SAGE Research Methods Cases Psychology Submission for Consideration

Case Title

Using Amazon Mechanical Turk to Recruit Participants with Disabilities

Author Name(s)

Emily M. Lund[1,2], Michael R. Nadorff[3,4], Kate Galbraith[5], Katie B. Thomas[6]

Author Affiliation & Country of Affiliation

1. Center for Psychiatric Rehabilitation, Boston University, USA
2. Department of Special Education and Rehabilitation, Utah State University, USA
3. Department of Psychology Mississippi State University, USA
4. Baylor College of Medicine, USA
5. Utah State Office of Rehabilitation, USA
6. James A. Haley Veterans Hospital, USA

Lead Author Email Address

Email: emlund@bu.edu / emily.m.lund@gmail.com

Discipline: D3 [please do not alter]

Sub-discipline

General and Applied Psychology [SD-Psych-1]

Academic Level

Intermediate Undergraduate

Contributor Biographies

Emily M. Lund, Ph.D., is a postdoctoral fellow at the Center for Psychiatric Rehabilitation at Boston University. She holds a PhD in Disability Disciplines from Utah State University, with specializations in both rehabilitation counseling and special education. She has published extensively on topics relating to victimization and trauma in people with disabilities, suicide and non-suicidal self-injury in the context of disability and trauma, and the experiences of psychology graduate students with disabilities.

Michael R. Nadorff, Ph.D., is a clinical psychologist and Assistant Professor of Psychology at Mississippi State University. His research is focused on the intersection between behavioral sleep medicine and suicide. He is the chair of the 2017 American Association of Suicidology conference, and his work has been supported by SAMHSA and NIMH.

Kate Galbraith, MRC, CRC, is a vocational rehabilitation counselor for the Utah State Office of Rehabilitation. She holds a masters degree in rehabilitation counseling from Utah State University. She has published research on suicide and disability and the transition from high school for students with disabilities.

Katie B. Thomas, Ph.D., is a clinical psychologist in the Trauma Recovery Program at the James A. Haley Veterans Hospital in Tampa, FL. She obtained her doctorate from the University of North Dakota and recently completed her post-doctoral fellowship in Advanced Women's Health through the Veteran's Affairs San Diego Healthcare System and University of California, San Diego. She has published research on sexual assault and other interpersonal trauma, emotion dysregulation, and suicidal and self-injurious behavior.

Published Articles

Lund, E. M. (2016). Suicide and disability: Three different analyses of a nation-wide sample of American adults (unpublished doctoral dissertation). Utah State University, Logan, UT.

Lund, E. M**.**, Nadorff, M. R., & Seader, K. (2016). The relationship between suicidality and disability when accounting for depressive symptomology. *Rehabilitation Counseling Bulletin*, *59*, 185-188. doi: https://doi.org/10.1177/0034355215586388

Lund, E. M., Nadorff, M. R., Winer, E. S., & Seader, K. (2016). Is suicide an option?: The impact of disability on suicide acceptability in the context of depression, suicidality, and demographic factors. *Journal of Affective Disorders*, *189*, 25-35. doi: https://doi.org/10.1016/j.jad.2015.08.028

Abstract

Amazon Mechanical Turk (MTurk) is an increasingly popular, cost effective, and quick way to recruit participants. In general, researchers have found samples recruited from MTurk to produce reliable, valid data and to be fairly diverse and representative. However, little is known about the prevalence of disability in MTurk samples. This case study reviews the mechanics of MTurk, previous research on MTurk samples, and discusses the results of a study that included an item regarding disability identity. In this study, we found that our sample reported rates of disability roughly similar to that seen in the general population and that physical, chronic health, and psychiatric disabilities were the most commonly reported conditions.

Learning Outcomes

By the end of this case study, students should be able to:

1. Analyze the strengths and limitations of MTurk as a participant recruitment system.
2. Describe the general composition of reported MTurk samples.
3. Compare the potential effects of different ways of asking about disability status.

Case Study

Project Overview

This project was a study of disability, suicidality, and attitudes towards suicide in a sample of American adults. Our primary research question was about the ways in which the disability status, or lack thereof, of a hypothetical suicidal individual influences suicide acceptability and permissiveness in respondents. To this end, we had participants read five pairs of brief vignettes about individuals who were experiencing suicidal ideation stemming from a life stressor (e.g., academic problems, isolation and lack of sense of purpose, break up of a romantic relationship, chronic unemployment). In each vignette pair, participants read one vignette about a suicidal individual who was stated to have a disability (e.g., congenital blindness, spinal cord injury, chronic health condition, bipolar disorder, traumatic brain injury) and another vignette about a suicidal individual of a similar background and experiencing a similar life stressor who was not stated to have any disabilities. Participants completed five Likert-type rating questions for each vignette from 1 ("strongly disagree") to 5 ("strongly agree"). The items included statements like "I believe [hypothetical individual's name] should have the right to kill himself/herself" and "If I were in [name's] situation, I would feel in the same way." The scores were compared both overall and for each vignette pair. The results of these analyses are published in Lund, Nadorff, Winer, and Seader (2016). In addition, participants also completed measures of depressive symptoms, suicidality, attitude towards disability, and a variety of demographic items. These other variables were used in both the primary analyses and in secondary analyses of suicidality in people with and without disabilities (see, for example, Lund, Nadorff, & Seader, 2016).

Recruitment Source: Amazon Mechanical Turk

We used Amazon Mechanical Turk (often abbreviated as MTurk) to recruit participants. MTurk is a participant recruitment system managed by Amazon. Participants, called "workers," complete tasks, called "Human Intelligence Tasks" (HITs) in exchange for small amounts of monetary compensation. HITs are typically brief, single-session tasks, and payment amounts for HITs tend to be relatively small amounts, such as $.25 USD, $.50 USD, or $1.00 USD; this is sometimes referred to as "micro compensation." Individuals who need workers ("requesters") post HITs that include information on the amount of compensation available and the approximate length and nature of the task. Before a HIT is posted on MTurk, requesters set the compensation level and number of participants desired and pay the total compensation amount plus a posting fee. MTurk automatically distributes compensation once a worker's HIT is approved. Requesters can approve all HIT submissions manually or set MTurk to automatically approve all submissions after a certain amount of time (e.g., one minute, five minutes, one hour, one day). Once the target number of participants is reached, MTurk automatically removes the HIT notice from its listings. As Mason & Suri (2012) note, requesters can limit their HIT audiences to individuals from certain countries, as based on internet provider (IP) address, or allow workers from all countries to view and complete their HIT. Buhrmester, Kwang, and Gosling (2011) found that higher compensation rates tended to be associated with more quickly reaching the target number of completed HITs (i.e., participants) but did not appear to affect the quality of responses. Even higher compensation rates were relatively small, given the practice of micro compensation on MTurk

Quality and Nature of Data from MTurk Samples

As with any recruitment source, it is important that researchers have confidence in the quality of the data obtained from participants recruited from that source. Buhrmester and colleagues (2011) found that data collected from MTurk generally has good to excellent psychometric properties, including internal consistency and test-retest reliability, which we also found in the present study. In their review of the literature on MTurk data validity, Mason and Suri (2012) reported that data from MTurk has been consistently shown to be valid and reliable, and others researchers have found evidence in support of the validity of the data obtained from MTurk samples. For example, Thomas, Lund, and Bradley (2015) conducted a study on MTurk in which participants completed a measure of nonsuicidal self-injury that contained open-ended items; they found that responses to the open-ended items were overwhelmingly logical, appropriate, and consistent with the questions asked. Bogart, Rottenstein, Lund, and Bouchard (in press) included two "attention check" items (e.g., "For this item, select 'strongly agree'") near the end of a long MTurk (300-900 variable) survey. Only 126 out of 1,105 participants (11.4%) failed an attention check, again suggesting that

MTurk produces valid data.

Demographic Characteristics of MTurk Samples

In terms of demographics, both Buhrmeister and colleagues (2011) and Thomas and colleagues (2015) found that MTurk recruitment results in samples with a roughly equal male/female gender ratio and a mean age in the mid-30s. Buhrmester and collegues (2011) also noted that MTurk samples tend to be more racially and demographically diverse than most college student samples (Buhrmester et al., 2011). Both Buhrmester and colleagues and Thomas and colleagues (2015) reported recruiting samples from MTurk that are about 76%-80% White. This is roughly in-line with the percentage of Americans who described themselves as White of any ethnicity (77.1%), as reported by the United States Census Bureau (2015).

The use of MTurk for recruiting marginalized populations of particular interest is a somewhat open topic. Shapiro, Chandler, and Mueller (2013) found that their MTurk samples reported roughly the same or higher rates of psychopathology as general population samples, with the rates of social anxiety being reported being markedly higher than that seen in non-online samples. They concluded that MTurk may be a convenient and inexpensive platform from which to recruit participants with both clinical and subclinical levels of psychological distress, particularly in the area of social anxiety. Although they noted that misrepresentation is a potential issue, they also noted the overall high quality and good reliability of the data.

In addition to their findings on psychopathology, Shapiro and colleagues (2013) found that about 9-10% of their sample identified as non-heterosexual. In contrast, Lund, Thomas, Sias, and Bradley (2016) found that about 29% of their sample reported something other than exclusively heterosexual and heteroromantic attraction. However, they asked about attraction rather than identity and separated out romantic and sexual attraction. This may have resulted in some participants who may have identified simply as "heterosexual" on a forced-choice identity—such as individuals who experience bisexual sexual attraction but only heterosexual romantic attraction or those who experience heterosexual sexual attraction but no romantic attraction—to be classified as sexual or romantic minorities (Lund, Thomas, et al., 2016). This also highlights the possible importance of how demographic questions are worded, broken down, or asked in influencing responses. As another example, individuals who identify as Hispanic or Latino may be under-represented in the demographics of many MTurk samples. For example, Thomas and colleagues (2015) reported that 4.1% of their sample identified as Hispanic or Latino, and Shapiro and colleagues (2013) also reported that 4.1% of their sample was Hispanic. This is markedly lower than the 17.6% of Americans who identify as Hispanic or Latino on the U. S. Census (2015). However, it is important to note that the U.S. Census data asks about Hispanic/Latino ethnicity separate from race and that about two thirds of Hispanic/Latino individuals also identify as White alone (U.S. Census Bureau, 2014), so it is unclear how those individuals would identify if presented with a forced-choice option. Future research should examine MTurk racial and ethnic demographics with an open-choice race/ethnicity item or separate items for race and ethnicity, as this may also better capture the 2.9% of Americans who identify as multiracial (U.S. Census Bureau, 2015).

MTurk and Disability

Like individuals who are sexual minorities or those who identify as non-White or Hispanic, people with disabilities may be a population of specific interest to researchers. People with disabilities often have reduced access to many community settings and thus may be missed by community-based recruitment methods. Similarly, data collection techniques such as random digit phone dialing may exclude individuals who are Deaf or hard of hearing or have significant speech disabilities, and individuals who are blind or visually impaired may be excluded by study materials that are provided only in hardcopy. Because computers allow for a wide range of accessibility features, such as text enlargement or the use of text-to-speech software, they may provide a higher level of access to people with diverse disabilities (Oschwald et al., 2014). As reported by Lund, Andrews, and Holt (2014), many online survey systems, such as Qualtrics, can be made to be screen-reader accessible and can be checked for web accessibility compliance.

There is relatively little available information on the prevalence of disability in MTurk samples. Shapiro and colleagues (2013) reported that 16.3% of the participants in their sample reported that they had been "ever diagnosed with a chronic illness or physical disability" and 21.0% reported that they had been "ever diagnosed with a psychiatric or psychological condition." Thomas and colleagues (2015) reported that 28.0% of their sample had ever received a mental health diagnosis from a doctor or therapist. In general, they reported that these diagnoses were primarily anxiety and depression, a similar finding to that of Shapiro and colleagues (2013). However, as Bogart and colleagues (in press) note, diagnosis of a physical or mental health condition is different from disability identity; not all individuals who receive such a diagnosis may identify as a person with a disability. In addition, some conditions may not reach the level of impairment of a disability or may resolve over time or with treatment. Thus, asking about physical or mental health diagnosis may be different than asking about disability.

Recruitment for the Present Study

The study was advertised under the title "Examining Our Perceptions of Individuals With and Without Disabilities" with the stated purpose of "understand[ing] the relation between attitudes toward suicidal thoughts and behavior and disability." Demographic items, including items about disability, were placed towards the end of the survey. In order to be eligible for participation in the present study, participants had to be United States residents and be age 18 or older. Participants were paid $.25 for participation. In order to protect the anonymity and confidentiality of participant responses, all data were collected on a secure, university-based Qualtrics server and could not be linked to participant MTurk identification numbers or other identifying information. In order to receive compensation, participants entered a code that they received at the end of the survey into MTurk and automatically received compensation at that point. The survey took approximately 30 minutes to complete. All materials and procedures were approved by a university institutional review board (IRB) prior to data collection.

Participant Demographics

Five hundred participants completed the survey. A majority were female (60.4%; *n*=302) and Caucasian (74.4%; *n*=372). As with other MTurk samples, such as those used by Shapiro and colleagues (2013) and Thomas and colleagues (2015), the mean age of our sample was in the mid-30s, specifically 35.92 years (*SD*=13.85, range=18-75). Similar to the findings of Buhrmester and colleagues (2011) and Shapiro and colleagues (2013), participants were fairly well-educated, with two-fifths of respondents (*n*=203; 40.6%) having earned a bachelor's degree or higher and an additional 43.0% (*n*=215) having completed an associate's degree or some college. Most reported working full-time (*n*=178; 35.7%) or part-time (*n*=75; 15.0%). Additionally, just under a fifth were full-time students (*n*=94; 18.8%). Participants represented 49 states and Puerto Rico. Additional demographic information can be found in Lund, Nadorff, Winer, & Seader (2016).

Participant Disability Status

Of the 500 participants, 15 (3%) selected the response "prefer not to disclose" to the item, "Do you identify as having a disability or disabilities?" Although these individuals were included in the main analyses on suicide acceptability (Lund, Nadorff, Winer, & Seader, 2016), they were excluded from analyses focusing on disability status and suicidality (Lund, Nadorff, & Seader, 2016). Of the remaining 485 participants, 94 initially answered “yes” to the item regarding disability identity. These participants were then shown an item that read, “Please describe the disability or disabilities you have. If you prefer not to disclose, please say so,” with a space for text entry. Two of the 94 respondents were recoded as “no” responses to the disability item due to responses to the open-ended item that explicitly denied disability status (“I don’t” and “I do not have any disabilities”). Thus, 92 of the 485 participants who responded to the item (19.0%) identified as having one or more disabilities. This is similar to the roughly 20% of American who have a disability according to the United States government (Brault, 2012).

Responses to the open-ended item were recoded into disability categories by the principal investigator. Seven participants stated that they preferred not to disclose, and an additional five participants gave responses that were unclear or uninterpretable. Thus, a total of 12 responses were coded as “unknown disability category.” The remaining 80 responses were coded as follows: psychiatric (*n*=24), chronic health condition (*n*=19), physical (*n*=20), learning disability (*n*=4), speech disability (*n*=3), hearing impairment (*n*=3), autism spectrum disorder (*n*=2), chronic health condition and physical (*n*=2), chronic health condition and visual impairment (*n*=1), physical and psychiatric disability (*n*=1), and visual impairment (*n*=1). It is difficult to compare these specific category prevalences to national data, as U.S. Census estimates report area of functional impairment—such as “difficulty walking” or “difficulty concentrating”—rather than type of disability (Brault, 2012). However, disability identification categories are commonly used other forms of research and reporting. For example, the Association of Psychology Postdoctoral and Internship Centers (APPIC) asks internship applicants to classify their disability by similar categories and reported that chronic health conditions were the most commonly endorsed category, followed by learning disability and "mental illness" (Andrews & Lund, 2015). The representation of disability may change over the lifetime; as Brault (2012) notes, people tend to acquire disabilities, especially physical and chronic health conditions, with age. In general, our results are somewhat similar to those of Shapiro and colleagues (2013) who found that mental health conditions and chronic health conditions, such as chronic pain, were most commonly reported in their MTurk sample.

Limitations and Consideration

As a whole, our findings suggest that MTurk may be a good platform with which to recruit American adults with disabilities. We found that just under 20% of participants who responded to the item about disability identified as having a disability, a proportion close to the 20% of Americans who are estimated to have a disability. Additionally, 97% of the sample was willing to answer the question, suggesting a willingness to disclose disability status in an anonymous, online survey. We found that physical, psychiatric, and chronic health disabilities were the most commonly reported disability types. MTurk may be more useful for recruiting individuals with these types of disabilities as opposed to learning disabilities, sensory impairments, or autism spectrum disorders. Additionally, the nature of MTurk may exclude participants with more severe disabilities who cannot use a computer or understand the platform or informed consent processes. Thus, individuals with significant cognitive disabilities or other disabilities that significantly interfere with computer use or comprehension may be excluded from participation in MTurk surveys.

One important limitation and consideration was our use of an open-ended disability type identification item versus checkbox-style reporting of disability type (e.g., "What type or types of disability do you have? Check all that apply"). Although the use of an open-ended item provides more information on an individual's specific condition, it also creates some subjectivity in coding of disability type, increases the response cost of responding to the item, and creates the possibility of unclear or uninterpretable responses, some of which we did receive in our survey. Additionally, some participant responses were cut-off due to length, leaving the possibility that some participants may have disclosed other conditions or details that could not be captured in our coding due to the length limitation for responses. This, combined with the additional response burden placed by typing, may explain the relatively low rates of multiple types of disability relative to that reported by Lund and colleagues (2014), who used checkbox-style reporting of disability type. Researchers conducting future MTurk studies may wish to compare the results of checkbox-style reporting versus open-ended items for disability type.

Exercises and Discussion Questions

1. What do you think about MTurk's use micro-compensation? How might it affect the types of participants recruited?
2. Why might asking individuals if they have been diagnosed with a chronic health condition yield different responses than asking if they have a disability?
3. What are the advantages and disadvantages of using a checkbox style system of disability identification versus an open-ended item?
4. Why might an individual identify as having a disability but prefer not to disclose the nature of their disability?

Web Resources

Amazon Mechanical Turk: https://www.mturk.com/mturk/welcome

References

Andrews, E. E., & Lund, E. M. (2015). Disability in professional psychology training: Where are we? *Training and Education in Professional Psychology, 9*, 210-216. doi: https://doi.org/10.1037/tep0000085

Bogart, K. R., Rottenstein, A., Lund, E. M., & Bouchard, L. M. (in press). Disability self-identification in adults with chronic health conditions: Impairment and contextual influences. *Rehabilitation Psychology*.

Brault, M. W. (2012). *Americans with disabilities: 2010. Current Population Reports*. Retrieved from http://www.census.gov/prod/2012pubs/p70-131.pdf

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. doi: https://doi.org/10.1177/1745691610393980

Lund, E. M., Andrews, E. E., & Holt, J. M. (2014). How we treat our own: Characteristics and experiences of psychology trainees with disabilities. *Rehabilitation Psychology, 54*, 367-375. doi: https://doi.org/10.1037/a0037502

Lund, E. M., Nadorff, M. R., & Seader, K. (2016). The relationship between suicidality and disability when accounting for depressive symptomology. *Rehabilitation Counseling Bulletin*, *59*, 185-188. doi: https://doi.org/10.1177/0034355215586388

Lund, E. M., Nadorff, M. R., Winer, E. S., & Seader, K. (2016). Is suicide an option?: The impact of disability on suicide acceptability in the context of depression, suicidality, and demographic factors. *Journal of Affective Disorders*, *189*, 25-35. doi:

Lund, E. M., Thomas, K. B., Sias, C. M. & Bradley, A. R. (2016). Examining concordant and discordant sexual and romantic attraction in American adults: Implications for counselors. *Journal of LGBT Issues in Counseling, 10*, 211-226. doi: https://doi.org/10.1080/15538605.2016.1233840

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

Oschwald, M., Leotti, S., Raymaker, D., Katz, M., Goe, R., Harviston, M., Wallington, A., Howard, L., Beers, L., Nicolaidis, C., Robinson-Whelen, S., Hughes, R. B., Lund, E., Powers, L. E., & the Partnering with People with Disabilities to Address Violence Consortium (2014). Development of an audio-computer assisted self-interview to investigate violence and health in the lives of people with developmental disabilities. *Disability and Health Journal, 7*, 292-301. doi: https://doi.org/10.1016/j.dhjo.2014.01.002

Shapiro, D. N., Chandler, J., & Mueller, P. A. (2013). Using Mechanical Turk to study clinical populations. *Clinical Psychological Science*, *1*(2), 213-230. doi: https://doi.org/10.1177/2167702612469015

Thomas, K. B., Lund, E. M., & Bradley, A. R. (2015). Composite trauma and mental health diagnosis as predictors of lifetime non-suicidal self-injury history in an adult online sample. *Journal of Aggression, Maltreatment, and Trauma, 26*, 623-635. doi: https://doi.org/10.1080/10926771.2015.1049766

U.S. Census Bureau (2014). *Hispanic or Latino Origin by race: 2014 American Community Survey 1-Year estimates*. Retrieved from http://factfinder.census.gov/ faces/tableservices/jsf/pages/productview.xhtml?src=bkmk

U.S. Census Bureau (2015). *QuickFacts: United States*. Retrieved from https://www.census.gov/quickfacts/table/PST045215/00