

Penny for Your Thoughts: Using Amazon MTurk for Research

Megan Holly, Alexandra Washburn, Michael Curtis, & Kimberly A. Barchard

University of Nevada, Las Vegas

UNLV

ABSTRACT

Recruiting participants is a perennial challenge for researchers. Crowdsourcing can provide a solution to this challenge. Crowdsourcing is the act of outsourcing small jobs to large groups of people (Miziolek, 2011). Recently, crowdsourcing websites – such as Amazon Mechanical Turk (MTurk) – have appeared across the Internet, making it easy for ordinary people to use crowdsourcing. Researchers have been using these sites to collect data from all over the world quickly and cheaply – many MTurk studies pay only a few pennies. Hundreds of research studies have recruited participants through MTurk, and 16 of the top 30 universities in the United States using MTurk to collect data in the areas of social psychology, cognitive psychology, linguistics, and other social sciences (Goodman, Cryder, & Cheema, 2013).

MTurk provides a number of useful features that can be used by researchers with a moderate amount of computer knowledge. One such feature is the ability to hide the study from people who have already completed it, thus preventing multiple submissions, which might otherwise distort the findings. Another useful feature is the ability to invite panel members. A panel member is a person who has completed a pre-screening study and is then invited to participate in the main study or in a series of studies. MTurk allows researchers to invite panel members by sending study invitations to specific people.

Some researchers are concerned about online crowdsourcing because they worry that conducting studies online may be a less valid way of collecting data, especially when compared to traditional studies that take place in person (Kesebir, 2010). However, studies in a variety of topic areas have concluded that data collected through MTurk are just as valid as – and perhaps even more valid than – data obtained through traditional means (Buhmester, Kwang, & Gosling, 2011; Casler, Bickel, & Hackett, 2013).

This poster will explain the benefits of using MTurk to recruit research participants. Next, it will teach people how to hide the study from participants who have already completed it and how to invite panel members. Finally, this poster will discuss which types of studies would most benefit from using MTurk to recruit participants.

CROWDSOURCING

Gathering participants is a recurrent challenge for psychological researchers. For a study to have adequate power, many participants are usually necessary. Psychologists often recruit participants through a university subject pool because this is inexpensive and easy, but university students may not be representative of the populations of interest. Online crowdsourcing platforms – such as Amazon Mechanical Turk (Amazon, n.d.) – allow researchers to quickly recruit larger and more diverse samples of participants. Moreover, this data collection can often be accomplished with a much smaller research team than in-person studies require (Miziolek, 2011). Thus, MTurk can cut down the overall cost and time that it takes to complete a study (Buhmester, Kwang & Gosling, 2011).

Amazon's Mechanical Turk (MTurk; Amazon, n.d.) is an online crowdsourcing platform. Initially marketed predominantly to businesses, MTurk is now often used to conduct psychological research (Paolacci & Ipeirotis, 2010). Hundreds of research studies have been conducted through MTurk, and 16 of the top 30 universities in the United States have used MTurk to collect data in the areas of social psychology, cognitive psychology, linguistics, and other social sciences (Goodman et al., 2013). To post a study on MTurk, the researcher first signs up for a requester account. Then the researcher creates a Human Intelligence Task (HIT), which provides a link to the online study materials. Finally, MTurk workers see the HIT and have the option of becoming participants in the study.

MTurk advertisements are an effective way to recruit participants for several reasons. First, data collected on MTurk have been shown to be at least as reliable as data collected through traditional means (Buhmester, Kwang & Gosling, 2011). Second, MTurk samples are slightly more demographically diverse than are standard Internet samples and are significantly more diverse than typical American college samples (Buhmester, Kwang & Gosling, 2011). Third, MTurk is helpful for researchers that wish to recruit participants from diverse populations that may not be readily available at their institution or in their immediate communities. For example, a researcher might want to compare cultural norms in Japan and the United States. By posting their study on MTurk, the researcher could limit the participants to those who live in the United States and Japan and could collect data without having to physically meet or speak to participants directly. Finally, researchers can advertise their studies to specific types of participants, rather than having to manually remove unwanted participants after the data have been collected. This can be done using MTurk qualifications. Qualifications are traits assigned to workers. MTurk has a list of default qualifications, such as the worker's location and total approved HITs. Requesters can create their own qualifications for workers.

Concerns About Crowdsourcing

As online crowdsourcing gains popularity with researchers, it faces increased scrutiny. Many dismiss the benefits of online crowdsourcing by claiming that the online medium is a less valid way of collecting data than collecting data in person. However, research conducted about studies on MTurk have concluded that data collected through MTurk are just as valid as data obtained through traditional studies (Buhmester et al., 2011; Casler et al., 2013).

Critics of online crowdsourcing also claim that compensation affects the quality of data. Indeed, research has shown that rates of compensation affect the speed of data collection. Higher monetary compensation increases the speed at which participants are recruited (Kesebir, 2010). However, compensation does not affect the actual quality of data in any significant way (Buhmester et al., 2011).

USING MTURK FOR RESEARCH

MTurk allows researchers to create studies using two methods. First, MTurk provides a number of templates for different kinds of HITs, such as surveys, transcriptions, categorization, and picture tagging. Once a template is selected, the researcher uses MTurk's formatting tools to customize the materials for their study. Alternatively, researchers can choose to provide a link to an external survey website that was created using software such as Qualtrics or SurveyMonkey. In this case, the researcher must create the study using the external website. In that external survey, the researcher should include a unique code that study. Within MTurk, the researcher uses the template for an external survey link, and enters the link to the external survey. When participants complete the survey, they see the unique code and then enter that into a box in MTurk to prove that they completed the study.



Researchers can send emails through MTurk to workers who have completed HITs for them in the past. Researchers often use this to send participants information about new studies. This allows researchers to stay in touch with workers who have a specific background or skills that are useful for future studies.

Designing and running studies using the above methods is effective and easy. Researchers can collect large samples in a matter of days, often for significantly less

than it would cost to pay in-person participants. The functionality of MTurk can be greatly increased, however, by learning how to invite panel members and to prevent duplicate participation.

Inviting Panel Members

A panel is a group of people who completes multiple studies. Researchers might want a panel for several reasons. First, if the studies are long and pay a lot of money, the researcher might want to limit participation to workers that the researcher already knows are reliable and trustworthy workers. Second, the researcher might want to limit the study to certain types of participants. For example, if the researcher wants participants who are left-handed, they can conduct a screening study to ask people if they are left-handed (along with other background information) and then invite the left-handed people to subsequent studies. Third, if the researcher wants to collect a lot of data from the same participants, they may have more surveys or tasks than can be completed during a single setting. In that case, the tasks could be divided between several HITs and the researcher could ask the same people to complete all of those HITs.

To set up the panel, the researcher must first publish an initial HIT, which asks participants the screening questions. Once enough participant data has been gathered, the researcher removes the HIT and analyzes the data. If the study used MTurk exclusively, the researcher can download the .csv file that contains a list of all Worker IDs and the set of each Worker's responses. If third-party software (such as Qualtrics) is being used, the MTurk .csv file will only contain Worker IDs. Instead, the researcher must obtain the data from the site being used. Therefore, the researcher must design the screening survey so that it includes Worker IDs.

The researcher uses this file to decide which workers will be placed in the panel, depending on the researcher's needs. For example, a researcher may choose panel members that are left-handed, that completed all of the previous studies, or that provided well-written responses to previous survey questions.

There is a page on MTurk that allows the requester to email workers one at a time, but this can be tedious and time-consuming when contacting many workers. Instead, the researcher can write a program to email workers for them. Figure 2 contains an example of this program.

The program uses Amazon's Requester API (Application Programming Interface), which is a set of operations, data structures, and parameters that MTurk uses to complete tasks for the user. These operations can perform actions such as creating HITs, granting bonuses, or checking your account balance. In this case, the program will use an operation called NotifyWorkers, which is designed to email a provided list of workers.

When writing the program, the researcher sends a digital signature to MTurk, so that it knows who is sending the request. This digital signature contains the operation name, service name (AWSMechanicalTurkRequester), the service version, a timestamp, the researcher's access key ID, and the researcher's secret access key. These last two can be found by logging onto Amazon Web Services and checking Your Security Credentials). This allows Amazon to verify that the request is coming from a legitimate source.

The researcher writes the subject and body of the email that the program will send. In the case of a panel invitation, the email invites each worker from the panel to participate in a future study.

Lastly, the program must include a list of Worker IDs. To recruit participants from the panel for a new study, the researcher creates a text file that contains the Worker IDs of all those wanted for the study.

Once the program has been created, the researcher compiles and runs the program. When the re-

Figure 2
Portion of Code from AssignQualification Program

```
##### MAKE 1ST CHANGE HERE #####
#Insert your Qualification ID in my $Qual_ID, replacing "QUAL ID GOES HERE"
#Insert the integer value you want to assign the workers that get this qualification in my
$Pct_val
#If you want the workers to receive a notification email, change my $Notify from 0 to 1
my $Qual_ID = "QUAL ID GOES HERE"
my $Pct_val = 100
my $Notify = 0
##### MAKE 2ND CHANGE HERE #####
#Sign up for Amazon web services here:
https://aws-portal.amazon.com/gp/aws/developer/registration/index.html
#Look up your personal requester "Access Key ID" with this link and insert below (replacing
"ACCESS KEY ID GOES HERE")
https://aws-portal.amazon.com/gp/aws/developer/account/index.html?action=access-key
#From the same page, look up your secret access key and insert below (replacing "SECRET ACCESS KEY
GOES HERE")
my $AWS_SECRET_ACCESS_KEY = "SECRET ACCESS KEY GOES HERE"
##### LEAVE BELOW UNCHANGED #####
my $SERVICE_NAME = "AWSMechanicalTurkRequester";
my $SERVICE_VERSION = "2013-11-25";
```

searcher runs the program, MTurk sends emails containing the provided information to each worker in the list. Using this process, researchers can invite dozens (or hundreds) of panel members very efficiently.

Preventing Repeat Participation

Typically, when a worker completes a HIT, that specific worker is not allowed to complete it again. However, if a researcher posts the same study multiple times, a worker could complete the study once each time it appears as a new HIT. Also, if a researcher does several studies that are similar to each other, it may be inappropriate for a worker to participate in more than one of these studies. For these reasons, it is sometimes necessary to exclude people who have done a previous study.

Researchers have two options. They can either identify people who have done their study more than once and remove them from the database or they can somehow completely prevent people from doing their study more than once. This second solution is ideal and can be accomplished using MTurk's qualification system to track workers that have completed studies. Researchers can create HITs that require certain qualifications in order to prevent the wrong people from completing them.

To do this, the researcher publishes a HIT and collects data from participants. Then the researcher downloads the .csv file from MTurk to get the list of workers that have completed that HIT. The researcher copies those Worker IDs into a text file. Next, the researcher logs onto MTurk and creates a qualification type that will be used to distinguish workers who have completed the study from those that have not. MTurk allows researchers to assign this qualification to each worker individually, but it is more efficient to use a program when handling a large number of workers. Figure 2 provides a program that will do this.

This program completes the operation named AssignQualification, which is another feature in Amazon's Requester API.

The contents of this program are similar to the contents of the program that invites panel members, except this program includes the qualification ID, the integer value each worker gets with the qualification, and a true-or-false value indicating whether the workers will receive a notification email.

When the program compiles and runs, the program assigns this qualification type to all the listed workers.

Whenever the researcher publishes a future HIT for the same study or a related study, the researcher should specify that the worker must not have this qualification.

DISCUSSION

MTurk provides a large and easy-to-access worker base from people all around the globe. With MTurk, it has become significantly easier for psychological researchers to find and engage with diverse potential participant populations for their studies (Buhmester et al., 2011). The worries about marketing studies, soliciting participants within the logistically available community, or scheduling appointment times are largely removed. Instead, a study is posted to MTurk, and the researcher simply waits for workers to complete the study. If the researcher pays participants well, the time it takes to collect data is significantly lower, with little risk of lowering the reliability of the data (Buhmester et al., 2011).

No method of data collection is perfect. As such, MTurk has its limitations. One such limitation is MTurk's anonymity. By default, the only personally identifiable information that HITs record is the worker's identification number. Participants might lie about background information that is needed to participate in the study, and researchers would have no way to verify if it is true. For this reason, studies that focus on a specific population (such as Hispanic women) may have difficulty finding a genuine set of participants. Second, because MTurk is an online system, it requires access to an electronic device with an Internet connection. Therefore, it may not be as easy to recruit people who have low socioeconomic status, such as the unemployed or homeless. Third, MTurk is also a poor choice for studies that involve children, require carefully controlled environments, or require special equipment (such as eyetrackers) or specimens (like blood samples). Because of this, online studies in general – and MTurk studies in particular – will never completely replace in person studies. Despite these limitations, MTurk provides a readily available, diverse population of people willing and able to complete psychological studies (Buhmester et al., 2011).

REFERENCES

- Amazon Mechanical Turk. (n.d.). Retrieved November 21, 2014, from <http://docs.aws.amazon.com/AWS-MechTurk/latest/AWSMTurkAPI/Welcome.html>
- Amazon Mechanical Turk. (n.d.). Retrieved November 21, 2014, from <https://www.mturk.com/mturk/help?helpPage=overview>
- Buhmester, M., Kwang, T., & Gosling, S. (2011). Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6(3), 3-5. Doi:10.1177/1745691610393980
- Kesebir, S. (2010). *Notes from the Amazon Mechanical Turk tutorial*. University of Virginia: Behavioral Research at Darden.
- Miziolek, J. (2011). CROWD SOURCING: CREATIVITY 2.0. *BrandPackaging*, 15(1), 16-17. Retrieved from <http://ezproxy.library.unlv.edu/login?url=http://search.proquest.com/docview/866321726?accountid=3611> Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5, 411.
- Worker Web Site FAQs. (2015). Retrieved November 21, 2015, from <https://www.mturk.com/mturk/help?helpPage=worker>