbook of online research methods* (pp. 58–78). Sage.

Hewson, C., Yule, P., Laurent, D. & Vogel, C. (2003). *Internet research methods: A practical guide for the social and behavioural sciences*. Sage Publications, Ltd.

Hilbig, B. E. & Thielmann, I. (2021). On the (mis) use of deception in web-based research.

Zeitschrift für Psychologie, *229*(4), 225–229. https://doi.org/10.1027/2151-2604/a000466

Horton, J. J., Rand, D. G. & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, *14*(3), 399–425. https://doi.org/10.1007/s10683-011-9273-9

Hou, J., Zhu, Y., Tong, J., Ndasauka, Y. & Fang, X. (2020). Preparation of multi-motive grid questionnaire for social networking sites use. *PloS one*, *15*(5), e0233205. https://doi.org/10.1371/journal.pone.0233205

Hox, J. (2002). *Quantitative methodology series. Multilevel analysis techniques and applications*. Lawrence Erlbaum Associates Publishers.

Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M. & DeShon, R. P. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology*, *27*(1), 99–114. https://doi.org/10.1007/s10869-011-9231-8

Ipsen, C., Kurth, N. & Hall, J. (2021). Evaluating MTurk as a recruitment tool for rural people with disabilities. *Disability and Health Journal*, *14*(1), 100991. https://doi.org/10.1016/j.dhjo.2020.100991

Ismail, S. N. M. S., Aziz, N. A. A., Ibrahim, S. Z., Khan, C. T. & Rahman, M. A. (2021). Selecting video stimuli for emotion elicitation via online survey. *Human-Centric Computing and Information Sciences*, *11*(36), 1–18. https://doi.org/10.22967/HCIS.2021.11.036

James, N. & Busher, H. (2015). Ethical issues in online research. *Educational Research and Evaluation*, *21*(2), 89–94. https://doi.org/10.1080/13803611.2015.1024420

Jia, F., Moore, E. W. G., Kinai, R., Crowe, K. S., Schoemann, A. M. & Little, T. D. (2014). Planned missing data designs with small sample sizes: How small is too small? *International Journal of Behavioral Development*, *38*(5), 435–452. https://doi.org/10.1177/0165025414531095

Johnson, J. A. (2005). Ascertaining the validity of individual protocols from web-based personality inventories. *Journal of Research in Personality*, *39*(1), 103–129. https://doi.org/10.1016/j.jrp.2004.09.009

Joinson, A. N. & Reips, U.-D. (2007). Personalized salutation, power of sender and response rates to web-based surveys. *Computers in Human Behavior*, *23*(3), 1372–1383. https://doi.org/10.1016/j.chb.2004.12.011

Jones, M. S., House, L. A. & Gao, Z. (2015). Respondent screening and revealed preference axioms: Testing quarantining methods for enhanced data quality in web panel surveys. *Public Opinion Quarterly*, *79*(3), 687–709. https://doi.org/10.1093/poq/nfv015

Jorgensen, T. D., Rhemtulla, M., Schoemann, A., McPherson, B., Wu, W. & Little, T. D. (2014). Optimal assignment methods in three-form planned missing data designs for longitudinal panel studies. *International Journal of Behavioral Development*, *38*(5), 397–410. https://doi.org/10.1177/0165025414531094

Joshi, H. (2016). Why do we need longitudinal survey data? *IZA World of Labor* (November).

Jun, E., Jo, B. A., Oliveira, N. & Reinecke, K. (2018). Digestif: Promoting science communication in online experiments. *Proceedings of the ACM on Human-Computer Interaction*, *2*(84), 1–26. https://doi.org/10.1145/3274353

Kalleitner, F., Mühlböck, M. & Kittel, B. (2020). What's the benefit of a video? The effect of non-material incentives on response rate and bias in web surveys. *Social Science Computer Review*, Advance online publication. https://doi.org/10.1177/0894439320918318

Keith, M., Shao, B. & Steinbart, P. J. (2007). The usability of passphrases for authentication: An empirical field study. *International Journal of Human-Computer Studies*, *65*(1), 17–28. https://doi.org/10.1016/j.ijhcs.2006.08.005

Keller, F., Gunasekharan, S., Mayo, N. & Corley, M. (2009). Timing accuracy of web experiments: A case study using the WebExp software package. *Behavior Research Methods*, *41*(1), 1–12. https://doi.org/10.3758/BRM.41.1.12

Kennedy, R., Clifford, S., Burleigh, T., Waggoner, P. D., Jewell, R. & Winter, N. (2019). The shape of and solutions to the MTurk quality crisis. *Political Science Research and Methods*, *8*(4), 614–629. https://doi.org/10.1017/psrm.2020.6

Kim, S., Cho, S. & Park, Y. (2021). Daily micro-breaks in a self-regulatory resources lens: Perceived health climate as a contextual moderator via microbreak autonomy. *Journal of Applied Psychology*, *107*(1), 60–77. https://doi.org/10.1037/apl0000891

Kim, S., Lee, J. & Gweon, G. (2019, May). Comparing data from chatbot and web surveys: Effects of platform and conversational style on survey response quality. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, *86*, 1–12. https://doi.org/10.1145/3290605.3300316

Knowles, S. & Stahlmann-Brown, P. (2021). Cash is not king in incentivizing online surveys. *Applied Economics Letters*, *28*(2), 105–108. https://doi.org/10.1080/13504851.2020.1734524

Kochari, A. R. (2019). Conducting web-based experiments for numerical cognition research. *Journal of Cognition*, *2*(1), 1–21. https://doi.org/10.5334/joc.85

Kohavi, R., Tang, D. & Xu, Y. (2020). *Trustworthy online controlled experiments: A practical guide to a/b testing*. Cambridge University Press.

Korman, J., Van Quaquebeke, N. & Tröster, C. (2021). Managers are less burned-out at the top: The roles of sense of power and self-efficacy at different hierarchy levels. *Journal of Business and Psychology*, *37*, 151–171. https://doi.org/10.1007/s10869-021-09733-8

Kraut, R., Olson, J., Banaji, M. R., Bruckman, A., Cohen, J. & Couper, M. P. (2004). Psychological research online: Report of board of scientific affairs' advisory group on the conduct of research on the internet. *American Psychologist*, *59*(2), 105–117. https://doi.org/10.1037/0003-066X.59.2.105

Krosnick, J. A., Holbrook, A. L., Berent, M. K., Carson, R. T., Hanemann, M. W., Kopp, R. J., Mitchell, R. C., Presser, S., Ruud, P. A., Smith, K. V., Moody, W. R., Green, M. C. & Conaway, M. (2002). The impact of 'No opinion' response options on data quality: Non-attitude reduction or an invitation to satisfice? *Public Opinion Quarterly*, *66*, 371–403. https://doi.org/10.1086/341394

Kühne, S. & Kroh, M. (2018). Personalized feedback in web surveys: Does it affect respondents' motivation and data quality? *Social Science Computer Review*, *36*(6), 744–755. https://doi.org/10.1177/0894439316673604

Kulik, C. T. (2011). Climbing the higher mountain: The challenges of multilevel, multisource, and longitudinal research designs. *Management and Organization Review*, *7*(3), 447–460. https://doi.org/10.1111/j.1740-8784.2011.00226.x

Kung, F. Y. H., Kwok, N. & Brown, D. J. (2018). Are attention check questions a threat to scale validity? *Applied Psychology*, *67*(2), 264–283. https://doi.org/10.1111/apps.12108

Kunz, T., Landesvatter, C. & Gummer, T. (2020). Informed consent for paradata use in web surveys. *International Journal of Market Research*, *62*(4), 396–408. https://doi.org/10.1177/1470785320931669

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. https://doi.org/10.1017/iop.2015.13

Lavan, N., Burston, L. F. & Garrido, L. (2019). How many voices did you hear? Natural variability disrupts identity perception from unfamiliar voices. *British Journal of Psychology*, *110*(3), 576–593. https://doi.org/10.1111/bjop.12348

Leung, C. B. & Unal, Z. (2013). Anonymity and confidentiality in the conduct of online surveys. In C. S. Rhodes & K. Weiss (Eds.), Ethical issues in literacy research. (pp. 129–138). New York: Routledge.

Lewis, T. & Hess, K. (2017). The effect of alternative e-mail contact timing strategies on response rates in a self-administered web survey. *Field Methods*, *29*(4), 351–364. https://doi.org/10.1177/1525822X17715865

Lindsey, A. P., Avery, D. R., Dawson, J. F. & King, E. B. (2017). Investigating why and for whom management ethnic representativeness influences interpersonal mistreatment in the workplace. *Journal of Applied Psychology*, *102*(11), 1545–1563. https://doi.org/10.1037/apl0000238

Litman, L., Robinson, J. & Abberbock, T. (2017). Turkprime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences.

Behavior Research Methods, *49*(2), 433–442. https://doi.org/10.3758/s13428-016-0727-z

Lu, L., Li, F., Leung, K., Savani, K. & Morris, M. W. (2018). When can culturally diverse teams be more creative? The role of leaders' benevolent paternalism. *Journal of Organizational Behavior*, *39*(4), 402–415. https://doi.org/10.1002/job.2238

Lucas, R. E., Wallsworth, C., Anusic, I. & Donnellan, M. B. (2021). A direct comparison of the day reconstruction method (DRM) and the experience sampling method (ESM). *Journal of Personality and Social Psychology*, *120*(3), 816–835. https://doi.org/10.1037/pspp0000289

Luce, K. H., Winzelberg, A. J., Das, S., Osborne, M. I., Bryson, S. W. & Taylor, C. B. (2007). Reliability of self-report: Paper versus online administration. *Computers in Human Behavior*, *23*(3), 1384–1389. https://doi.org/10.1016/j.chb.2004.12.008

Lučić, L., Brkljačić, T. & Lipovčan, L. K. (2018). Longitudinal online survey in well-being research: First experiences from CRO-WELL project. *ILIRIA International Review*, *8*(1), 205–227. https://doi.org/10.21113/iir.v8i1.392

Lüdtke, O., Robitzsch, A. & Grund, S. (2017). Multiple imputation of missing data in multilevel designs: A comparison of different strategies. *Psychological Methods*, *22*(1), 141–165. https://doi.org/10.1037/met0000096

Macik-Frey, M., Quick, J. C. & Nelson, D. L. (2007). Advances in occupational health: From a stressful beginning to a positive future. *Journal of Management*, *33*(6), 809–840. https://doi.org/10.1177/0149206307307634

Magalhães, A., dos Santos, N. R. & Pais, L. (2019). Multi-source research designs on ethical leadership: A literature review. *Business and Society Review*, *124*(3), 345–364. https://doi.org/10.1111/basr.12179

Mahon-Haft, T. A. & Dillman, D. A. (2010, May). Does visual appeal matter? Effects of web survey aesthetics on survey quality. *Survey Research Methods*, *4*(1), 43–59. 10.18148/srm/2010.v4i1.2264

Marcus, B., Bosnjak, M., Lindner, S., Pilischenko, S. & Schütz, A. (2007). Compensating for low topic interest and long surveys: A field experiment on nonresponse in web surveys. *Social Science Computer Review*, *25*(3), 372–383. https://doi.org/10.1177/0894439307297606

Marcus, B., Weigelt, O., Hergert, J., Gurt, J. & Gelléri, P. (2017). The use of snowball sampling for multi source organizational research: Some cause for concern. *Personnel Psychology*, *70*(3), 635–673. https://doi.org/10.1111/peps.12169

Martin, S. R., Côté, S. & Woodruff, T. (2016). Echoes of our upbringing: How growing up wealthy or poor relates to narcissism, leader behavior, and leader effectiveness. *Academy of Management Journal*, *59*(6), 2157–2177. https://doi.org/10.5465/amj.2015.0680

Mason, W. & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, *44*(1), 1–23. https://doi.org/10.3758/s13428-011-0124-6

Mathieu, J. E., Aguinis, H., Culpepper, S. A. & Chen, G. (2012). Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. *Journal of Applied Psychology*, *97*(5), 951–966. https://doi.org/10.1037/a0028380

Mavletova, A. & Couper, M. P. (2014). Mobile web survey design: Scrolling versus paging, SMS versus e-mail invitations. *Journal of Survey Statistics and Methodology*, *2*(4), 498–518. https://doi.org/10.1093/jssam/smu015

Mayer, D. M., Thau, S., Workman, K. M., Dijke, M. v. & Cremer, D. d. (2012). Leader mistreatment, employee hostility, and deviant behaviors: Integrating self-uncertainty and thwarted needs perspectives on deviance. *Organizational Behavior and Human Decision Processes*, *117*(1), 24–40. https://doi.org/10.1016/j.obhdp.2011.07.003

McCambridge, J., Kypri, K. & Wilson, A. (2012). How should debriefing be undertaken in web-based studies? Findings from a randomized controlled trial. *Journal of Medical Internet Research*, *14*(6), e157. https://doi.org/10.2196/jmir.2186

McInroy, L. B. (2016). Pitfalls, potentials, and ethics of online survey research: LGBTQ and other marginalized and hard-to-access youths. *Social Work Research*, *40*(2), 83–94. https://doi.org/10.1093/swr/svw005

McInroy, L. B. (2017). Innovative ethics: Using animated videos when soliciting informed consent of young people for online surveys. *Social Work Research*, *41*(2), 121–128. https://doi.org/10.1093/swr/svx004

Meade, A. W. & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, *17*(3), 437–455. https://doi.org/10.1037/a0028085

Meuleman, B., Langer, A. & Blom, A. G. (2018). Can incentive effects in web surveys be generalized to non-western countries? Conditional and unconditional cash incentives in a web survey of Ghanaian university students. *Social Science Computer Review*, *36*(2), 231–250. https://doi.org/10.1177/0894439317699577

Miller, J. D. & Baker-Prewitt, J. (2009). Beyond 'trapping' the undesirable panelist: The use of red herrings to reduce satisficing. *CASRO Panel Quality Conference*, *2009*(CP51).

Miner, M. H., Bockting, W. O. & Swinburne, R. (2012). Conducting internet research with the transgender population: Reaching broad samples and collecting valid data. *Social Science Computer Review*, *30*(2), 202–211. https://doi.org/10.1177/0894439311404795

Molina-Azorín, J. F., Pereira-Moliner, J., López-Gamero, M. D., Pertusa-Ortega, E. M. & Tarí, J. J. (2019). Multilevel research: Foundations and opportunities in management. *BRQ Business Research Quarterly*, *23*(4), 319–333. https://doi.org/10.1016/j.brq.2019.03.004

Molnar, A. (2019). SMARTRIQS: A simple method allowing real-time respondent interaction in qualtrics surveys. *Journal of Behavioral and Experimental Finance*, *22*, 161–169. https://doi.org/10.1016/j.jbef.2019.03.005

Nayak, M. S. D. P. & Narayan, K. A. (2019). Strengths and weaknesses of online surveys. *Journal of Humanities and Social Sciences*, *24*(5), 31–38. https://doi.org/10.9790/0837-2405053138

Nestler, S., Thielsch, M., Vasilev, E. & Back, M. D. (2015). Will they stay or will they go? Personality predictors of dropout in an online study. *International Journal of Internet Science*, *10*(1), 37–48.

Newman, A., Bavik, Y. L., Mount, M. & Shao, B. (2021). Data collection via online platforms: Challenges and recommendations for future research. *Applied Psychology*, *70*(3), 1380–1402. https://doi.org/10.1111/apps.12302

Niemann, J. & Schenk, M. (2014). Online experience sampling method: Measuring user behavior combining tracking and survey methodology. Conference presentation. University of Hohenheim (Issue April).

Oppenheimer, D. M., Meyvis, T. & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867–872. https://doi.org/10.1016/j.jesp.2009.03.009

Paolacci, G., Chandler, J. & Ipeirotis, P. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*(5), 411–419. https://doi.org/10.1037/t69659-000

Pedersen, M. J. & Nielsen, C. V. (2014). Improving survey response rates in online panels: Effects of low-cost incentives and cost-free text appeal interventions. *Social Science Computer Review*, *34*(2), 1–15. https://doi.org/10.1177/0894439314563916

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. https://doi.org/10.1016/j.jesp.2017.01.006

Petriglieri, G., Ashford, S. J. & Wrzesniewski, A. (2019). Agony and ecstasy in the gig economy: Cultivating holding environments for precarious and personalized work identities. *Administrative Science Quarterly*, *64*(1), 124–170. https://doi.org/10.1177/0001839218759646

Pittenger, D. J. (2003). Internet research: An opportunity to revisit classic ethical problems in behavioral research. *Ethics and Behavior*, *13*(1), 45–60. https://doi.org/10.1207/S15327019EB1301_08

Porter, C., Outlaw, R., Gale, J. P. & Cho, T. S. (2019). The use of online panel data in management research: A review and recommendations. *Journal of Management*, *45*(1), 319–344. https://doi.org/10.1177/0149206318811569

Porter, S. R. & Whitcomb, M. E. (2007). Mixed-mode contacts in web surveys: Paper is

not necessarily better. *Public Opinion Quarterly*, *71*(4), 635–648. https://doi.org/10.1093/poq/nfm038

Prince, K. R., Litovsky, A. R. & Friedman-Wheeler, D. G. (2012). Internet-mediated research: Beware of bots. *The Behavior Therapist*, *35*(5), 85–88.

Questback (2013). Enterprise feedback suite EFS survey. In User Manual for EFS Survey.

Ramo, D. E. & Prochaska, J. J. (2012). Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. *Journal of Medical Internet Research*, *14*, e28. https://doi.org/10.2196/jmir.1878

Recklitis, C. J., Campbell, E. G., Kutner, J. S. & Bober, S. L. (2009). Money talks: Non-monetary incentive and internet administration fail to increase response rates to a physician survey. *Journal of Clinical Epidemiology*, *62*(2), 224–226. https://doi.org/10.1016/j.jclinepi.2008.09.005

Reh, S., Tröster, C. & van Quaquebeke, N. (2018). Keeping (future) rivals down: Temporal social comparison predicts coworker social undermining via future status threat and envy. *Journal of Applied Psychology*, *103*(4), 399–415. https://doi.org/10.1037/apl0000281

Reips, U.-D. (2002). Standards for internet-based experimenting. *Experimental Psychology*, *49*(4), 243–256. https://doi.org/10.1026/1618-3169.49.4.243

Reips, U.-D. (2010). Design and formatting in internet-based research. In *Advanced methods for conducting online behavioral research* (pp. 29–43). American Psychological Association. https://doi.org/10.1037/12076-003

Reips, U.-D. (2013). Internet-based studies. In *Encyclopedia of behavioral medicine* (pp. 1097–1102). Springer New York.

Reips, U.-D. & Krantz, J. H. (2010). Conducting true experiments on the web. *Advanced Methods for Conducting Online Behavioral Research*, 193–216. https://doi.org/10.1037/12076-013

Rennung, M., Blum, J. & Göritz, A. S. (2016). To strike a pose: No stereotype backlash for power posing women. *Frontiers in Psychology*, *7*, 1463. https://doi.org/10.3389/fpsyg.2016.01463

Revilla, M., Couper, M. P., Bosch, O. J. & Asensio, M. (2020). Testing the use of voice input in a smartphone web survey. *Social Science Computer Review*, *38*(2), 207–224. https://doi.org/10.1177/0894439318810715

Reyes, D. L. (2020). Combatting carelessness: Can placement of quality check items help reduce careless responses? *Current Psychology*, Advance online publication, 1–9. https://doi.org/10.1007/s12144-020-01183-4

Richter, M., König, C. J., Geiger, M., Schieren, S., Lothschütz, J. & Zobel, Y. (2018). "Just a little respect": Effects of a layoff agent's actions on employees' reactions to a dismissal notification meeting. *Journal of Business Ethics*, *153*(3), 741–761. https://doi.org/10.1007/s10551-016-3372-7

Riediger, M.. (2009). Experience sampling. German Council for Social and Economic Data, No 62. Online Publication.

Roberts, L. D. & Allen, P. J. (2015). Exploring ethical issues associated with using online surveys in educational research. *Educational Research and Evaluation*, *21*(2), 95–108. https://doi.org/10.1080/13803611.2015.1024421

Rowbotham, M. C., Astin, J., Greene, K. & Cummings, S. R. (2013). Interactive informed consent: Randomized comparison with paper consents. *PloS one*, *8*(3), e58603. https://doi.org/10.1371/journal.pone.0058603

Saari, L. M. & Scherbaum, C. A. (2020). From identified surveys to new technologies. In Macey, W. & Fink, A. (Eds.), *Employee surveys and sensing: Challenges and opportunities* (Vol. 391, pp. 391–407).

Salem, M., Van Quaquebeke, N., Besiou, M. & Meyer, L. (2019). Intergroup leadership: How leaders can enhance performance of humanitarian operations. *Production and Operations Management*, *28*(11), 2877–2897. https://doi.org/10.1111/poms.13085

Sammut, R., Griscti, O. & Norman, I. J. (2021). Strategies to improve response rates to web surveys: A literature review. *International Journal of Nursing Studies*, *123*, 1040–1058. https://doi.org/10.1016/j.ijnurstu.2021.104058

Sánchez-Fernández, J., Muñoz-Leiva, F. & Montoro-Ríos, F. J. (2012). Improving retention rate and response quality in web-based surveys. *Computers in Human Behavior*, *28*(2), 507–514. https://doi.org/10.1016/j.chb.2011.10.023

Sánchez-Fernández, J., Muñoz-Leiva, F., Montoro-Ríos, F. J. & Ibáñez-Zapata, JÁ (2010). An analysis of the effect of pre-incentives and post-incentives based on draws on response to web surveys. *Quality and Quantity*, *44*(2), 357–373. https://doi.org/10.1007/s11135-008-9197-4

Sauermann, H. & Roach, M. (2013). Increasing web survey response rates in innovation research: An experimental study of static and dynamic contact design features. *Research Policy*, *42*(1), 273–286. https://doi.org/10.1016/j.respol.2012.05.003

Sauter, M., Draschkow, D. & Mack, W. (2020). Building, hosting and recruiting: A brief introduction to running behavioral experiments online. *Brain Sciences*, *10*(4), 251. https://doi.org/10.3390/brainsci10040251

Scherbaum, C. A. & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods*, *12*(2), 347–367. https://doi.org/10.1177/1094428107308906

Schubert, T. W., Murteira, C., Collins, E. C. & Lopes, D. (2013). ScriptingRT: A software library for collecting response latencies in online studies of cognition. *PLoS ONE*, *8*(6), e67769. https://doi.org/10.1371/journal.pone.0067769

Schwarz, N. (2011). Why researchers should think "real-time.". In *Handbook of research methods for studying daily* (pp. 22–42). Guilford Press.

Schyns, B. & Day, D. (2010). Critique and review of leader–member exchange theory: Issues of agreement, consensus, and excellence. *European Journal of Work and Organizational Psychology*, *19*(1), 1–29. https://doi.org/10.1080/13594320903024922

Schyns, B., Felfe, J. & Schilling, J. (2018). Is it me or you?—how reactions to abusive supervision are shaped by leader behavior and follower perceptions. *Frontiers in Psychology*, *9*, 1309. https://doi.org/10.3389/fpsyg.2018.01309

Scott, B. A., Garza, A. S., Conlon, D. E. & Kim, Y. J. (2014). 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), 1571–1591. https://doi.org/10.5465/amj.2012.0644

Semmelmann, K., Hönekopp, A. & Weigelt, S. (2017). Looking tasks online: Utilizing webcams to collect video data from home. *Frontiers in Psychology*, *8*, 1582. https://doi.org/10.3389/fpsyg.2017.01582

Sessions, H., Nahrgang, J. D., Vaulont, M., Williams, R. & Bartels, A. L. (2020). Do the hustle! empowerment from side-hustles and its effects on full-time work performance. *Academy of Management Journal*, *64*(1), 235–264. https://doi.org/10.5465/amj.2018.0164

Shapiro-Luft, D. & Cappella, J. N. (2013). Video content in web surveys: Effects on selection bias and validity. *Public Opinion Quarterly*, *77*(4), 936–961. https://doi.org/10.1093/poq/nft043

Shaw, A. & Hargittai, E. (2021). Do the online activities of Amazon mechanical Turk workers mirror those of the general population? A comparison of two survey samples. *International Journal of Communication*, *15*(16), 4383–4398.

Sheehan, K. B. (2018). Crowdsourcing research: Data collection with Amazon's Mechanical Turk. *Communication Monographs*, *85*(1), 140–156. https://doi.org/10.1080/03637751.2017.1342043

Shropshire, K. O., Hawdon, J. E. & Witte, J. C. (2009). Web survey design: Balancing measurement, response, and topical interest. *Sociological Methods & Research*, *37*(3), 344–370. https://doi.org/10.1177/0049124108327130

Singh, A., Taneja, A. & Mangalaraj, G. (2009). Creating online surveys: Some wisdom from the trenches tutorial. *IEEE Transactions on Professional Communication*, *52*(2), 197–212. https://doi.org/10.1109/TPC.2009.2017986

Smyth, J. D., Dillman, D. A., Christian, L. M. & Mcbride, M. (2009). Open-ended questions in web surveys: Can increasing the size of answer boxes and providing extra verbal instructions improve response quality? *Public Opinion Quarterly*, *73*(2), 325–337. https://doi.org/10.1093/poq/nfp029

Spector, P. E. (2019). Do not cross me: Optimizing the use of cross-sectional designs. *Journal of Business and Psychology*, *34*(2), 125–137. https://doi.org/10.1007/s10869-018-09613-8

Spitz, G., Pepper, J., Chakravarti, V., Adler, T. & Niles, F. (2013). Using a web-based longitudinal panel to

measure customer satisfaction. *Transportation Research Record Journal of the Transportation Research Board*, *53*(9), 1689–1699. https://doi.org/10.3141/1887-19

Spreen, T. L., House, L. A. & Gao, Z. (2020). The impact of varying financial incentives on data quality in web panel surveys. *Journal of Survey Statistics and Methodology*, *8*(5), 832–850. https://doi.org/10.1093/jssam/smz030

Spurk, D., Volmer, J., Orth, M. & Göritz, A. S. (2020). How do career adaptability and proactive career behaviours interrelate over time? An inter- and intraindividual investigation. *Journal of Occupational and Organizational Psychology*, *93*(1), 158–186. https://doi.org/10.1111/joop.12288

Stanton, J. M. & Rogelberg, S. G. (2001). Using internet/intranet web pages to collect organizational research data. *Organizational Research Methods*, *4*(3), 200–217. https://doi.org/10.1177/109442810143002

Storozuk, A., Ashley, M., Delage, V. & Maloney, E. A. (2020). Got bots? Practical recommendations to protect online survey data from bot attacks. *The Quantitative Methods for Psychology*, *16*(5), 472–481. https://doi.org/10.20982/tqmp.16.5.p472

Su, J., Shao, P. & Fang, J. (2008). Effect of incentives on web-based surveys. *Tsinghua Science and Technology*, *13*(3), 344–347. https://doi.org/10.1016/S1007-0214(08)70055-5

Sue, V. M. & Ritter, L. A. (2012). Planning the online survey. In *Conducting online surveys* (pp. 107–138). Sage Publications Inc.

Taylor, S. G., Griffith, M. D., Vadera, A. K., Folger, R. & Letwin, C. R. (2019). Breaking the cycle of abusive supervision: How disidentification and moral identity help the trickle-down change course. *Journal of Applied Psychology*, *104*(1), 164–182. https://doi.org/10.1037/apl0000360

Tenney, E. R., Costa, E., Allard, A. & Vazire, S. (2021). Open science and reform practices in organizational behavior research over time (2011 to 2019). *Organizational Behavior and Human Decision Processes*, *162*, 218–223. https://doi.org/10.1016/j.obhdp.2020.10.015

Thau, S., Pitesa, M. & Pillutla, M. (2014). Experiments in organizational behavior. In Webster, M. & Sell, J. (Eds.), *Laboratory experiments in the social sciences* (Vol. Second Edition, pp. 433–447).

Thompson, L. & Surface, E. A. (2007). Employee surveys administered online: Attitudes toward the medium, nonresponse, and data representativeness. *Organizational Research Methods*, *10*(2), 241–261. https://doi.org/10.1177/1094428106294696

Toepoel, V. & Couper, M. P. (2011). Can verbal instructions counteract visual context effects in web surveys? *Public Opinion Quarterly*, *75*(1), 1–18. https://doi.org/10.1093/poq/nfq044

Toepoel, V. & Lugtig, P. (2015). Online surveys are mixed-device surveys. Issues associated with the use of different (mobile) devices in web surveys. *Methods, Data, Analyses*, *9*(2), 155–162. https://doi.org/10.12758/mda.2015.009

Troester, C. & Van Quaquebeke, N. (2020). When victims help their abusive supervisors: The role of LMX, self-blame, and guilt. *Academy of Management Journal*, *64*(6), 1793–1815. https://doi.org/10.5465/amj.2019.0559

Trübner, M. (2020). Effects of header images on different devices in web surveys. *Survey Research Methods*, *14*(1), 43–53. https://doi.org/10.18148/srm/2020.v14i1.7367

Valentijn, P. P., Vrijhoef, H. J. M., Ru waard, D., de Bont, A., Arends, R. Y. & Bruijnzeels, M. A. (2015). Exploring the success of an integrated primary care partnership: A longitudinal study of collaboration processes. *BMC Health Services Research*, *15*(1), 1–18. https://doi.org/10.1186/s12913-014-0652-8

Van Dijke, M., Leunissen, J. M., Wildschut, T. & Sedikides, C. (2019). Nostalgia promotes intrinsic motivation and effort in the presence of low interactional justice. *Organizational Behavior and Human Decision Processes*, *150*, 46–61. https://doi.org/10.1016/j.obhdp.2018.12.003

van Gils, S., van Quaquebeke, N. & van Knippenberg, D. (2010). The X-factor: On the relevance of implicit leadership and followership theories for leader–member exchange agreement. *European Journal of Work and Organizational Psychology*, *19*(3), 333–363. https://doi.org/10.1080/13594320902978458

Van Mol, C. (2017). Improving web survey efficiency: The impact of an extra reminder and reminder content on web survey response. *International Journal of Social Research Methodology*, *20*(4), 317–327. https://doi.org/10.1080/13645579.2016.1185255

Veen, F. v., Göritz, A. S. & Sattler, S. (2016). Response effects of prenotification, prepaid cash, prepaid vouchers, and postpaid vouchers: An experimental comparison. *Social Science Computer Review*, *34*(3), 333–346. https://doi.org/10.1177/0894439315585074

Vitak, J., Shilton, K. & Ashktorab, Z. (2016). Beyond the Belmont principles: Ethical challenges, practices, and beliefs in the online data research community. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, CSCW '16*, (February), 941–953. https://doi.org/10.1145/2818048.2820078

Vogel, R. M., Rodell, J. B. & Lynch, J. W. (2016). Engaged and productive misfits: How job crafting and leisure activity mitigate the negative effects of value incongruence. *Academy of Management Journal*, *59*(5), 1561–1584. https://doi.org/10.5465/amj.2014.0850

von Bastian, C. C., Locher, A. & Ruflin, M. (2013). Tatool: A Java-based open-source programming framework for psychological studies. *Behavior Research Methods*, *45*(1), 108–115. https://doi.org/10.3758/s13428-012-0224-y

Walston, J. T., Lissitz, R. W. & Rudner, L. M. (2006). The influence of web-based questionnaire presentation variations on survey cooperation and perceptions of survey quality. *Journal of Official Statistics*, *22*(2), 271–291.

Walter, S. L., Seibert, S. E., Goering, D. & O'Boyle, E. H. (2019). A tale of two sample sources: Do results from online panel data and conventional data converge? *Journal of Business and Psychology*, *34*, 425–452. https://doi.org/10.1007/s10869-018-9552-y

Wang, J. Y. & Kitsis, E. A. (2013). Tangling the web: Deception in online research. *The American Journal of Bioethics*, *13*(11), 59–61. https://doi.org/10.1080/15265161.2013.840868

Wang, M., Beal, D. J., Chan, D., Newman, D. A., Vancouver, J. B. & Vandenberg, R. J. (2017). Longitudinal research: A panel discussion on conceptual issues, research design, and statistical techniques. *Work, Aging and Retirement*, *3*(1), 1–24. https://doi.org/10.1093/workar/waw033

Ward, M. K. & Meade, A. W. (2018). Applying social psychology to prevent careless responding during online surveys. *Applied Psychology*, *67*(2), 231–263. https://doi.org/10.1111/apps.12118

Wenzel, R. & Van Quaquebeke, N. (2018). The double-edged sword of big data in organizational and management research: A review of opportunities and risks. *Organizational Research Methods*, *21*(3), 548–591. https://doi.org/10.1177/1094428117718627

Whelan, T. J. (2007). Anonymity and confidentiality: Do survey respondents know the difference? *30th Annual Meeting of the Society of Southeastern Social Psychologists*, (October), 1–11.

Wilson, A. C., Schwannauer, M., McLaughlin, A., Ashworth, F. & Chan, S. W. (2018). Vividness of positive mental imagery predicts positive emotional response to visually presented project soothe pictures. *British Journal of Psychology*, *109*(2), 259–276. https://doi.org/10.1111/bjop.12267

Wilt, J., Condon, D. M. & Revelle, W. (2012). Telemetrics and online data collection: Collecting data at a distance. In Laursen, B., Little, T. D. & Card, N. A. (Eds.), *Handbook of developmental research methods* (pp. 163–180). The Guilford Press.

Woods, A. T., Velasco, C., Levitan, C. A., Wan, X. & Spence, C. (2015). Conducting perception research over the internet: A tutorial review. *PeerJ*, *3*, e1058. https://doi.org/10.7717/peerj.1058

Woods, K. J., Siegel, M. H., Traer, J. & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, & Psychophysics*, *79*(7), 2064–2072. https://doi.org/10.3758/s13414-017-1361-2

Ye, T., Reinecke, K. & Robert, L. P. (2017, February). Personalized feedback versus money: The effect on reliability of subjective data in online experimental platforms. In *Companion of the 2017 ACM conference on compute(r supported cooperative work and social computing* (pp. 343–346). University of Washington.

Zack, E. S., Kennedy, J. & Long, J. S. (2019). Can nonprobability samples be used for social

science research? A cautionary tale. *Survey Research Methods*, *13*, 215–227. https://doi.org/10.18148/srm/2019.v13i2.7262

Zirkel, S., Garcia, J. A. & Murphy, M. C. (2015). Experience-Sampling research methods and their potential for education research. *Educational Researcher*, *44*(1), 7–16. https://doi.org/10.3102/0013189X14566879

Zong, J. & Matias, J. N. (2018, October). Automated debriefing: Interface for large-scale research ethics. In *Companion of the 2018 ACM conference on computer supported cooperative work and social computing* (pp. 21–24). Princeton University.

Author biographies

Niels Van Quaquebeke is professor of leadership and organizational behavior at the Kühne Logistics University in Germany and distinguished research professor at the University of Exeter in the United Kingdom. Having worked for AOL before (which used to be a well-known company), he seeks to combine his knowledge and skills in the online domain to further the methods of leadership research.

Mojtaba Salem is a Postdoctoral Researcher at the Chair of Research and Science Management in Technical University of Munich in Germany. He holds a PhD in Leadership and Management from Kühne Logistics University with a dissertation on "Leadership Issues in Humanitarian Operations".

Marius van Dijke is professor of behavioral ethics at Rotterdam School of Management, Erasmus University, the Netherlands. His research interests lie at the intersection of morality, social justice, power, trust and leadership. He publishes on these topics in a range of leading journals in management and psychology.

Ramon Wenzel is a Social Data Scientist at Minderoo Foundation and Adjunct Research Fellow in Work Psychology at Curtin University. He seeks to enable actionable research to make work better. His research is characterised by advanced survey designs, data pipelines, web apps, modern analytics and responsible research data governance protocols.

Appendix 1

Journals of Articles Used in Literature Search for Best Practices.

Type	*Journal*
Research method journals	Advances in Methodology & Statistics
	Behavior Research Methods
	Bulletin of Sociological Methodology
	Field Methods
	International Journal of Social Research Methodology
	Journal of Medical Internet Research
	Journal of Research Practice
	Journal of Survey Statistics and Methodology
	Methods in Psychology
	Organizational Research Methods
	Psychological Methods
	Quality and Quantity
	Research Ethics
	Research Policy
	Social Research Practice
	Sociological Methods & Research
	Survey Methods: Insights from the Field
	Survey Practice
Psychology and organizational behavior journals	American Psychologist
	Annual Review of Clinical Psychology
	Applied Psychology
	British Journal of Psychology
	Current Psychology
	Employee Surveys and Sensing: Challenges and Opportunities
	Ethics and Behavior
	Experimental Psychology
	Frontiers in Psychology
	Journal of Applied Psychology
	Journal of Business Ethics
	Journal of Business and Psychology
	Journal of Experimental Social Psychology
	Journal of Research in Personality
	International Journal of Business and Information
	Organizational Behavior and Human Decision Processes
	Personnel Psychology
	PLoS ONE
	Studia Psychologica
	Translational Issues in Psychological Science
	Zeitschrift für Psychologie

(continued)

(continued)

Type	*Journal*
Business, economics, social and natural sciences	American Journal of Evaluation American Journal of Health Education American Journal of Political Science Applied Economics Letters Attention, Perception, & Psychophysics BMC Health Services Research Brain Sciences Canadian Journal of Agricultural Economics Computers in Human Behavior Educational Research and Evaluation Ethical Issues in Literacy Research Experimental Economics Human-Centric Computing and Information Sciences International Journal of Behavioral Development International Journal of Internet Science International Journal of Market Research International Journal of Nursing Studies International Journal of Public Opinion Research International Review of Social Research Journal of Behavioral and Experimental Finance Journal of Clinical Epidemiology Journal of Consumer Research Journal of Counseling & Development Journal of Human Lactation Journal of Humanities and Social Sciences Journal of the Economic Science Association Journal of Travel & Tourism Marketing PeerJ Political Analysis Public Opinion Quarterly Safety Science Social Science Computer Review Social Work Research The American Journal of Bioethics Tsinghua Science and Technology