**Survey research methodology: An annotated bibliography**

Prof Gavin T L Brown ©2022, *The University of Auckland*

This bibliography is designed for the provisional year doctoral student who thinks that a survey would be a good research tool. This bibliography focuses on the various methodological choices and challenges faced when trying to demonstrate that the candidate knows what and why a survey should be implemented. Scholarship of method is important because method matters. The results we depend on are a function of the validity of the methods we use. This document is open-access on a CC-BY basis and is available at <https://doi.org/10.17608/k6.auckland.19791727>

**Planning**

What are the philosophic assumptions you make that legitimise carrying out survey research? I recommend scientific hypothetico-deductive and realist positions, rather than pragmatic or qualitative frameworks. How would you defend such a stance?

* Chakravartty, A. (2016). Scientific Realism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 ed.). Stanford University, Metaphysics Research Lab. <https://plato.stanford.edu/archives/win2016/entries/scientific-realism/>
* Lawson, A. E. (2015). Hypothetico-deductive Method. In R. Gunstone (Ed.), *Encyclopedia of Science Education* (pp. 471-472). Springer Netherlands. <https://doi.org/10.1007/978-94-007-2150-0_260>
* Manicas, P. T. (2006). *A realist philosophy of social science: Explanation and understanding*. Cambridge University Press.
* Nola, R. (2007). The hypothetico-deductive method. In *Theories of Scientific Method: An Introduction* (pp. 170-184). Acumen Publishing. https://doi.org/DOI: 10.1017/UPO9781844653881.008
* Philips, D. C., & Burbules, N. C. (2000). *Postpositivism and Educational Research*. Rowman & Littlefield Publishers.

Why is a survey a good method to use? What can you achieve with it that you can’t do otherwise? What are the downsides of this method?

* Fowler Jr., F. J. (1993). *Survey research methods* (2nd ed.). Sage.
* Rea, L. M., & Parker, R. A. (2005). *Designing and conducting survey research: A comprehensive guide*. Jossey-Bass.
* Visser, P. S., Krosnick, J. A., & Lavrakas, P. J. (2000). Survey research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of Research Methods in Social and Personality Psychology* (pp. 223-252). Cambridge University Press.

How should questions or items be written?

* Frary, R. B. (1996). Hints for designing effective questionnaires. *Practical Assessment, Research & Evaluation*, *5*(3). <https://doi.org/10.7275/h53m-b438>
* Oppenheim, A. N. (1966). *Questionnaire design and attitude measurement.* Gower.
* Siniscalco, M. T., & Auriat, N. (2006). *Questionnaire design*. UNESCO, International Institute for Educational Planning. <http://www.sacmeq.org/sites/default/files/sacmeq/training-modules/sacmeq-training-module-8.pdf>
* Sudman, S., & Bradburn, N. M. (1982). *Asking Questions: A Practical Guide to Questionnaire Design*. Jossey-Bass.
* Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about Answers: The Application of Cognitive Processes to Survey Methodology.* Jossey-Bass Publishers.

What kind of response system should you use assuming you will use structured responding?

* Brown, G. T. L., & Shulruf, B. (in review). Response Option Design in Surveys. In L. R. Ford & T. A. Scandura (Eds.), *The SAGE Handbook of Survey Development and Application*. Sage.
* Taherdoost, H. (2019). What Is the Best Response Scale for Survey and Questionnaire Design; Review of Different Lengths of Rating Scale / Attitude Scale / Likert Scale. *International Journal of Academic Research in Management (IJARM)*, *8*. <https://hal.archives-ouvertes.fr/hal-02557308>
* Toepoel, V., Das, M., & van Soest, A. (2009). Design of Web Questionnaires: The Effect of Layout in Rating Scales. *Journal of Official Statistics*, *25*(4), 509-528. <https://doi.org/10.1177/0049124108327123>
* Weng, L.-J. (2004). Impact of the number of response categories and anchor labels on Coefficient Alpha and test-retest reliability. *Educational and Psychological Measurement*, *64*(6), 956-972. <https://doi.org/10.1177/0013164404268674>

How many people do you need?

* Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160. <https://doi.org/10.3758/BRM.41.4.1149>
* Hibberts, M., Johnson, R. B., & Hudson, K. (2012). Common survey sampling techniques. In L. Gideon (Ed.), *Handbook of Survey Methodology for the Social Sciences* (pp. 53-74). Springer. <https://doi.org/10.1007/978-1-4614-3876-2_5>
* Maxwell, S. E., Kelley, K., & Rausch, J. R. (2008). Sample Size Planning for Statistical Power and Accuracy in Parameter Estimation. *Annual Review of Psychology*, *59*(1), 537-563. <https://doi.org/10.1146/annurev.psych.59.103006.093735>

How many items do you need for each construct?

* Marsh, H. W., Hau, K.-T., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research*, *33*(2), 181-220. <https://doi.org/10.1207/s15327906mbr3302_1>
* Osborne, J. W., & Costello, A. B. (2004). Sample size and subject to item ratio in principal components analysis. *Practical Assessment, Research, and Evaluation*, *9*. <https://doi.org/10.7275/ktzq-jq66>

**Administration**

How do you reach your participants? You probably can’t use a true random sample because you can’t get lists of people in your population. So how do you recruit?

* Christensen, T., Riis, A. H., Hatch, E. E., Wise, L. A., Nielsen, M. G., Rothman, K. J., Toft Sørensen, H., & Mikkelsen, E. M. (2017). Costs and efficiency of online and offline recruitment methods: A web-based cohort study. *Journal of Medical Internet Research*, *19*(3), e58. <https://doi.org/10.2196/jmir.6716>
* Farrokhi, F., & Mahmoudi-Hamidabad, A. (2012). Rethinking convenience sampling: Defining quality criteria. *Theory and Practice in Language Studies*, *2*(4), 784-792. <https://doi.org/10.4304/tpls.2.4.784-792>

Sometimes you just have to use the people who show up, on a convenience basis.

* Sedgwick, P. (2013). Convenience sampling. *BMJ*, *347*, f6304. <https://doi.org/10.1136/bmj.f6304>

Other times you can ask people to pass on your survey to others. This is sometimes called chain-referral and often snowball recruitment.

* Gile, K. J., & Handcock, M. S. (2010). Respondent-driven sampling: An assessment of current methodology. *Sociological methodology*, *40*(1), 285-327. <https://doi.org/10.1111/j.1467-9531.2010.01223.x>
* Sedgwick, P. (2013). Snowball sampling. *BMJ : British Medical Journal*, *347*, f7511. <https://doi.org/10.1136/bmj.f7511>

How will you survey people? Telephone, mail, web/internet? Are these equivalent modes? What kinds of threat do they introduce? Does it change how many people you need?

* Alvarez, R. M., & Van Beselaere, C. (2005). Web-based survey. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (Vol. 3, P-Y, pp. 955-962). Elsevier Academic Press.
* Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method* (4th ed.). Wiley.
* Mei, B., & Brown, G. T. L. (2018). Conducting online surveys in China. *Social Science Computer Review*, *36*(6), 721-734. <https://doi.org/10.1177/0894439317729340>
* Ritter, P., Lorig, K., Laurent, D., & Matthews, K. (2010). Internet versus mailed questionnaires: A randomized comparison. *Journal of Medical Internet Research*, *6*(3), 3-7. <https://doi.org/10.2196/jmir.6.3.e29>

What about missing data? Depending on the mode of administration there might be missing values. How much missing is too much? How can you keep participants who might not have answered 100% of questions?

* Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of abnormal psychology*, *112*, 545-557. <https://doi.org/10.1037/0021-843X.112.4.545>
* Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). John Wiley & Sons.
* Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147-177. <https://doi.org/10.1037//1082-989X.7.2.147>

Sometimes, to save participants, you might choose not to give everyone all the same items or questions. How do you handle data when the missingness is planned?

* Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods*, *11*(4), 323-343. <https://doi.org/10.1037/1082-989X.11.4.323>
* Rhemtulla, M., & Hancock, G. R. (2016). Planned Missing Data Designs in Educational Psychology Research. *Educational Psychologist*, *51*(3-4), 305-316. <https://doi.org/10.1080/00461520.2016.1208094>

Not everyone who participated is out to tell you the honest truth. Sometimes they try to fool your survey. How would you detect this if it is possible in the population you are surveying?

* Fan, X., Miller, B. C., Park, K.-E., Winward, B. W., Christensen, M., Grotevant, H. D., & Tai, R. H. (2006). An exploratory study about inaccuracy and invalidity in adolescent self-report surveys. *Field Methods*, *18*, 223-244. <https://doi.org/10.1177/152822X06289161>

**Analysis**

Once you get your data back how will you analyse it? Have you used a conventional psychological approach that infers a latent construct from responses to manifest variables that conceptually aggregate into a factor? The basis for doing this is latent trait theory and the Multi-Item, Multi-Indicator (MIMIC) framework:

* Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*. Cambridge University Press.
* Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, *70*(351a), 631-639. <https://doi.org/10.1080/01621459.1975.10482485>

What is your approach to determining if items group into pools according to your design? The general technique is called common factor analysis. This presumes that we can isolate ‘error’ from ‘true’ signal in the pattern of responses to detect items go together. If you have a design or have borrowed someone else’s inventory then you should test the inventory first with confirmatory factor analysis.

* Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.
* Bandalos, D. L., & Finney, S. J. (2010). Factor analysis: Exploratory and confirmatory. In G. R. Hancock & R. O. Mueller (Eds.), *The Reviewer's Guide to Quantitative Methods in the Social Sciences* (pp. 93-114). Routledge.

There are lots of technical issues in doing CFA but what you need to know is what it means and how to decide if the result supports your conclusion. This is to do with model fit. There are debates about what constitutes acceptable or good fit.

* Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research*, *42*(3), 509–529. <https://doi.org/10.1080/00273170701382864>
* Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1-55. <https://doi.org/10.1080/10705519909540118>
* Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling*, *10*(3), 333-351.
* Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*(3), 320-341. [https://doi.org/doi:10.1207/s15328007sem1103\_2](https://doi.org/doi%3A10.1207/s15328007sem1103_2)

If your preferred model does not fit the data, what will you do? The usual solution is to investigate how many factors there are using an exploratory technique. You need to be able to explain why you didn’t do this first, by the way.

* Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, *10*(7). <https://doi.org/10.7275/jyj1-4868>
* Courtney, M. G. R. (2013). Determining the Number of Factors to Retain in EFA : Using the SPSS R-Menu v2 . 0 to Make More Judicious Estimations. *Practical Assessment Research & Evaluation*, *18*. <https://doi.org/10.7275/9cf5-2m72>
* Hoyle, R. H., & Duvall, J. L. (2004). Determining the number of factors in exploratory and confirmatory factor analysis. In D. Kaplan (Ed.), *The SAGE Handbook of Quantitative Methodology for Social Sciences* (pp. 301-315). Sage.

Assuming you get a working solution for each construct, how should you relate the various constructs to each other? While inter-correlated models are easy to test and will usually fit well, you might actually think there is a causal relationship (i.e., A makes changes in B). Here are some ideas about causation.

* Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect*. Hachette Book Group.
* Wegener, D. T., & Fabrigar, L. R. (2000). Analysis and design for nonexperimental data: Addressing causal and noncausal hypotheses. In H. T. Reis & C. M. Judd (Eds.), *Handbook of Research Methods in Social and Personality Psychology* (pp. 412-450). Cambridge University Press.
* 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.

The statistical approach to causal analysis is called structural equation modeling which uses many of the same ideas as CFA (esp. in terms of fit). The difference is that it comes after establishing that each construct has been measured in a robust way.

* Bollen, K. A. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons, Inc. [https://doi.org/doi:10.1002/9781118619179](https://doi.org/doi%3A10.1002/9781118619179)
* Klem, L. (2000). Structural equation modeling. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and Understanding More Multivariate Statistics* (pp. 227-260). APA.
* Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling* (3rd ed.). Guilford Press.
* Maruyama, G. M. (1998). *Basics of Structural Equation Modeling*. Sage.
* Mulaik, S. A. (2009). *Linear causal modeling with structural equations*. <https://doi.org/10.1201/9781439800393>
* Raykov, T., & Marcoulides, G. A. (2007). *A first course in structural equation modeling*. Psychology Press.

When you plan to make comparisons between groups, you have to first establish if the measurement models are equivalent. There are 3 levels of equivalence that matter: configural (the same pattern of items and paths), metric (the same strength of loading from factors to items), and scalar (the same starting point for the regression path).

* Brown, G. T. L., Harris, L. R., O'Quin, C., & Lane, K. E. (2017). Using multi-group confirmatory factor analysis to evaluate cross-cultural research: identifying and understanding non-invariance. *International Journal of Research & Method in Education, 40*(1), 66-90. <https://doi.org/10.1080/1743727X.2015.1070823>
* Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233-255. [https://doi.org/doi:10.1207/S15328007SEM0902\_5](https://doi.org/doi%3A10.1207/S15328007SEM0902_5)
* Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(4), 4-70. <https://doi.org/10.1177/109442810031002>

How do you think you will analyse the data? In addition to path values, we are often interested in mean scores, especially when a model is equivalent between groups.

* DiStefano, C., Zhu, M., & Mîndrilă, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation*, *14*(20). <https://doi.org/10.7275/da8t-4g52>
* Sörbom, D. (1974). A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematical and Statistical Psychology, 27*(2), 229-239. <https://doi.org/10.1111/j.2044-8317.1974.tb00543.x>
* Thompson, M. S., & Green, S. B. (2006). Evaluating between-group differences in latent variable means. In G. R. Hancock & R. O. Mueller (Eds.), *A second course in structural equation modeling* (pp. 119-169). Information Age Publishing.

If you are evaluating change over time how will you do that? The simplest way is to look at change scores. This is valid if you have a pre- post- design, where you have 2 time points.

* Bonate, P. L. (2000). *Analysis of Pretest-Posttest Designs.* Chapman and Hall/CRC. <https://doi.org/10.1201/9781420035926>
* Hovland, C. I., Lumsdaine, A. A., & Sheffield, F. D. (1955). A baseline for measurement of percentage change. In P. F. Lazarsfeld & M. Rosenberg (Eds.), *The language of social research: A reader in the methodology of social research* (pp. 77-82). The Free Press.
* Zimmerman, D. W., & Williams, R. H. (1982). Gain scores in research can be highly reliable. *Journal of Educational Measurement*, *19*(2), 149-154. <https://doi.org/10.1111/j.1745-3984.1982.tb00124.x>

If you are interested in multiple time points there are multiple ways of analysing change over time.

* Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. Wiley-Interscience.
* Curran, P. J., & Bollen, K. A. (2001). The best of both worlds: Combining autoregressive and latent curve models. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 107-135). APA.
* Fitzmaurice, G. M., Laird, N. M., & Ware, J. W. (2004). *Applied longitudinal analysis*. Wiley.
* McArdle, J. J. (2007). Five steps in the structural factor analysis of longitudinal data. In R. Cudeck & R. C. MacCallum (Eds.), *Factor analysis at 100: Historical developments and future directions* (pp. 99-130). LEA.
* Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press.

Your data might come from individuals who are in groups, such as schools or classes. The influence of that situation can be understood using hierarchical or multi-level models.

* Goldstein, H. (2011). *Multilevel Statistical Models* (4th ed.). John Wiley & Sons, Ltd
* Hoyt, W. T., & Kenny, D. A. (2013). *Multilevel Modeling in R , Using the nlme Package.* Available: [http://davidakenny.net/papers/k&h/MLM\_R.pdf](http://davidakenny.net/papers/k%26h/MLM_R.pdf)
* Osborne, J. W. (2000). Advantages of hierarchical linear modeling. *Practical Assessment, Research & Evaluation*, *7*(1). <https://doi.org/10.7275/pmgn-zx89>
* Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage.
* Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage.