Longitudinal Planned Missing Designs in Education Research

Jessica Logan

Menglin Xu

The Ohio State University

Longitudinal Planned Missing Designs in Education Research

Introduction. Planned missing designs are relatively new to the field of psychological and developmental research. These designs allow researchers to randomly assign participants to missing data conditions; participants are either randomly selected to receive or *not* receive particular measures or items during data collection (Little & Rhemtulla, 2013). Conventionally, researchers would spare no effort to avert incomplete data due to the negative consequences along with it, e.g., smaller sample size and thus power, inaccurate estimation (Graham, 2009). However, when the participants have been randomly assigned to have that missing data, the data can be assumed to be missing completely at random (MCAR), because the missingness arises from a completely random process (Graham, 2012).

In this study, we provide a methodological application of one type of planned missingness design: the Two Method Planned Missingness TMM design. This design was specifically developed to address the tradeoff between statistical power and measurement error, and to reduce the costs and burden of data collection, while not sacrificing estimation accuracy (Graham & Shevock, 2012; Graham, Taylor, Olchowski, & Cumsille, 2006). These designs are applicable to research questions where the primary construct can be measured both by an expensive or "gold-standard" measure as well as by an inexpensive "efficient" measure. Gold standard measures should have excellent reliability, and are usually more costly in terms of financial expense, administration, and personnel training (e.g., classroom observations, standardized tests). Such measures typically take up more of a participant's time, often must be administered to participants individually by project staff, and typically have a high per-unit cost. Efficient measures are cheaper and much faster to administer, but are also typically less reliable or measured with inherent bias. In the TMM design, all participants receive the efficient measure, but only a randomly selected subset of participants receive the more expensive gold-standard measure. By selecting (randomly or purposefully) a subset of the sample to receive the expensive measures, costs to the project are minimized.

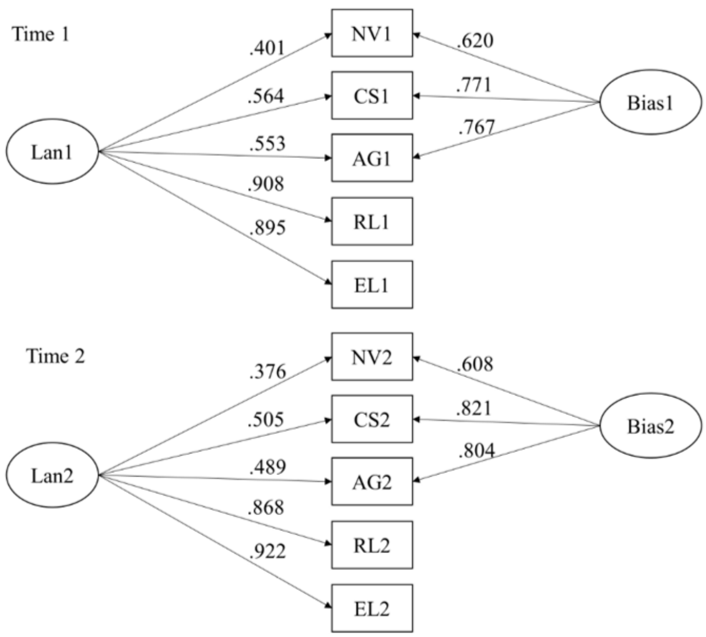
Educational Applications. Regarding applications in education and developmental research, researchers are not only interested in students' academic performance at a single time point, but also how children learn and grow over time. Above and beyond the dilemma faced at a single time point in terms of trade-offs between budget and data quality, longitudinal design adds additional burden in terms of cost and effort. Drawing from these concepts, Garnier-Villarreal, Rhemtulla, and Little developed a longitudinal extension of TMM design (2014). Just as with the original TMM, the gold standard measure is only administered to a subset of participants, but they also posit that it could potentially be administered at only a subset of occasions as well. To analyze the data, this model essentially fits the TMM design separately at each measurement occasion, then estimates relations between measurement occasions. *However, to date the longitudinal TMM design has yet to be tested with empirical educational data, which is the key goal of the present work.*

The present study. In summary, planned missingness designs provide researchers a method to measure a subset of people with expensive gold-standard assessments, while maximizing power and obtaining accurate estimation for a much larger total sample. Work by Garnier-Villarreal et al. (2014) proposed a longitudinal application of the TMM design, and demonstrate through simulation studies that it is an efficient and accurate tool for longitudinal research. However, to date, no empirical applications of the model have yet been published. The goal of the present study is to first, inform the planned missingness design literature by providing an empirical application of the model. Second, we provide an extension to the existing literature on planned missing designs to demonstrate how to examine causality and intervention response within the longitudinal planned missing data framework.

Sample. We provide an illustrative example to demonstrate how the model fitting procedure works with educationally relevant data. In our example dataset (Star2, Justice, 2017; Retrieved from ICPSR doi.org/10.3886/ICPSR36738.v1), we examine the language development of 683 preschool children across two time points using the Comprehensive Evaluation of Language Fundamentals–Preschool (CELF: P-2; Wiig, Secord, & Semel, 2004). The CELF includes both gold standard measures, which are completed by trained project staff, and efficient measures, completed by the parents of the participating children.

Results. The model fitting process was conducted in four basic steps. In Step 1, the TMM was fit to the data at each measurement occasion separately. A figure representing this stage of the model building process is included below. In Step 2, we used model comparisons to identify whether one or two bias factors were needed to most appropriately fit the data. In this step, the optimal longitudinal TMM model was determined to have a separate response bias factor at each measurement occasion. In Step 3, we established cross-time measurement invariance for both the bias factors and the construct factors. Finally, in Step 4 we added the predictors (covariates and treatment variable) to the structural equation model.

Conclusion. The primary goal of this work is to provide evidence that TMM designs are viable for use in developmental and educational contexts. Through our model fitting process, we demonstrate that these models can provide an excellent fit to such data. The second aim of this study was to provide an empirical example of how the TMM design could be used to examine treatment effects in a residualized change format. This model was able to do so without introducing model misfit, and with no estimation problems. Residualized change models are almost ubiquitously used by researchers who are evaluating efficacy and effectiveness trials, and use of a TMM design has the potential to greatly reduce the total cost of the project. In sum, the current study demonstrates the promise of the TMM design for longitudinal, causal educational research.



Figure： Two-Method Measurement Model for Time 1 and Time 2 with standardized estimates.
Note. Lan: language skill factor; Bias: response bias factor. NV: Nonverbal communication; CS: Conversational routines and skills; AG: Asking for, giving and responding to information; RL: receptive language standard score; EL: expressive language standard score. 1 or 2 indicates time points. All the loadings are significant at $p < .001$.