JOURNAL FOR DERADICALIZATION

Spring 2021
Nr. 26
ISSN: 2363-9849

Research Methods Brief: Attrition Happens (and What to Do About It)

Michael J. Williams[a1]
[a]The Science of P/CVE

Abstract
Attrition (participant "dropout") is the loss of participants from a program/initiative or longitudinal (e.g., pre/post) data collection. If participants dropout for non-random, systematic reasons, those factors bias the sample and limit the study or evaluation's generalizability. The importance of statistically diagnosing participant attrition can scarcely be overstated, given that P/CVE research and evaluations are commonly concerned, not merely with the results from a given sample of participants, but whether, how, or to what extent the results might generalize to other, perhaps much broader samples. Therefore, the threat to generalizability, posed by non-random participant attrition, threatens the very reason for conducting many, if not most, P/CVE-related research and evaluations.

Non-random attrition prevents research and evaluations from making valid claims or inferences about their target populations, and to know whether attrition likely threatens the validity of a project's findings, one must test for it. The present article includes step-by-step guidance on how to diagnose participant attrition, including discussion of the implications: implications that potentially can salvage a P/CVE-related program from seemingly problematic participant attrition.

Article History
Received Jan 18, 2021
Accepted Mar 16, 2021
Published Mar 26, 2021

Keywords: Attrition, Generalizability, Countering Violent Extremism, Preventing Violent Extremism, CVE, PVE, Evaluation, Research Method

Introduction

Attrition (aka participant "dropout") is the loss of participants from a program/initiative or longitudinal (e.g., "pre/post") data collection, and it tends to be worse the longer the timeframe (West et al., 2004). Of course, the best way to deal with attrition is to prevent it,

[1] Corresponding Author Contact: Michael J. Williams, Email: m.williams@thescienceofpcve.org, Twitter: @MickWilliamsPhD, ORCID iD https://orcid.org/0000-0001-5630-9814

Spring 2021
Nr. 26
ISSN: 2363-9849

and there exists a body of knowledge, and associated techniques, designed to enhance participant retention (e.g., see Davis et al., 2002). Nevertheless, attrition happens; the question is, to what extent it matters.

If participants dropout *completely at random*, that is not a problem (from a data analytic perspective); that doesn't threaten the findings' generalizability. The only analytic drawback to completely random attrition is the reduction in sample size and commensurate reduction in statistical power (i.e., the ability to detect significant program effects/outcomes). Even if attrition isn't completely random, but is entirely explainable by other measured variables (e.g., attrition attributable to age, sex, etc.) the analyses still can be performed (Bhaskaran & Smeeth, 2014).[2]

However, if participants dropout *for a reason*—because of something systematic (perhaps the program is disproportionately uncomfortable for certain participants, or there are logistical barriers that tend to cause certain participants to dropout)—such factors bias the sample and limit the study or evaluation's generalizability to the participant characteristics of those remaining in the study (Kazdin, 2003). In other words, **non-random attrition prevents research and evaluations from making valid claims or inferences about their target populations**, and *to know whether attrition likely threatens the validity of a project's findings, one must test for it.* (See below for the recommended method for this test.)

The importance of statistically diagnosing participant attrition can scarcely be overstated, given that P/CVE research and evaluations are commonly concerned, not merely with the results from a given sample of participants, but whether, how, or to what extent the results might generalize to other, perhaps much broader samples. Therefore, the threat to

[2] If missing data are attributable to other measured variables, the data within the strata of those measured variables might be missing completely at random, and—as such—can by analyzed within those strata without concern for the biasing effects of missing data (Bhaskaran & Smeeth, 2014).

Alternatively, data that are missing either completely at random (MCAR) or missing at random (MAR) can be estimated (e.g., through Multiple Imputation or Full Information Maximum Likelihood [FIML] estimation), and analyses can continue as otherwise intended.

Bear in mind that if more than one variable has missing data, those variables' data might be missing for different reasons (e.g., MCAR, MAR, or missing not at random MNAR). Therefore, although it can be time-consuming, it is necessary to diagnose the degree of randomness of missing data per variable (Grace-Martin, n.d.).

generalizability, posed by non-random participant attrition, threatens the very reason for conducting many, if not most, P/CVE-related research and evaluations. **Consequently, it is astonishing that analysis and discussion of attrition has not been routinely integrated into P/CVE research and evaluations**. Consider the following thought experiment: when was the last time that you read a P/CVE-related report where participant attrition was reported, not to mention statistically diagnosed, and (if applicable) mitigated?[3]

How to Diagnose Participant Attrition

To diagnose attrition, the first data-analytic step is to create a variable in the dataset (call it "Attrition," or the like) and code each participant (i.e., each case, each unit of analysis) on that variable using a binary system (e.g., 1 = attriter vs. 0 = completer). Then, using logistic regression, the variable "attrition" is set as the dependent variable, and the variable that codes each case for its group membership (e.g., comparison group 1 vs. comparison group 2, etc.) is set as the independent variable—along with any pretest/program intake measure(s) of interest (hereafter "pretest"), *and* the interaction of group membership with the pretest measure(s)—to determine if group membership, and/or any pretest measures, are significantly associated with attrition. In other words, this test reveals the extent to which attrition significantly varies between and across the groups. Of course, one hope that it does not significantly vary: by convention, resulting in *p*-values greater than .05.[4] In performing this test, with one or more pretest variables, it is important *not* to adjust the *p*-values for alpha slippage (aka "alpha inflation," West et al., 2000). Such adjustments (e.g., the Bonferroni correction) are

[3] See Koehler (2017) for an early work that spoke to the importance of measuring attrition of P/CVE programs.
[4] This assumes that the analysis has been performed properly: notably, that any/all outliers have been identified through inspection of the residuals and their associated metrics of influence (e.g., Cook's Distance and Leverage). Typically, any case that is beyond conventional tolerances for the residuals, and one or more influence metrics, may be considered an outlier and should be considered for exclusion from the analysis.

Additionally, logistic regression is sensitive to high correlations (multicollinearity) between predictor variables. Therefore, if more than just the variable for "group membership" is entered as an independent variable, collinearity diagnostics must be examined. Although there is no universally accepted cutoff value for determining the presence of multicollinearity (Senaviratna & Cooray, 2019), tolerance values as severe as ≤ 0.1 can be considered cause for concern (ibid.), though tolerance values < .2 might also indicate problematic multicollinearity (Menard, 2002).

counterproductive to the sensitivity of this procedure intended to identify possible causes of attrition (ibid.).

Main Effect for Pretest Measure(s)

If the logistic regression reveals a significant main effect for a given pretest measure, that indicates that one should be especially reluctant to generalize the results of the program to the broader population of interest (i.e., the population from which the sample was selected, West et al., 2004). Such a finding indicates that participants remaining in the program are no longer equivalent to the population from which they were sampled; specifically, they differ (at least) according to the pretest measure(s) that demonstrated a significant main effect. In short, the difference(s) between those remaining in the program, vs. those of the sampled population, might be at least partially responsible for a program's outcomes.

Main Effect for Group Membership/Condition

If the logistic regression reveals that attrition is significantly associated with a main effect for group membership (i.e., condition)—without a significant interaction between group membership and pretest scores—that indicates that at least one of the groups is systematically, and especially, prone to dropouts/failures. In other words, if there is significantly different attrition, based on group membership, it indicates that there is something about the program/intervention that affects the likelihood that participants will (dis)continue their participation (West et al., 2004). For example, one of the conditions might be relatively too difficult (or otherwise frustrating) to participants, or—conversely—it might be too easy (or otherwise boring) to them. Therefore, if there is a main effect for group membership, posttest outcomes of the program cannot be attributed solely (if at all) to the intervention itself, but to another factor that, at least partially, influences the extent to which individuals are willing to continue with the program.

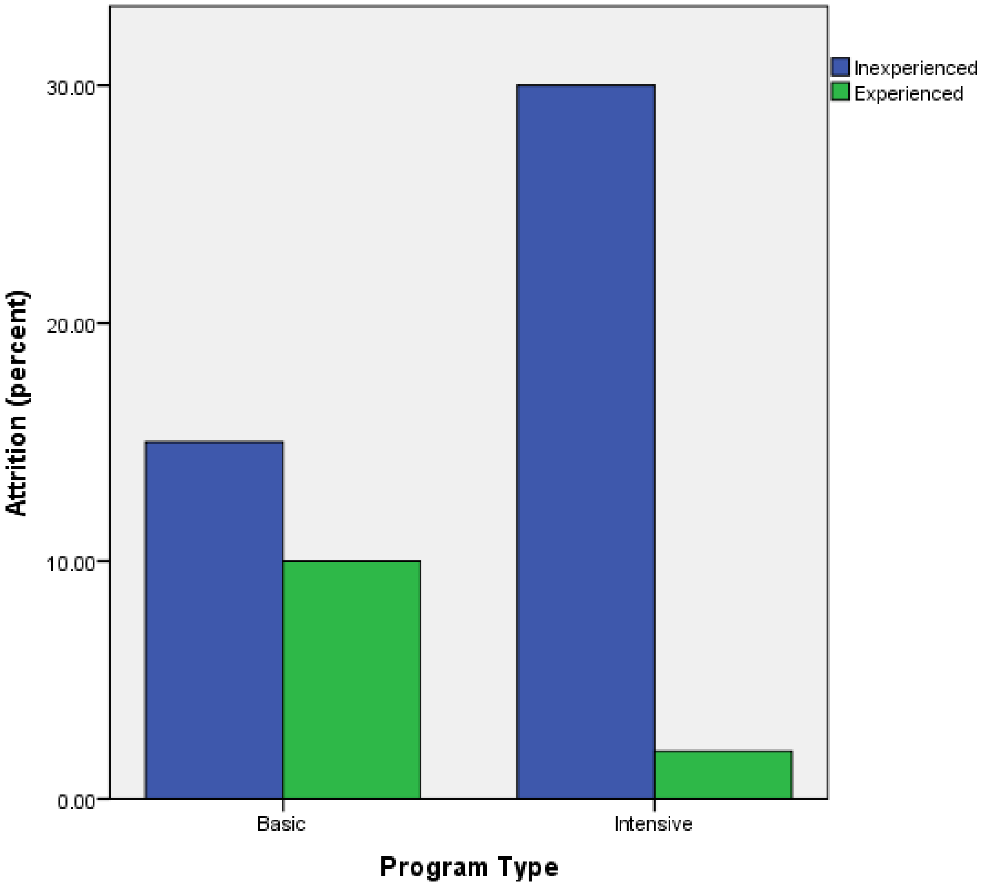
Interaction of Pretest Measure(S) with Group Membership/Condition

If the logistic regression reveals that attrition is significantly associated with the interaction of participants' pretest measure(s) and their group membership, that indicates that whether or not participants dropout from (or continue with) the program is associated with (i.e., predictable by) the measured pretest attribute(s). At first glance, a significant interaction would seem to indicate yet another shortcoming, if not failure, of the program. However, diagnosing the interaction (i.e., parsing the so-called "simple effects" of the significant pretest measure, per group) will reveal how attrition varies between the groups, according to the pretest measure(s). By doing so, the simple effects might reveal that attrition is relatively low for a certain type of participant.

As a hypothetical example, consider two types of P/CVE-related training programs: one featuring an intensive (e.g., lengthier and/or challenging) curriculum vs. one featuring a basic (e.g., shorter and/or less challenging) curriculum. It would not be surprising to find a main effect of attrition for the intensive program. However, curriculum type might interact with certain pretest attributes: for example, participants' level of prior experience in the domain taught by the curriculum. If so, as depicted in Figure 1, one might find that attrition is especially low (perhaps at its lowest)—even in the intensive program—among those who are relatively more experienced. In short, the more challenging curriculum appealed to this type of participant, who tended to continue with the program.

Figure 1

Attrition by Program Type and Participants' Level of Experience



This example also touches upon a common misunderstanding that, if attrition rates are equivalent between program types (i.e., between groups), there is unlikely to be any interpretive problems with respect to comparing those programs' outcomes (Tabachnick & Fidell, 2007; West et al., 2004). To illustrate why that cannot be assumed, consider (for example) that in a basic/easy program, the "most skilled" 5% of participants might drop out (e.g., they become bored, or have little need for what the program has to offer); whereas, in an intensive/challenging program, the "least skilled" 5% of participants might drop out (e.g., due to frustration with the challenges of the program). Such attrition rates, though equivalent, might make one of the programs appear more/less effective—or, conversely, might make the programs appear to have null/equivalent outcomes—not because of the programs' approaches per se, but because of the performance of those who did not drop out of the programs. Therefore, even equivalent attrition across groups should be subject to closer inspection (i.e., tested for an interaction between group membership and participants' pretest/intake attributes, Tabachnick & Fidell, 2007).

Implications of an Interaction of Pretest Measure(S) with Group Membership

As illustrated by Figure 1, diagnosing a significant interaction between the pretest measure(s) and group membership may reveal, not only who is most likely to discontinue their participation, but—conversely—who is most likely to continue their participation. As mentioned, in the present example, attrition was at its lowest, in the intensive program, among those who were relatively more experienced. This statistical diagnosis is important for (at least) the follow three reasons. First, it serves as evidence that the program might not be critically flawed. Instead, it points to whom the program might be relatively successful (i.e., the type of participants who remained in the program). As such, the second way that this diagnosis is important is that it might help program managers to refocus the program's content or activities to orient them toward those for whom attrition was relatively low. Concomitantly, the third way that this diagnosis is important is that it might help program managers to refocus the program's marketing or recruitment strategies to orient them either toward those for whom attrition was relatively low, or to vary the marketing or recruitment strategies according to the pretest-by-group interaction. (In the previous example, marketing the intensive curriculum to those who have greater experience would likely be more appealing/effective, than to those with lesser experience, and vice versa with respect to marketing the basic curriculum.)

Therefore, this bit of "bad news"—that there is an aspect of the intervention that is influencing dropouts—could be good news in disguise: as mentioned, that the program could be revised to tailor it to the types of participants who might benefit most from a given program type. Furthermore, this might provide ancillary benefits, such as informing a streamlined (perhaps more efficient) curriculum, or informing a more focused (perhaps less costly) marketing strategy, or informing participant triage, so that participants are enrolled in the program type that they are most likely to complete, based on their pretest attributes.

Conclusion

One would be unable to take advantage of the aforementioned insights unless attrition is both measured and diagnosed as described above. Furthermore, as mentioned, presumed programmatic outcomes might actually be illusory: at least partially (if not largely) the result of attrition, instead of the result of the program's content or activities per se. In short, there is simply no sound reason not to measure and diagnose attrition, and—at the risk of placing too strong a point on the matter—to fail to do so is tantamount to malpractice.

Consequently, attrition should be analyzed at the earliest opportunity—if possible, while the program is ongoing, or (better yet) in a program's pilot phase—to diagnose whether attrition represents a systemic problem with the program that might warrant immediate correction. Even if analyses reveal that attrition is not significantly associated with group membership and/or pretest measures, it is ethical practice to describe the sample characteristics (e.g., demographics) of attriters in reports about a given P/CVE program (Williams, 2020).

If your projects have never measured or diagnosed attrition, so be it, but—from this time forward—resolve always to measure and diagnose this critical (and potentially program-salvaging) factor. We owe this, not only to those who fund our P/CVE programs, but to our P/CVE program participants who might benefit from a program that is better tailored to reach them and serve their needs.

Spring 2021
Nr. 26
ISSN: 2363-9849

References

Bhaskaran, K., & Smeeth, L. (2014). What is the difference between missing completely at random and missing at random? *International Journal of Epidemiology*, *43*(4), 1336–1339. https://doi.org/10.1093/ije/dyu080

Davis, L. L., Broome, M. E., & Cox, R. P. (2002). Maximizing retention in community-based clinical trials. *Journal of Nursing Scholarship*, *34*(1), 47–53.

Grace-Martin, K. (n.d.). *How to Diagnose the Missing Data Mechanism*. The Analysis Factor. Retrieved December 19, 2020, from https://www.theanalysisfactor.com/missing-data-mechanism/

Kazdin, A. E. (2003). Drawing valid inferences I: Internal and external validity. In *Research design in clinical psychology*. Allyn & Bacon.

Koehler, D. (2017). *Structural quality standards for work to intervene with and counter violent extremism*. Counter Extremism Network Coordination Unit (KPEBW).

Menard, S. (2002). *Applied logistic regression analysis* (Vol. 106). Sage.

Senaviratna, N. A. M. R., & Cooray, T. M. J. A. (2019). Diagnosing multicollinearity of logistic regression models. *Asian Journal of Probability and Statistics*, *5*(2), 1–9. https://doi.org/10.9734/ajpas/2019/v5i230132

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Allyn & Bacon/Pearson Education.

West, S. G., Biesanz, J. C., & Kwok, O. M. (2004). Within-subject and longitudinal experiments: Design and analysis issues. In C. Sansone, C. C. Morf, & A. T. Panter (Eds.), *The SAGE Handbook of Methods in Social Psychology* (pp. 287–312). https://doi.org/10.4135/9781412976190.n13

West, S. G., Biesanz, J. C., & Pitts, S. C. (2000). Causal inference and generalization in field settings: Experimental and quasi-experimental designs. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 40–84). Cambridge University Press.

Williams, M. J. (2020). *Preventing & countering violent extremism: Designing and evaluating evidence-based programs*. Routledge. https://thescienceofpcve.org/book/

About the JD Journal for Deradicalization

The JD Journal for Deradicalization is the world's only peer reviewed periodical for the theory and practice of deradicalization with a wide international audience. Named an "essential journal of our times" (Cheryl LaGuardia, Harvard University) the JD's editorial board of expert advisors includes some of the most renowned scholars in the field of deradicalization studies, such as Prof. Dr. John G. Horgan (Georgia State University); Prof. Dr. Tore Bjørgo (Norwegian Police University College); Prof. Dr. Mark Dechesne (Leiden University); Prof. Dr. Cynthia Miller-Idriss (American University Washington); Prof. Dr. Julie Chernov Hwang (Goucher College); Prof. Dr. Marco Lombardi, (Università Cattolica del Sacro Cuore Milano); Dr. Paul Jackson (University of Northampton); Professor Michael Freeden, (University of Nottingham); Professor Hamed El-Sa'id (Manchester Metropolitan University); Prof. Sadeq Rahimi (University of Saskatchewan, Harvard Medical School), Dr. Omar Ashour (University of Exeter), Prof. Neil Ferguson (Liverpool Hope University), Prof. Sarah Marsden (Lancaster University), Prof. Maura Conway (Dublin City University), Dr. Kurt Braddock (Pennsylvania State University), Dr. Michael J. Williams (The Science of P/CVE), and Dr. Aaron Y. Zelin (Washington Institute for Near East Policy), Prof. Dr. Adrian Cherney (University of Queensland).

For more information please see: www.journal-derad.com

Twitter: @JD_JournalDerad
Facebook: www.facebook.com/deradicalisation

The JD Journal for Deradicalization is a proud member of the Directory of Open Access Journals (DOAJ).

ISSN: 2363-9849

Editor in Chief: Daniel Koehler