

2020

A Practical Guide for Using Electronic Surveys in Aviation Research: Best Practices Explained

Stephen Rice

ERAU, srice@outlook.com

Scott R. Winter

Embry-Riddle Aeronautical University, scott.winter@mac.com

Follow this and additional works at: <https://commons.erau.edu/ijaaa>



Part of the [Other Psychology Commons](#), and the [Transportation Commons](#)

Scholarly Commons Citation

Rice, S., & Winter, S. R. (2020). A Practical Guide for Using Electronic Surveys in Aviation Research: Best Practices Explained. *International Journal of Aviation, Aeronautics, and Aerospace*, 7(2). Retrieved from <https://commons.erau.edu/ijaaa/vol7/iss2/1>

This Position Paper is brought to you for free and open access by the Journals at Scholarly Commons. It has been accepted for inclusion in International Journal of Aviation, Aeronautics, and Aerospace by an authorized administrator of Scholarly Commons. For more information, please contact commons@erau.edu.

The use of surveys to collect data has a long and storied history in the social sciences. A review of top journals in some fields reveals that a majority of published research in those areas use surveys for data collection (Anderson et al., 2019). The reasons for this are varied and depend largely on the needs of the researcher. In some cases, researchers want to collect data quickly and inexpensively. Bringing participants into a laboratory setting can be costly and time-consuming. In contrast, passing out a survey exponentially speeds up the process and limits costs.

In other cases, researchers are unable to collect behavioral data due to practical or ethical limitations. For example, if one wants to know how pilots might respond to a hijacking scenario (Mehta, Rice, Winter, & Buza, 2017), then it would be impossible to replicate a real event in the aircraft to measure pilots' behavioral responses. Instead, the researcher would create a hypothetical scenario and present that to the pilot to see what their attitudes or intentions might be.

There are some obvious weaknesses to this approach. First, while attitudes and intentions correlate well with behaviors (Ajzen, 1985, 1991, 2002), this relationship is not guaranteed. A person might have racist attitudes towards a pilot of color, but never express them verbally due to the obvious consequences. Second, we can never be fully certain that respondents understand the questions they are being asked, are answering with due diligence, or are being honest in their responses.

While we acknowledge the strengths and weaknesses of using surveys for data collection (Rice et al., 2017), the purpose of this paper is not to debate those points. Instead, we wish to assume that the reader already plans to use a survey and wishes to have a practical guide for doing so which highlights best practices for successful data collection. We hope this paper provides such a guide when using surveys in aviation research.

What Research Questions Can Surveys Answer?

Most textbooks argue that your research question should drive your research method and research design (Edmonds & Kennedy, 2017; Gall, Gall, & Borg, 2007; Salkind, 2009). The research design relates to your overall plan in effectively addressing your research problem and questions. Various quantitative designs include experimental, quasi-experimental, and correlational, among others (Edmonds & Kennedy, 2017; Salkind, 2009). Some common qualitative designs include case studies, phenomenological, and ethnographies (Gall, Gall, & Borg, 2007). Both quantitative and qualitative methods are useful in different situations, and occasionally a researcher might wish to use both in a mixed-methods study (Edmonds & Kennedy, 2017).

The research methodology, on the other hand, refers to the process or strategy you use to implement your design. Some common examples include interviews, focus groups, and datamining (archival data analysis) (Leedy &

Ormrod, 2016). You can use these different methodologies with both quantitative and qualitative methods. Let's say that you wish to interview airport personnel about a rash of runway incursions in the past year. You can interview personnel and collect open-ended responses that you incorporate into your qualitative design. Or you can use the interviews to collect quantitative data via various scales and other numeric responses. Or you can do both.

A survey is an instrument that may be used by the researcher in both quantitative and qualitative research, and across a variety of methodologies (Leedy & Ormrod, 2016; Vogt, Gardner, & Haeffele, 2012). For example, we might conduct a phenomenological study on women in aviation, and one of the instruments could be a survey to capture their opinions about the difficulties of breaking into a largely male-dominated field. Another example might be that the researcher wishes to conduct a correlational study where they attempt to define the relationship between fear of heights and willingness to fly in an airplane.

We want to add that using surveys does not preclude the researcher from conducting a true experiment. Many studies published in the top aviation journals are in fact experiments. An experiment requires three conditions: a) the experimenter must manipulate the independent variable and at least have an experimental group and control group; b) the experimenter must control all other variables to ensure they are not causing the effect; and c) the experimenter must randomly assign participants to the different conditions. There are variations to these rules depending on the textbook; however, they all essentially say the same things (Edmonds & Kennedy, 2017; Gall, Gall, & Borg, 2007; Salkind, 2009).

Let's say that you wish to measure consumer willingness to fly in an autonomous air taxi (Ragbir, Rice, Winter, Choy, & Milner, 2020). You decide to conduct an experiment to measure a potential causal relationship between price and willingness to fly. In a very simple study, you might create two conditions: a) full price ticket, and b) free ticket. You provide participants with a survey that presents these two hypothetical situations to randomly assigned participants. The only difference between the two conditions is the price. You then measure willingness to fly using a valid scale (Rice et al., 2020) and you conclude that willingness to fly is greater when the ticket is free. While this was not conducted in a physical laboratory, you have still conducted an experiment.

Typically, if your research question addresses the need to capture a person's attitudes, intentions or opinions, then using a survey as your data collection instrument may be the best approach to use. If, however, you need to collect behavioral data to answer your research question, then you will probably not find a survey useful. As we mentioned above, we are assuming the reader has already made this decision, and now wishes to proceed with the survey.

Methods of Survey Distribution

Once you have decided on using a survey for your data collection, you may wish to consider the various ways to distribute your survey. First, you will want to decide if you need a paper survey or an electronic survey. There are advantages and disadvantages to each.

Paper Surveys

A paper survey is easy to pass out to the public or in a classroom setting. It ensures that everyone can participate even if they don't have an electronic device. They are generally easy to read, and some studies indicate that participants would rather fill out a paper survey compared to an electronic one (Hohwü et al., 2013; Wyatt, 2000). One downside of a paper survey is cost—you have to pay for the paper and printing costs. While this is not onerous for most experienced researchers with university or company funding, a student running a study might not find this affordable. Another downside is the time and energy needed to convert the data to electronic format for data processing. Lastly, paper surveys make randomizing any necessary components more difficult, often requiring multiple versions of the survey, which in turn complicates the coding of responses.

Electronic Surveys

Electronic surveys avoid the need for conversion of data from paper to electronic format. They also provide the researcher with a much wider and larger audience. Passing out paper surveys limits a researcher to those people she or he can directly interact with. Electronic surveys can be put online, where a researcher can instantly access the entire country or world. Furthermore, data collection goes much more quickly when using electronic surveys. One can put a survey online and collect thousands of responses in minutes. Many electronic survey programs allow the researchers to pre-code responses so when the data is exported, it is already numerically coded and ready for initial data analysis. In addition, one can stipulate certain demographics *a priori*. Do you want female British participants between the ages of 18-55? We discuss below how to go about doing this.

On the other hand, electronic surveys preclude the researcher from actually meeting their participants and one tends to lose control over the setting. If you need to ensure that your participants understand the verbiage and questions, then you may wish to conduct an in-person paper survey. An alternative is to use an iPad or laptop and have participants fill out their answers electronically in your presence, while still maintaining control over the setting.

Setting up the Survey

At this point, we assume the reader has decided to use an electronic survey. There are many choices of survey software. Perhaps the most famous is SurveyMonkey®. This is a software program that allows one to create a professional-looking survey in minutes. The layout and formatting are very nice, and give you choices of colors, font sizes, font types, and backgrounds, etc. You can put in text, pictures, videos, audio and other options. Once the data is collected, it can be easily exported to Microsoft® Excel® or a statistical processing software like IBM SPSS® or R®. Another example is Qualtrics®, which offers much of the same experience. Both of these programs require a paid subscription if you wish to get the most out of them. A free alternative is Google Forms®; however, this software does not have some of the features their competitors have, like randomizing the order of pages.

When creating your survey, you will have many choices to make. First, you must decide how many questions to ask. Our advice is to limit the number of questions to those that directly answer your research question. Many researchers are tempted to add in additional *curiosity* questions, but this may have the downside of boring your participants with too many questions. We have found that requiring participants to answer more than 50-60 questions typically results in a degradation of data; however, there are exceptions depending on the pay scale.

Next, you must decide what type of questions to ask. Do you want to use validated scales to measure your dependent variable, or are you looking for open-ended questions? There are many choices for types of questions, including Likert-scales, Likert-type scales, sliders, multiple choice grids, etc. You must determine the pros and cons of each and decide which one best serves your purposes.

Once you have completed the survey, you must test it. It seems obvious to say, but in our experience, bad surveys are often due to a lack of testing. Double and triple check everything. Run through the survey yourself and have someone else go through it as well. You want to make sure your participant understands everything you are asking. Nothing is more frustrating than getting bad data because you confused your participants. Due to the large populations of online platforms, it is fairly easy to conduct an adequate pilot study to verify the validity of your instrument before conducting the main data analysis. A number of these programs then also allow you to exclude the participants from the pilot study when you conduct the main data collection to ensure the assumption of independence is maintained.

If you are conducting an experiment, then you want to make sure you are randomly assigning your participants to the various groups. This can be done in many ways. For example, if you have two groups, then you may have a starter question that asks them for the last digit of their home address number. You can then assign the odd numbers to one group and even numbers to another. Or if you

need three groups, then you might use birthdays. This can be tricky because some values, such as birthdays, are not as random as they may appear — there is a slight bias to the latter part of the year and moms wanting a C-section tend to avoid the holidays.

Getting Participants

There are many sources for researchers to gather participants to fill out their surveys. Some choose to post a link to their survey on Twitter, Facebook, Reddit, Instagram or other online social media outlets. This can be effective if you have a large following and do not want to pay for participants. However, you are often limited to people who know you or follow you online. Sometimes your friends or followers will repost to their followers, but this is hit or miss.

Another source for getting participants is via an online paid portal. SurveyMonkey® offers a service where they will provide you with participants for a fee, but there are cheaper alternatives. There are a variety of other online paid portals (Peer, Brandimarte, Samat, & Acquisti, 2017), but the biggest and most popular one in the social sciences is Amazon's® Mechanical Turk® (MTurk).

MTurk is a portal developed by Amazon ® that was originally intended to be used by companies seeking one-off tasks completed by humans (called Turkers). Companies can post a Human Intelligence Task (HIT) and find someone to complete that task for pay without having to formally hire them. This is a very effective way of making brochures, editing papers, etc., without committing to a long-term hire.

Sometime around 2008, a few researchers in the social sciences began using this portal for their research studies (Paolacci, Chandler, & Ipeirotis, 2010; Schnoebelen & Kuperman, 2010). It gained popularity quickly in Psychology and expanded rapidly after Burhmester, Kwang, and Gosling (2011) published a sort of how-to guide for MTurk. Other researchers quickly tested the portal and found that the data collection was at least as reliable as laboratory data, and more generalizable since one could capture a larger and wider audience compared to university students (Berinsky, Huber, & Lenz, 2012; Chandler, Mueller, & Paolacci, 2014; Garrow, Chen, Ilbeigi, & Lurkin, 2020; Germine et al., 2012; Lee, Seo, & Siemsen, 2018; Smith, Sabat, Martinez, Weaver, & Xu, 2015). The authors of this paper began using it at about the same time and have to date conducted approximately 250 studies using MTurk (e.g., Rice, Winter, Mehta, & Ragbir, 2019; Winter, Rice, Rains, Milner, & Mehta, 2017).

Amazon's ® Mechanical Turk ®

The internet is full of guides on how to use MTurk, so we will not spend an inordinate amount of time on this; however, given that we have used it for almost

a decade now, we have some tips that the reader might find useful. We offer these best practices to help ensure the successful use of this platform.

Basic Demographics

Sample size. There are over a million Turkers, and it is not difficult to get several thousand of them to respond to a survey if you pay them. The amount of Turkers present depends largely on the day of the week and time of day. You can usually find the most Turkers during the workday, as many of them are bored at their jobs and do online studies to fill in the time (Mason & Suri, 2012). If you need participants from other countries, keep the time differences in mind. For example, if we want to collect data from India, we usually post the HIT late in the evening our time and download the data the next morning as their day is ending. When done correctly, we have no problem getting 2,000+ Turkers in an afternoon.

Location. The distribution of Turkers is not equal across countries. About 40% of Turkers are located in the United States. Another 40% are located in India, and the rest are spread out around the world. Some countries have little to no MTurk presence, so if you need data from Bolivia, for example, then you will have a difficult time finding participants.

Age. One convenience that Amazon provides is ensuring that all the Turkers are adults before they are allowed to sign up. Thus, the concern about accidentally collecting data from minors is alleviated. In our experience, the average age of Turkers who respond to our studies is about 33 years old with a standard deviation of approximately 11-12 years. The range is from 18 to 82, although it is possible there are even older respondents. This age range is far superior to the typical 18-19 year old averages seen in university labs, where students tend to be much younger than the general population (Chandler et al., 2019; Goodman, Cryder, & Cheema, 2013). Conducting aging studies on MTurk is much easier than in the lab.

Gender. Up until a few years ago, about 60-70% of Turkers were male (Paolacci, Chandler, & Ipeirotis, 2010), but this has changed dramatically in the past three years. We are now seeing a 50-50 split, and even more females in some cases (Hitlin, 2016). This varies depending on the country, so you will need to check your data if you are comparing genders. Only about one in 200 Turkers identify as 'other' (neither male nor female) in our experience.

Sexual orientation. The vast majority of Turkers mimic society's norms. We see about 95% heterosexual self-reporting, with the rest identifying as gay, bisexual, etc. Surprisingly, there are more self-identifying bisexuals in India compared to the United States; however, we are not sure if this is a reflection of their entire society or just the online presence.

Tips and Tricks for Researchers: Best Practices for Success

Pay rate. There has been a lot of controversy about Turkers and the pay rate. Many years ago, Turkers were willing to do a survey for a few pennies. This has changed; however, they are still barely making minimum wage and many people feel that researchers are taking advantage of an ‘online sweatshop.’ The counterargument is that most Turkers do not do this for a living—they are bored and would rather make a few bucks doing surveys than randomly surf the internet (Mason & Suri, 2012; Paolacci, Chandler, & Ipeirotis, 2010). Either way, you as the researcher get to decide how much to pay them, and the Turkers get to decide whether they agree. If you want more Turkers of a higher quality, then offer more money for a HIT. We have found that 25-50 cents is a common amount for a survey of 10-20 questions that takes a couple of minutes to complete.

Choosing your Turkers. You can set some basic demographics *a priori* if you are looking for specific people to fill out your surveys. These choices include gender, whereby you can pick only women, or only men, for a particular study. Or you can limit your Turkers to people from the United States or some other country. You can even limit them to any country *except* the United States or other country. You can also pay an additional fee if you wish to use MTurk Master Workers. These workers have demonstrated a high degree of success over an extended period of time and are typically more trustworthy compared to regular Turkers. In our experience, this difference is negligible and not worth the extra cost, but you can decide for yourself.

You can also limit your Turker sample to those workers who have previously completed a certain number of successful HITs. For example, you may only want Turkers who have previously completed 100, 500 or even 1000 previous HITs. Anyone who does not meet that standard is automatically excluded. You can also limit your Turker sample to those who have been the most successful on previous HITs. After each HIT, the researcher can approve or disapprove of the performance, which results in a lifetime approval rate for each Turker. We typically only allow Turkers who have at least a 98% approval rate, with at least 100 previous successful HITs. This action helps to ensure quality data is collected as these Turkers have demonstrated high levels of performance.

Lastly, you can block Turkers from future studies. If you feel they are not performing up to your standards, then you block them, and they will never see another survey you post on MTurk. We advise only doing this if you strongly feel they have wronged you, because this affects their reputation with Amazon and the community.

Paying your Turkers. Once the HIT is complete, you will need to pay the Turker for her or his time. This is conveniently done through Amazon. The money in your account is moved over to their account depending on how you wish to set

this up. You can have it done instantly, so they get paid as soon as they complete the HIT. We recommend not doing this until you've checked their performance.

If you need a bit of time to check the work, then you can choose how many hours before they get auto-paid. Or you can choose to manually pay each participant when you want. We recommend not taking too long with payment, as this can affect your reputation as a lab. There are many websites out there keeping track of you as well, and if you mistreat your Turkers, they will ignore your future studies. We typically auto-pay after one hour.

In order to ensure that the Turker actually did your survey, we recommend using a code at the end of the survey. The Turker then inputs this code on the MTurk site and this lets you know they got paid. We typically ask the Turker to provide two letters and two numbers to avoid duplicates.

Title your HIT. This is more important than you might think. The title is what Turkers see first. They see a lot of titles as they peruse the boards, and yours needs to stand out. Using a title like 'Online Study' is not useful and does not attract attention. Using a title like 'Rate these Airplanes' is much more effective. People like airplanes and they like rating things. Play around with your title and if you have trouble finding Turkers, then try changing it to something more dynamic. It is important not to lie in your title. Remember your reputation if you wish to do more than one study.

Instructions. Give your Turkers clear instructions. There is nothing more frustrating to a Turker than being confused by what you want. They want to get paid and now they're not sure if they made you happy. Many of them will quit the HIT to avoid receiving a Disapproval rating from you.

Answer emails. Turkers will sometimes need, or want, to email you. Perhaps they didn't understand the instructions and need clarification. Perhaps they messed up the code and want you to know so they don't get their work rejected. Sometimes we get a nice email telling us that they liked the survey and hope to see more. Answer all your emails politely and promptly. It helps to keep your reputation high and brings back repeat participants.

Conclusions

The purpose of this paper was to provide researchers with a guide to conducting survey research within the aviation field, particularly as it relates to using Internet-based samples. The paper provides a brief review of research methods, research designs, and research methodologies. Following the assumption that researchers have already decided to use a survey to answer their proposed research questions, a discussion on the advantages and disadvantages of paper or electronic administration of the instrument was discussed. Information was also presented regarding the use of online platforms for survey research, such as Amazon's Mechanical Turk. Basic demographic information was

provided for this population, and lastly, recommended best practices were shared to help researchers maximize the success of their future studies.

References

- Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckman (Eds.), *Action Control* (pp. 11-39). Berlin, Germany: Springer.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211.
- Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology*, 32(4), 665-683. doi:10.1111/j.1559-1816.2002.tb00236.x
- Anderson, C. A., Allen, J. J., Plante, C., Quigley-McBride, Lovett, A., & Rokkum J. N. (2019). The MTurkification of social and personality psychology. *Personality and Social Psychology Bulletin*, 45(6), 842-850.
- 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, 351-368.
- 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.
- Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods*, 46, 112-130.
- Chandler, J., Rosenzweig, C., Moss, A. J., Robinson, J., & Litman, L. (2019). Online panels in social science research: Expanding sampling methods beyond Mechanical Turk. *Behavior Research Methods*, 51, 2022-2038.
- Edmonds, W. A., & Kennedy, T. D. (2017). *An applied guide to research designs: Quantitative, qualitative, and mixed methods*. Los Angeles, CA: Sage.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction* (8th edition). Boston, MA: Pearson.
- Garrow, L. A., Chen, Z., Ilbeigi, M., & Lurkin, V. (2020). A new twist on the gig economy: Conducting surveys on Amazon Mechanical Turk. *Transportation*, 47, 23-42.
- Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the web as good as the lab? Comparable performance from web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin and Review*, 19, 847-857.
- Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making*, 26, 213-224.

- Hitlin, P. (2016). *Research in the crowdsourcing age, a case study*. Pew Research Center. Retrieved on March 23, 2020 from: <http://www.pewinternet.org/2016/07/11/research-in-the-crowdsourcing-age-a-case-study/>
- Hohwü, L., Lyshol, H., Gissler, M., Jonsson, S. H., Petzold, M., & Obel, C. (2013). Web-based versus traditional paper questionnaires: A mixed-mode survey with a Nordic perspective. *Journal of Medical Internet Research*, *15*(8), e173. Retrieved from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3757995/>
- Lee, Y. S., Seo, Y. W., & Siemsen, E. (2018). Running behavioral operations experiments using Amazon's Mechanical Turk. *Production and Operations Management*, *27*(5), 973-989.
- Leedy, P. D. & Ormrod, J. E. (2016). *Practical research: Planning and design* (11th edition). Boston, MA: Pearson.
- Mason, W. & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavioral Research*, *44*, 1-23.
- Mehta, R., Rice, S., Winter, S. R., & Buza, P. (2017). Cabin depressurization as a hijacking mitigation tactic: A consumer perceptions study. *Collegiate Aviation Review International*, *35*(1), 13-24.
- Mellis, A. M. & Bickel, W. K. (2020). Mechanical Turk data collection in addiction research: Utility, concerns and best practices. *Addiction*, *115*, 1-9.
- Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*(5), 411-419.
- Peer, E., Brandimarte, L., Samat, A., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153-163.
- Ragbir, N. K., Rice, S., Winter, S. R., Choy, E. C., & Milner, M. N. (2020). How weather, distance, flight time, and geography affect consumer willingness to fly in autonomous air taxis. *Collegiate Aviation Review International*, *38*(1), 69-87.
- Rice, S., Winter, S. R., Capps, J., Trombley, J., Robbins, J. Milner, M. N., & Lamb, T. L. (2020). Creation of two valid scales: Willingness to fly in an aircraft and willingness to pilot an aircraft. *The International Journal of Aviation, Aeronautics, and Aerospace*, *7*(1), 1-21.
- Rice, S., Winter, S. R., Doherty, S. & Milner, M. N. (2017). Advantages and disadvantages of using internet-based survey methods in aviation-related research. *Journal of Aviation Technology and Engineering*, *7*(1), 58-65.
- Rice, S., Winter, S. R., Mehta, R., & Ragbir, N. K. (2019). What factors predict the type of person who is willing to fly in an autonomous commercial airplane? *Journal of Air Transport Management*, *75*, 131-138.
- Salkind, N. (2009). *Exploring research* (7th edition). Upper Saddle River, NJ:

Pearson.

- Schnoebelen, T. & Kuperman, V. (2010). Using Amazon Mechanical Turk for linguistic research. *Psihologija*, 43(4), 441-464.
- Smith, N. A., Sabat, I. E., Martinez, L. R., Weaver, K., & Xu, S. (2015). A convenient solution: Using MTurk to sample from hard-to-reach populations. *Industrial and Organizational Psychology*, 8, 220-228.
- Vogt, W. P., Gardner, D. C., & Haeffele, L. M. (2012). *When to use what research design*. New York, NY: The Guilford Press.
- Winter, S. R., Rice, S., Rains, T., Milner, M., & Mehta, R. (2017). A longitudinal study on the alteration of consumer perceptions and the use of pilot medication. *Journal of Air Transport Management*, 59, 100-106.
- Wyatt, J. C. (2000). When to use web-based surveys. *Journal of the American Medical Informatics Association*, 7(4), 426-430.