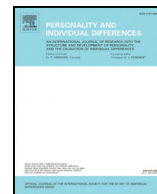




Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Exploring beyond simple demographic variables: Differences between traditional laboratory samples and crowdsourced online samples on the Big Five personality traits☆☆☆

Douglas E. Colman^{a,*}, Jared Vineyard^b, Tera D. Letzring^a

^a Department of Psychology, Idaho State University, United States

^b Idaho Center for Health Research, Idaho State University, United States

ARTICLE INFO

Article history:

Received 30 December 2016

Received in revised form 23 May 2017

Accepted 15 June 2017

Available online xxxx

Keywords:

Crowdsourcing

Mechanical Turk

Personality

Individual differences

Big Five traits

ABSTRACT

Amazon's Mechanical Turk (MTurk), a popular crowdsourcing website, is increasingly being utilized by researchers to obtain psychological data. This transition has prompted evaluation of sourcing costs, psychometric properties, and motivations of participants. However, research is limited comparing traditional and crowdsourced participants on personality measures. Therefore, in the current study laboratory participants (drawn from three universities) and MTurk workers completed the Big Five inventory and provided demographic information using web-based surveys. Controlling for age and gender, laboratory participants were significantly lower in Openness ($d = 0.26$), and higher in Extraversion ($d = 0.37$), Agreeableness ($d = 0.15$), and Neuroticism ($d = 0.05$) than MTurk participants. However, pairwise comparisons among individual sites revealed there were means above and below that for MTurk participants for Openness and Conscientiousness. Given these differences, researchers are encouraged to consider how such personality characteristics may influence the outcomes of their research when designing and conducting psychological studies that use crowdsourcing techniques to recruit participants.

© 2017 Elsevier Ltd. All rights reserved.

It has been said that what is done alone is sometimes better accomplished in a crowd. Modern behavioral researchers are now turning more frequently to cost-effective online solutions to sample and collect data from human participants (Howe, 2006). What is collectively known as crowdsourcing has become a popular avenue for data collection, with Amazon's Mechanical Turk (MTurk) becoming one of the most commonly used services. Crowdsourcing offers a significant advancement in the study of large populations (Paolacci & Chandler, 2014), but questions remain concerning the importance of individual differences between the more traditional laboratory samples in which data are collected in-person in a laboratory and MTurk samples in which data are collected online without participants coming to a laboratory. In addition, recent research has demonstrated that relatively wide variation exists in personality characteristics when participants are assessed with traditional in-person data collection across 30 colleges

and universities (Corker, Donnellan, Kim, Schwartz, & Zamboanga, 2015). The current study explored differences across settings by comparing a large sample of MTurk participants to several college student samples that used in-person data collection.

Crowdsourcing refers to using the internet to distribute tasks or work among a large group of individuals for compensation (Behrend, Sharek, Meade, & Wiebe, 2011; Chandler & Shapiro, 2016; Howe, 2006). MTurk provides a platform for researchers to post tasks for "workers" to complete for relatively little compensation. Compensation ranges from a few cents for short studies up to several dollars, although the majority of tasks posted by researchers fall below \$1.00. Recently, psychologists have increasingly used the internet, and crowdsourcing in particular, to recruit samples for studies and experiments traditionally gathered from community or university samples (Chandler & Shapiro, 2016; Skitka & Sargis, 2006).

The use of crowdsourcing websites provides a significant advantage to researchers in that large samples can be collected in a relatively short amount of time (Buhrmester, Kwang, & Gosling, 2011), compared to months (and sometimes years!) for data collection occurring in-person. In short, crowdsourcing is more efficient because it does not require physical lab space, eliminates the need for data entry, and allows for data collection at any time of the day or week. However, there has been concern over the equivalence of data quality among the various

* Portions of this research were presented at the 4th Biennial Association for Research in Personality Conference in St. Louis, MO in 2015.

☆☆ We thank Jeremy Biesanz and Simine Vazire for sharing data which comprise two of the samples reported in this article. We also thank Dolan Ward for his assistance with data collection and early contributions to this manuscript.

* Corresponding author at: Idaho State University, Department of Psychology, 921 S. 8th Ave, Stop 8112, Pocatello, 83209, United States.

E-mail address: colmdoug@isu.edu (D.E. Colman).

participant recruitment and data collection methods (Buhrmester et al., 2011; Gosling, Vazire, Srivastava, & John, 2004; Ward, 1993).

It has been argued that platform (paper-and-pencil, lab computer, and crowdsourced) differences could significantly alter the results of a study. Yet, evidence from several studies has demonstrated that differences between traditional in-person and crowdsourced participants are minimal. At the assessment level, there is already strong evidence of measurement equivalence/invariance for personality measures taken by MTurk workers and traditional in-person participants. Specifically, the 100 item IPIP version (Goldberg et al., 2006) of the NEO-PI-R was found to have equivalent psychometric properties across samples (Behrend et al., 2011). Likewise, the Big Five inventory (John, Naumann, & Soto, 2008) showed measurement invariance when MTurk participants were restricted to being from countries in which English is the primary language (Feitosa, Joseph, & Newman, 2015). In addition to efficiency and data equivalency, MTurk participants tend to be more diverse on important demographic variables such as age, education, and ethnicity than traditional in-person college/university participants (Behrend et al., 2011; Paolacci & Chandler, 2014). This greater demographic variability directly addresses one of the common limitations of studies conducted in-person at a single location.

Although there are benefits to crowdsourcing psychological data, such samples may not always produce conclusions that can be generalized to a non-MTurk population. One characteristic of crowdsourced participants that has the potential to inhibit generalizability is personality. Previous research has found that crowdsourced samples, in comparison to in-person samples, were lower on Extraversion, Neuroticism, Openness to experience, and Conscientiousness¹ when the Big Five traits were assessed with a 10-item inventory, as well as lower on trait level self-esteem when assessed with a single item (Goodman, Cryder, & Cheema, 2013; Kosara & Ziemkiewicz, 2010). A second study found that a crowdsourced sample again revealed lower levels of Extraversion and Neuroticism, but found higher levels of Openness and lower levels of Agreeableness (Kosara & Ziemkiewicz, 2010). Additionally, personality traits have also been found to be quite variable across traditional samples collected from different regions of the US (Corker et al., 2015), and therefore it is not yet clear whether the magnitude of the differences found between MTurk and traditional samples is within the range that would be expected based on regional differences.

Furthermore, it is well-known that differences in personality are related to behavioral differences that can result in variation in important life outcomes (Ozer & Benet-Martinez, 2006; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007), and therefore meaningful differences in personality between crowdsourced and traditional samples could be muddying the conclusions that can be drawn about basic psychological processes. Therefore, if consistent differences in personality exist between crowdsourced and traditional participants, it would behoove researchers to uncover and consider such differences when using crowdsourced samples.

Measuring and reporting variance in personality patterns across settings is a key aspect of understanding the impact of these differences for research efforts. Samples using crowdsourcing methods, university students, or community members differ on important variables. Therefore, replication of previous work and comparing multiple samples should be employed to strengthen the understanding of these differences. Using multiple samples allows for improved comparison within traditional settings and improves reliability of estimates of personality and related individual differences. Previous work has shown variability between crowdsourced and traditional samples (Goodman et al., 2013; Kosara & Ziemkiewicz, 2010), however a multiple-sample strategy is needed from both crowdsourced and traditional milieu to better examine the pattern of differences.

¹ Differences for Extraversion, Neuroticism, and self-esteem were found in two studies, while differences for Conscientiousness and Openness were only found in one study.

Table 1
Demographic characteristics by sample.

	Sample					
	A	B	C	D1	D2	D3
Gender						
Male	389	145	137	31	39	36
Female	880	355	279	135	88	79
Age						
M	37.15	21.77	19.43	21.00	19.50	20.41
SD	12.55	5.32	2.32	4.09	2.02	3.43
Range	18–78	18–53	18–39	17–50	17–35	17–44

1. Summary

Recent research has noted significant variability in participant personality across the United States and concluded that studies comparing a limited number of participant sources run the risk of over-generalizing research findings (Corker et al., 2015). On the other hand, data from crowdsourced participants are thought to provide high quality data (Buhrmester et al., 2011) and incorporate a wider range of individuals with regard to age and ethnic background (Behrend et al., 2011). Less clear, however, is the personality profile of crowdsourced participants compared to traditional in-person participants. Thus, the current analyses expand upon the framework of Corker et al. (2015) by examining the individual differences among participants from several in-person samples and a larger sample of crowdsourced participants (Table 1).

2. Method

2.1. Participants

2.1.1. Sample A

These data ($N = 1279$) were collected between Fall 2014 and the end of 2015 via an online survey platform. Data were collected as part of several projects for which all participants were recruited from MTurk, but restricted to individuals with an approval rating of 95% or greater who reside in the United States.² Participants were financially compensated between \$0.50 and \$2.00, depending upon the project.

2.1.2. Sample B

These data ($N = 500$) were collected between Fall 2014 and Spring 2016 in person at Idaho State University located in the West region of the United States. Data were collected as part of two separate projects for which all participants were recruited from a department subject pool. Announcements about each project were also made in some courses where research was a required element. As such, all participants were undergraduate students and were remunerated with research credits.

2.1.3. Sample C

These data ($N = 418$) were collected between Fall 2012 and the end of 2013 in person at Washington University in St. Louis, a private university in the Midwest region of the United States (Vazire et al., 2016). These participants were taking part in a longitudinal study for which they receive financial compensation. Recruitment occurred through various methods, including a department subject pool, flyers, and announcements in classes. Most of the participants were undergraduate students but a few (about 7%) were graduate students.

2.1.4. Sample D

These data ($N = 408$) were collected in or after 2007 and prior to 2014 in person at University of British Columbia located in British Columbia, Canada. Data were collected as part of several projects in

² All participants met inclusion criteria, which consisted of successfully answering $\geq 80\%$ of embedded attention checks and completing $\geq 80\%$ of the procedure in the given study.

Download English Version:

<https://daneshyari.com/en/article/7248480>

Download Persian Version:

<https://daneshyari.com/article/7248480>

[Daneshyari.com](https://daneshyari.com)