Search for efficient complete and planned missing data designs for analysis of change

Wei Wu[1] · Fan Jia[1] · Mijke Rhemtulla[2] · Todd D. Little[3]

Abstract The design of longitudinal data collection is an essential component of any study of change. A well-designed study will maximize the efficiency of statistical tests and minimize the cost of available resources (e.g., budget). Two families of designs have been used to collect longitudinal data: complete data (CD) and planned missing (PM) designs. This article proposes a systematic and flexible procedure named SEEDMC (SEarch for Efficient Designs using Monte Carlo Simulation) to search for efficient CD and PM designs for growth-curve modeling under budget constraints. This procedure allows researchers to identify efficient designs for multiple effects separately and simultaneously, and designs that are robust to MCAR attrition. SEEDMC is applied to identify efficient designs for key change parameters in linear and quadratic growth models. The identified efficient designs are summarized and the strengths and possible extensions of SEEDMC are discussed.

Keywords Planned missing data designs · Growth curve modeling · Efficiency · Longitudinal data collection

✉ Wei Wu
  wwei@ku.edu

[1] Department of Psychology, University of Kansas, 1415 Jayhawk Blvd., Lawrence, KS 66044, USA

[2] University of Amsterdam, Amsterdam, The Netherlands

[3] Institute for Measurement, Methodology, Analysis, and Policy, Texas Tech University, Lubbock, TX, USA

The design of longitudinal data collection is an essential component of any study of change. Factors that researchers must consider when planning such designs include but are not limited to the number and allocation of repeated measures, sample size, and minimizing unwanted factors such as carry-over effects and attrition. In addition to these considerations, researchers are typically constrained by a finite budget and limited resources which should also be taken into account at the design stage. A well-designed study will lead to statistical tests that are maximally efficient, that is, precise and powerful (Berger & Wong, 2009), given available resources.

Two families of designs can be used to collect longitudinal data: complete data (CD) and planned missing (PM) designs. Each of these designs represents a unique combination of design factors (e.g., number of repeated measures per participant and allocation of the repeated measures to time points), leading to the differential efficiency of target statistical procedures. Given two designs, the more efficient one produces parameter estimates that have smaller sampling variability, leading to smaller standard errors, tighter confidence intervals, and thus higher power to detect non-zero effects. Thus, more efficient designs are preferred.

Although past research has studied CD designs and specific PM designs, no study so far has searched for the most efficient design among all possible CD and PM designs. Furthermore, the effect of unplanned missingness (missingness that is not due to design, e.g., attrition) has been studied for CD designs but not for PM designs. Given that unplanned missingness is ubiquitous in longitudinal studies, this is an important issue to address. Identifying efficient designs is not a trivial matter. As shown in the examples below, using an efficient design may save up to 75 % of the research budget while achieving the same efficiency to detect target effects when compared to commonly used designs. In the current article, we propose a systematic procedure called SEEDMC to search for efficient

CD and PM designs with attrition for longitudinal growth curve designs under budget constraints.

The article is organized as follows. First, we briefly introduce growth-curve modeling with linear and quadratic change trajectories. We then explain CD and PM designs in detail and review the literature on these designs. Next, we describe the SEEDMC procedure and use it to identify efficient designs for linear and quadratic growth models. We conclude by discussing the strengths and limitations of SEEDMC and possible future extensions.

Growth-curve modeling (GCM)

GCM allows researchers to examine the average change trajectory of a trait, behavior, or ability over time, the individual variability in the change trend, and the predictors (e.g., treatment) that explain individual variability in the change trajectory (Wu, Selig, & Little 2012). Most commonly, GCM is applied to longitudinal data with a limited and fixed number of measurement occasions (e.g., 5-yearly measures of school achievement).

In GCM, the shape of a change curve is determined by so-called growth coefficients. If the curve is linear, the growth coefficients are an intercept and linear slope (constant rate of change over time). If the curve is quadratic, the growth coefficients are an intercept, linear slope (instantaneous rate of change at some time point), and quadratic rate (the change rate of the linear slope over time). In GCM, growth coefficients are random as their values can vary across individuals.

GCM can be implemented as either mixed/multilevel models or as structural equation models (SEMs) (Chou, Bentler, & Pentz, 1998). In the SEM framework, the growth coefficients are treated as latent factors which predict the repeated measures (Bollen & Curran, 2006; Meredith & Tisak, 1990). Figure 1 displays a linear GCM of 5-yearly observations of a given
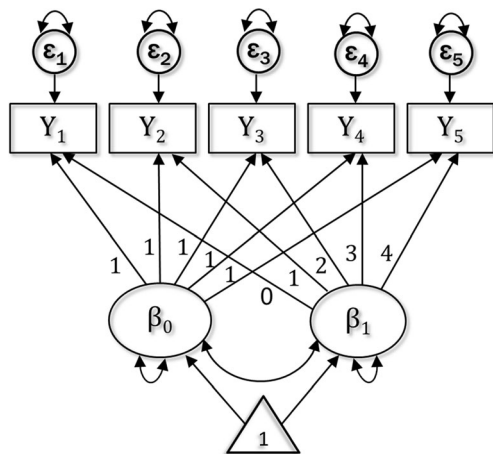
outcome (i.e., $Y_1 - Y_5$). The factor loadings associated with the latent slope factor are 0, 1, 2, 3, and 4, indicating the time elapsed since the beginning of the study. The intercept factor (β_0) in this case represents the subjects' initial status on Y (e.g., children's achievement in the first year of school) and the slope factor (β_1) represents the change in Y per year (e.g., children's linear growth in achievement over 5 years). $\varepsilon_1 - \varepsilon_5$ are residuals which capture the deviations of the observed data from the predicted individual curves, and they are typically assumed to be independent and identically normally distributed.

To extend the linear model to fit a quadratic curve, one more latent factor representing the quadratic term must be added to predict the repeated measures. The factor loadings associated with the quadratic term are squared time scores (i.e., 0, 1, 4, 9, and 16). In either the linear or quadratic model, the means of the growth coefficients (e.g., mean intercept and slope for a linear model) and the covariance matrix of the growth coefficients are typically estimated along with the residual variance.[1] Predictors of the growth coefficients can also be incorporated into the model by adding pathways from the predictors to the latent factors. Means of the growth coefficients and effects of predictors are often called fixed-effect parameters because they are constant across individuals. In contrast, the elements in the covariance matrix of the growth coefficients are called random-effect parameters as they capture individual variability on the growth coefficients. For ease of reference, we refer to the models with and without predictors of the growth coefficients as conditional and unconditional models, respectively.

Complete and planned missing designs

Having introduced the analysis method for change, we now turn to the CD and PM designs for longitudinal data collection. We define CD designs as those in which all participants participate at the same time points (i.e., designs with only one planned response pattern). For example, if a researcher is interested in the rate of change over 5 years and the outcome can only be measured once per year, a commonly used CD design is to collect data from all participants in each of the 5 years. This is not the only possible CD design, however. Alternate CD designs for the same research question could collect data from all participants in year 1, year 2, and year 5, or just year 1 and year 5. Given a fixed budget, these alternate designs could compensate for fewer measurement occasions by using a larger sample size. It would be interesting to know which CD design is most efficient for a specific research problem.

Fig. 1 A linear growth curve model of five repeated measures. *Note.* $Y_1 - Y_5$ are repeated measures, β_0 and β_1 are latent intercept and slope, respectively, and $\varepsilon_1 - \varepsilon_5$ are residuals. The paths from the triangle to β_0 and β_1 indicate the mean intercept and slope, respectively

---

[1] Depending on how the model is specified, there can be more or fewer parameters. For example, there will be fewer parameters if some of the parameters (e.g., slope variance) are constrained to be 0 and more parameters if the residuals have heterogeneous variance over time, autocorrelations, or both.

Since different CD designs have a different number and allocation of repeated measures, a systematic search for efficient CD designs can address two important design questions: How many measurement occasions should be included, and how should these be allocated across the time span of the study to achieve maximum efficiency?

In contrast to CD designs, PM designs divide individuals into groups, each of which is randomly assigned to a subset of the measurement occasions (Graham, Taylor, & Cumsille, 2001; Graham, Taylor, Olchowski, & Cumsille, 2006; Little & Rhemtulla, 2013; Rhemtulla, Jia, Wu, & Little, 2014). The resulting dataset has more than one response pattern and contains missing data due to design. For instance, a possible PM design for the 5-year example above includes three groups of participants, one measured at years 1, 2, and 5, another measured at years 1, 3, and 5, and the last measured at years 1, 4, and 5 (see Table 1). In this design, each participant has missing data on two of the five measurement occasions. Consequently, a missing data method has to be used to analyze the data.

Depending on the number of patterns, number of repeated measures in each pattern, and allocation of repeated measures in each pattern, a large number of PM designs are possible. It would be interesting to know which PM design is most efficient for a specific research problem as well as how these PM designs perform compared to CD designs for the same problem. A systematic search for efficient designs among both types of designs can address not only the two design questions mentioned above, but also the additional questions: How many response patterns should be used, and which measurement occasions should be included in each response pattern to achieve maximum efficiency?

CD designs are more widely used and easier to implement than PM designs. PM designs, however, have several potential advantages (Graham, Taylor, & Cumsille, 2001; Mistler & Enders, 2012). First, given a fixed budget, certain PM designs can show higher efficiency than CD designs that include the same number of total observed data points, because they can offer a better balance between the number of repeated measures per participant/pattern and sample size[2] (Galbraith, Stat, & Marschner, 2002; Mistler & Enders, 2012). Second, given a fixed sample size, PM designs require fewer repeated measures from each individual to achieve the same number of measurement occasions in the data as a CD design, at substantially reduced cost (Hogue, Pornprasertmanit, Fry, Rhemtulla, & Little, 2013). Third, because each participant receives fewer repeated measures, carry-over effects caused by repeatedly undergoing the same measurement procedure can be reduced (Jorgensen, Rhemtulla, Schoemann, McPherson, Wu, & Little, 2014). Finally, attrition rates may be lower when

Table 1 Example of a planned missing (PM) design

Pattern	Measurement occasion				
	1	2	3	4	5
1	1	1	0	0	1
2	1	0	1	0	1
3	1	0	0	1	1

Note. 1 = observed, 0 = missing

participants are not required to participate in a large number of measurement occasions (Harel, Stratton, & Aseltine, 2012). Note that the missing data in a PM design are missing completely at random (MCAR), which can be correctly handled by modern missing data treatments such as full information maximum likelihood (FIML) and multiple imputation (MI) (Enders, 2010; Rubin, 1987). The merits of these designs should be considered along with their efficiency.

Literature review

The majority of research on CD designs for the analysis of change is found in the optimal design literature in biomedical research. These studies typically use an efficiency criterion called the D-optimal criterion, according to which the design that minimizes the generalized variance of the target effects/parameters is most efficient (Ouwens, Tan, & Berger, 2002). The generalized variance is the determinant of the asymptotic covariance matrix of the target coefficients. In the following, we summarize the major findings from this literature.

First, given a fixed budget, it is not beneficial to have too many repeated measures as these will contain redundant information (Ouwens et al., 2002; Tekle, Tan, & Berger, 2011). For polynomial change trajectories (linear and quadratic trajectories are special cases), if the shape of the change trajectory is well understood and the goal is to achieve maximum efficiency to detect fixed-effect parameters (e.g., mean intercept and slope in a linear model), then the optimal number of repeated measures in a CD design should be as close as possible to the number of growth parameters (e.g., two or three for linear curves,[3] three or four for quadratic curves; Moerbeek, 2005; Tekle, Tan, & Berger, 2011; Willett, Singer, & Martin, 1998).

Second, the optimal allocation of repeated measures in a CD design is approximately equally spaced (Berger & Wong, 2009; Ouwens et al., 2002; Tekle et al., 2011). Although non-

---

[2] Between the two factors, holding one factor constant and increasing the other will usually increase the efficiency of a study. However, with a fixed budget, increasing one factor will cause a decrease of the other. In this case, finding the right balance between the two factors is the key to efficient designs.

[3] The optimal number of repeated measures was identified assuming that the model is identified with the number of repeated measures. For example, in a CD design with two repeated measures, the linear model cannot have more than five parameters. Such a model can be a GCM with only a random intercept or a GCM with autocorrelations among the repeated measures but not random effects on the intercept and slope.

equally spaced designs might turn out to be more efficient than equally spaced designs in some cases, the difference in efficiency between equally and nonequally spaced designs is generally small. In addition, equally spaced designs tend to be more practical and theoretically relevant than unequally spaced repeated measures (Berger & Wong, 2009; Galbraith & Marschner, 2002).

Third, the efficiency of optimal CD designs will be influenced by unplanned missingness incurred during data collection. Galbraith and Marschner (2002) found that for linear models, the loss of efficiency due to attrition was within 10 % if the attrition rate was no more than 30 %. Ortega-Azurbuy, Tan, and Berger (2008) found that the efficiency loss of D-optimal designs identified using complete data was within 15 % with a 70 % attrition rate at the last time point.

Only a few studies in the optimal design literature have considered PM designs (referred to as cohort designs in these studies; Gan, 2011; Ouwens et al., 2002; Tekle, Tan, & Berger, 2008, 2011). These studies are restricted to very specific types of PM designs. For example, Tekle et al. (2008, 2011) considered only non-overlapping missing data patterns (i.e., the missing data patterns share no common measurement occasion) and concluded that PM designs are less optimal than CD designs. Gan (2011) and Ouwens et al. (2002) examined PM designs with only two missing data patterns, both of which were measured at the first and last measurement occasions (each pattern captured a different intermediate measurement occasion). Ouwens et al. (2002) found that the efficiency of PM designs with three repeated measures in each of two groups was almost as high as that of CD designs with four repeated measures in detecting linear or quadratic trends under the same budget.

Outside of the optimal design literature, a few studies have paid special attention to more complex PM designs (Graham, Taylor, & Cumsille, 2001; Mistler & Enders, 2012). These studies compared several PM designs to the corresponding CD designs that included the same number of total measurement occasions. For example, Graham et al. (2001) investigated the efficiency of various PM designs in a longitudinal study with five measurement occasions. They considered designs with all combinations of response patterns with two, three, or four observations per pattern. They found that PM designs resulted in higher efficiency than the corresponding CD design to detect the effect of a binary predictor on the linear slope in a quadratic model. Graham et al. quantified efficiency using the asymptotic standard error of the regression coefficient. Table 2 shows the most efficient design in their study, which contains seven missing data patterns, including one pattern with complete data at every occasion and six patterns with complete data on three occasions each. This design assigned higher weight (i.e., more participants) to the groups measured at both the first and last measurement occasions.

Table 2 Design 3c in Graham et al. (2001)

Pattern	Measurement occasion					Percentage of sample
	1	2	3	4	5	
1	1	1	1	1	1	9.1 %
2	1	1	1	0	0	10.1 %
3	1	1	0	1	0	10.1 %
4	1	0	1	1	0	10.1 %
5	1	1	0	0	1	20.2 %
6	1	0	1	0	1	20.2 %
7	1	0	0	1	1	20.2 %

Note. 1= observed, 0 = missing

A follow-up study by Mistler and Enders (2012) examined PM designs for linear and quadratic growth models with six measurement occasions. They used power (the proportion of coefficients whose confidence intervals excluded zero) from Monte Carlo simulation as a proxy for efficiency, with higher power indicating higher efficiency. They compared three designs: design 1 was a CD design with six measurement occasions, design 2 was a PM design containing all possible combinations of 4 measurement occasions, and design 3 contained only those combinations of four measurement occasions that included the first and last measurement occasions (see Table 3). They found that design 3 led to the most efficient estimates of linear and quadratic growth trends given a fixed totally number of observed data points.

Limitations of previous research

Research on CD designs in the optimal design literature has focused on multiple fixed effects simultaneously (e.g., both intercept and slope). Thus, it is not clear whether the design that is most efficient for multiple effects is most efficient for any given effect. In addition, this research has only considered the efficiency of a design to detect the shape of a change trajectory. None have examined the predictor effects (e.g.,

Table 3 Design 3 in Mistler and Enders (2012)

Pattern	Measurement occasion						Percentage of sample
	1	2	3	4	5	6	
1	1	1	1	0	0	1	16.7 %
2	1	1	0	1	0	1	16.7 %
3	1	0	1	1	0	1	16.7 %
4	1	1	0	0	1	1	16.7 %
5	1	0	1	0	1	1	16.7 %
6	1	0	0	1	1	1	16.7 %

Note. 1= observed, 0 = missing. See Table 41.4 in Mistler and Enders (2012)

treatment) on change trajectories, which are often key effects in longitudinal studies.

The research on PM designs also has several limitations. First, many studies on PM designs compared PM designs to only a single CD design: the design with data on all measurement occasions, which is not necessarily the most efficient CD design for the target analysis. As a result, it is not clear how these PM designs would perform compared to more efficient CD designs (i.e., CD designs with fewer repeated measures). Second, these studies have not accounted for attrition or additional unplanned missingness in PM designs. Thus, it is not clear whether the designs that are efficient under planned missingness would continue to be efficient when unplanned missingness occurs. Third, although some specific PM designs are recommended in these studies, a systematic procedure to search for efficient PM designs is lacking.

In response to these limitations in the previous literature, we propose a systematic procedure to search for efficient designs for growth-curve modeling within a large pool of CD and PM designs using Monte Carlo Simulation. We call this procedure SEEDMC (SEarch for Efficient Designs using Monte Carlo Simulation). Although Monte Carlo simulation is not a new technique, this is the first attempt to develop a comprehensive search process that incorporates both CD and PM designs with attrition taken into account. More importantly, due to the flexibility of Monte Carlo simulation, this procedure can be readily extended to any change trajectory (linear or nonlinear), type of missing data whether planned or unplanned, type of variables (continuous or categorical), and estimation method (e.g., maximum likelihood-based estimation methods or weighted least square-based estimation methods; Hancock & French, 2013; Muthén & Muthén, 2002). This flexibility will be revisited in the discussion section.

SEEDMC (SEarch for Efficient Designs using Monte Carlo Simulation)

Three steps are involved in the search procedure. The first is to decide the maximum number of measurement occasions (MNM). "Maximum" is used here to indicate the fact that the total number of measurement occasions in a design may not necessarily include all possible measurement occasions (e.g., a CD design could include three, four, or five measurement occasions when MNM = 5). The MNM is defined based on a reasonable and practical unit of time (e.g., month, half a year, or year) and a desirable length of study. Second, based on the MNM, a design pool that contains all possible CD and PM designs is defined. Third, Monte Carlo simulation is

used to evaluate the efficiency of each parameter estimate for each design in the design pool, and the most efficient designs are identified for the target estimate(s). The second and third steps are explained further below.

In the second step, to reduce the size of the design pool to a manageable level, it is reasonable to impose certain constraints on the design pool. For example, to make sure that all CD designs share the same length of study, we restrict the CD designs to ones with observations at the first and last measurement occasions. In addition, a minimum number of repeated measures should be specified for CD designs so that the model under consideration can be identified. For example, a typical linear GCM requires at least three and a typical quadratic GCM requires at least four repeated measures to be identified.

We restrict our search for PM designs to ones that satisfy the following conditions: (1) the number of repeated measures in the sample is equal to the MNM but the number of repeated measures per pattern is less than the MNM (i.e., no pattern has complete data at all measurement occasions), (2) the response patterns in each design include the same number of repeated measures (ranges from 2 to MNM - 1), and (3) the proportion of participants assigned to each pattern is uniform, that is, the missing data patterns are equally weighted. We imposed these restrictions in order to keep the number of PM designs under consideration to a manageable number, though it is possible that the best PM design is excluded by these constraints.

The following notation is used for each design. Each response pattern is given as a list of the measurement occasions included in that pattern; multiple response patterns are separated by a bar, and the set of response patterns for a design is enclosed in brackets. For instance, the design shown in Table 3 is presented as {1,2,3,6 | 1,2,4,6 | 1,3,4,6 | 1,2,5,6 | 1,3,5,6 | 1,4,5,6}.

Having defined the design pool, SEEDMC then uses the Monte Carlo method to generate a large number of samples/ replications (e.g., 5,000) for each design based on the hypothesized model, the response pattern(s) in the design, and the pre-specified rates of attrition imposed at each of the measurement occasions (details about how attrition is imposed are provided in the first example).

Given a fixed budget, N of the simulated data varies across designs. Specifically, the N for a design is a function of the total budget (B), the cost of collecting each data point (C), and the number of repeated measures per pattern (T),

$$N = B\big/(T \times C). \tag{1}$$

Note that this function assumes that every data point is equally expensive, which may be an unreasonable assumption (i.e., collecting the first data point from a new participant may

be much more expensive than collecting a subsequent data point from someone who is already enrolled in the study). This possibility is discussed further in the Discussion. After the samples are generated, the hypothesized model is fitted to each sample. The analysis results are used to evaluate the efficiency of each design for each target parameter.

Relative efficiency

We measure the efficiency of an estimate in a design by the inverse of the sampling variance of the estimate resulting from the design. Sampling variance is inversely proportional to the amount of information available for estimating the parameter. Thus a design that minimizes the sampling variance would also maximize the amount of information available for estimating the parameter, leading to maximized efficiency. A lower sampling variance will also correspond to a smaller standard error, a larger test statistic value, leading to higher power to detect non-zero parameters. This measure of efficiency is consistent with the D-optimal criterion adopted in the optimal design literature (with the difference that the D-optimal criterion summarizes efficiency over all parameter estimates rather than considering them singly) and the criterion used in Graham et al. (2001). Note that in Monte Carlo simulation, we obtain the empirical sampling variance (i.e., the observed variance of parameter estimates over repeated samples) instead of the asymptotic variance of a parameter estimate. Compared to the asymptotic variance, the empirical variance better captures the sampling variability of the parameter estimate under finite sample sizes and missing data. Let θ be a given parameter and $\mathrm{var}\left(\hat{\theta}\right)$ be the empirical sampling variance of an estimate of θ ($\hat{\theta}$). $\mathrm{var}\left(\hat{\theta}\right)$ is computed as follows:

$$\mathrm{var}\left(\hat{\theta}\right) = \frac{\sum_{i=1}^{p}\left(\hat{\theta}_i - \overline{\theta}\right)^2}{p}, \tag{2}$$

where $\hat{\theta}_i$ is the parameter estimate from the ith replication, $\overline{\theta}$ is the average parameter estimate across replications, and p is the number of replications.

To compare the efficiency of multiple designs, we hold constant the total number of observed data points associated with the design, as a proxy for the cost of a design. To facilitate comparison of efficiency among designs, the relative efficiency of each design can be estimated for each parameter. The relative efficiency (RE) of a design is the efficiency of a design relative to the most efficient design. Let design 0 be the most efficient design (i.e., the design with the smallest sampling

variance), and design 1 be any other comparison design, the relative efficiency of design 1 is computed as follows.

$$\widehat{RE} = \frac{efficiency_{design1}}{efficiency_{design0}} = \frac{1/\mathrm{var}\left(\hat{\theta}\right)_{design1}}{1/\mathrm{var}\left(\theta\right)_{design0}}$$

$$= \frac{\mathrm{var}\left(\hat{\theta}\right)_{design0}}{\mathrm{var}\left(\theta\right)_{design1}}, \tag{3}$$

where $\widehat{RE}$ represents the estimated relative efficiency of design 1. $\widehat{RE}$ ranges from 0 to 1 with higher values indicating a more efficient design. For example, if the $\widehat{RE}$ is .80, then the sampling variance from design 0 is 80 % as large as the variance from design 1, and design 1 can be said to be 80 % efficient relative to design 0. $\widehat{RE}$ can be used to compute the additional number of data points (a proxy for cost, if the cost of collecting each data point is equal) necessary for a design to achieve the same efficiency as the most efficient design, using the formula $(1/RE - 1)\times100$ %. For instance, if the RE for a design is .80, then $(1/.80 - 1)\times100$ % = 25 % more data points (or a 25 % higher budget) are needed for the design to achieve the same efficiency as the most efficient design.

Due to the infinite number of replications used in the Monte Carol simulation, $\widehat{RE}$ is subject to random simulation error. In other words, $\widehat{RE}$ can change if one runs the search process for more than one time. The simulation error decreases as number of simulated samples/replications increases. We recommend at least 5,000 replications to minimize the simulation error.

Once the relative efficiency of each design is obtained for each parameter, one can start selecting the efficient designs. In this study, instead of focusing only on the design with the largest $\widehat{RE}$, we consider all designs with $\widehat{RE} \geq .90$ as efficient designs. There are two benefits of doing this. First, this will offer more design options, some of which might be preferable to the others due to practical concerns. Second, considering more than one efficient design for each effect can also facilitate the search for designs that are simultaneously efficient for multiple key effects (e.g., all fixed effects related to a slope factor). In this case, a design that is simultaneously efficient across multiple effects (e.g., has $\widehat{RE} \geq .90$ for all of the effects) is preferred to a design that is maximally efficient for one specific effect but may be inefficient for other key effects. Note that the threshold for efficient designs (.90 is used here) may be adjusted according to the extent to which a loss of efficiency can be tolerated.

In the next section, we use two examples to demonstrate how the SEEDMC procedure identifies efficient designs for key effects in linear and quadratic GCMs, respectively. In both examples, MNM = 5 which is a commonly seen number of

measurement occasions in practice. In addition, to examine the effect of attrition on the relative efficiency of a design and to identify efficient designs that are robust to attrition, we imposed three levels of unplanned attrition (none, low, and high).

Example 1: Linear GCM

In the first example, we applied SEEDMC to a linear growth curve model described in Biesanz, Deeb-Sossa, Papadakis, Bollen, & Curran (2004). This model had five measurement occasions (i.e., MNM = 5), resulting in seven possible CD designs and 48 PM designs that satisfy the restrictions described above. Details of how the designs are identified are given in Appendix 1. For ease of reference, we assign a label to each design consisting of a letter to indicate whether it is a complete data (C) or missing data (M) design and a number to indicate the order of the design in the design pool. The order is arbitrary and is not related to the efficiency of a design.

The SEEDMC algorithm simulated data from a linear growth model with $p = 5000$ replications for each design under each level of attrition. The population values of the parameters were taken from Biesanz et al. (2004). In the linear model, the latent intercept and slope had means of 39.46, and 8.06 respectively, variances of 28.78, and 8.20, respectively, and a covariance of 1.56. The residual variances were constrained to be equal to 30 at each measurement occasion. We adjusted the residual variances and the covariance between the linear slope and the intercept as the values reported in Biesanz et al. (2004) led to convergence problems. To search for efficient designs for predictor effects, we added a time-invariant predictor, Z, which had a .30 correlation with the intercept and the linear slope.

As mentioned above, we imposed three levels of unplanned attrition (none, low, or high). Following Ortega-Azurduy, Tan, and Berger (2008), for low or high attrition, the probability of attrition at each measurement occasion was a linear function of time.[4] This probability increased as time increased which is consistent with the common trend of attrition in longitudinal studies. Since the probability of attrition is only determined by time, which is treated as a fixed variable in the model, the missing data mechanism is MCAR.

$$\text{Low attrition}: \quad p\left(y_{ij} \text{ is missing}\right) = .075 \times time_j \qquad (4)$$

$$\text{High attrition}: \quad p\left(y_{ij} \text{ is missing}\right) = .175 \times time_j \qquad (5)$$

---

[4] Ortega-Azurduy, et al. (2008) also examined a scenario where the probability of attrition increased as a quadratic function of time. However, they did not find any meaningful difference in the result between the linear and quadratic functions.

where y_{ij} represents the outcome measure for the ith participant at the jth measurement occasion, and $time_j$ represents the elapsed time since the first measurement at the jth measurement occasion.

For CD designs, all participants enter the study at the first occasion. Thus, they all share the same probability of attrition at each occasion. For low attrition, the probabilities are 0 %, 7.5 %, 15 %, 22.5 %, and 30 % at the five measurement occasions, respectively (see Eq. 4). For high attrition, the probabilities are 0 %, 17.5 %, 35 %, 52.5 %, and 70 %, respectively (see Eq. 5). As indicated by a review conducted by Gustavson, von Soest, Karevold, and Røysamb (2012), attrition rates ranging from 30 % to 70 % are not uncommon in longitudinal studies.

For PM designs, however, the participants enter the study at different times depending on their designated response patterns. As a result, the probability of attrition at a specific time may also differ across participants. For instance, if a participant enters a study at the second measurement occasion and is followed at the fourth and fifth measurement occasions, then the time scores used to compute the probability of attrition will be 0, 2, and 3 at the three occasions, respectively. Based on Eq. 5, this will result in a probability of attrition of 0 %, 15 %, and 22.5 %. Compared to 7.5 %, 22.5 %, and 30 % at the same measurement occasions in a CD design, these probabilities were much lower. This way of computing probability of attrition reflects the advantage of PM designs as a tool to reduce unplanned attrition.

We arbitrarily set the cost of collecting each data point (C) to US$20 and the total budget to US$100,000, so that the sample size for the CD design with five repeated measures is 1,000. Note that although cost affects efficiency, we don't expect that the relative efficiency of a design given a fixed cost would change much across different levels of cost. In this sense, one can plug in any assumed cost values and obtain similar results. However, we encourage researchers to provide realistic values for the total cost and the cost of each data point so that the Monte Carlo simulation will be carried out based on a realistic N for each design. In the case where the N is too small, having additional attrition may cause convergence problems in some of the designs.

The sample size for each of the other designs is determined by Eq. 1. For example, for a PM design with three repeated measures per pattern ($T = 3$), given $B =$ US$100,000, and $C =$ US$20, $N = 100,000/(3×20) = 1,667$. By obtaining two fewer observations from each participant, this design allows researchers to collect data from 66.7 % more participants than the corresponding CD design under the same budget.

After generating the data, two models are fit to each data set. One is an unconditional linear model (without the predictor Z) and the other is a conditional linear model (with the predictor Z). The former is used to evaluate the designs with respect to the means and variances of the

growth factors. The latter is used to evaluate the designs for the predictor effect on the growth factors.

Example 1: Results

We focus on efficient designs (i.e., designs with $\widehat{RE} \geq .90$) for the fixed effects related to the slope (i.e, the mean slope and the effect of Z on the slope) as they are often more interesting to researchers than the fixed effects related to the intercept (i.e., the mean intercept and the effect of Z on the intercept) or the random effects of the intercept or slope. Efficient designs for these target effects are presented in Table 4. The results for the other parameters are included in the online supplementary material. Generally speaking, the designs that were efficient for the fixed effects related to the intercept are those with complete data at the first measurement occasion. This result makes intuitive sense, because the intercept represents the initial status of the outcome variable in the hypothesized model. The designs efficient for the random effects usually required at least one more repeated measure per pattern than the efficient designs for the fixed effects because the estimation of random effects requires more information than the estimation of fixed effects.

As can be seen in Table 4, there are multiple efficient designs for each target effect. For example, with no attrition, there are six efficient designs for the predictor effect on the linear slope. Given the space limit, we choose not to explain every efficient design but focus on the common features of the designs. Briefly, these designs are either PM designs with two or three repeated measures per pattern or CD designs with three repeated measures, suggesting that when the true change trajectory is linear, one does not need more than three repeated measures per participant to detect a linear trend or a predictor effect on a linear trend. These PM designs allow researchers to collect one fewer repeated measure per participant without substantially sacrificing efficiency. In addition, almost all designs include complete data (data on all participants) at the last occasion.

Based on Table 4, researchers can select the design that matches best with their research purposes, practical concerns, and expected attrition levels. For instance, if the research question centers around the predictor effect on the linear slope and unplanned attrition is not expected, then M18 = {1,5 | 2,5 | 3,5 | 4,5} ($\widehat{RE}$ = 1) is the most efficient design which contains all combinations of response patterns with two repeated measures per pattern and complete data at the last measurement occasion. C1 = {1,2,5} ($\widehat{RE}$ = .96) and C3 = {1,4,5} ($\widehat{RE}$ = .95) are also good choices in this case if a CD design is preferred.

In many cases, a researcher may prefer a design that is efficient for the mean linear slope and the predictor effect on the linear slope if they are both theoretically important, or a design that is efficient across all attrition levels for a key effect if prior knowledge on the expected attrition rates is not available. These designs can be found in the last row and last column of Table 4 respectively. For instance, if the efficient designs for both fixed effects is desired and the expected attrition level is low, then M18 = {1,5 | 2,5 | 3,5 | 4,5 }, C3 = {1, 4,5}, and M32 = {1,2,5 | 1,3,5 | 1,4,5} are all efficient designs (see the last row of Table 4). If the mean linear slope is of interest and the expected attrition rates are unknown, then M18 is the best choice as it is the only efficient design for the mean linear slope across all levels of attrition. In fact, for the linear model, M18 would be the best choice to detect either or both of the fixed effects related to the slope regardless of the attrition level.

The influence of attrition on $\widehat{RE}$ seemed to be dependent on the type of design and the attrition level. M18, the most efficient PM design for both effects was quite robust to MCAR attrition. However, the $\widehat{RE}$ of the other efficient PM designs (e.g., M32) tended to be decreased by high attrition but not by low attrition. For CD designs, attrition seemed to affect the allocation of the assessments but not the number of repeated measures per pattern. Specifically, given the attrition rate increasing over time, attrition tended to shift the data collection points toward the occasion with a higher probability of

Table 4 Efficient designs for the fixed effects related to the slope in the linear model

Effect	Attrition			
	None	Low	High	All levels
Mean linear slope	M18 (1.00) C3 (.90), C1 (.90)	M18 (1.00), C3 (.96) M32 (.90)	M18 (1.00)	M18
Z on linear slope	M18 (1.00), C1 (.96), C3 (.95) M32 (.94), C2 (.93), M14 (.90)	M18 (1.00), C3 (.97) M32 (.93), C2 (.90)	M18 (1.00)	M18
Both fixed effects	M18, C3, C1	M18, C3, M32	M18	M18

Note. The values in the parentheses are $\widehat{RE}$ s

C1 = {1, 2, 5}, C2 = {1, 3, 5}, C3 = {1, 4, 5}, M14 = {1, 2| 1, 3| 1, 4| 1, 5},

M18 = {1, 5| 2, 5| 3, 5| 4, 5}, M32 = {1, 2, 5| 1, 3, 5| 1, 4, 5}

attrition. For example, C3 = {1,4,5} was preferred to C1 = {1, 2,5} with low attrition although their efficiencies were comparable when there was no attrition.

It is useful to show how the designs considered in the previous studies perform. Here we reported the $\widehat{RE}$ s for three designs: (1) the CD design with five repeated measures (i.e., C7 = {1,2,3,4,5}). This is a commonly used design in practice and the design to which the past research on PM designs are compared, (2) Design 3c in Graham et al., (2001; see Table 2) (note that this design was not in the defined design pool because it has unequal assignment of participants to different patterns and combines patterns with different numbers of repeated measures; if this design is more efficient than the designs in our design pool, then we might need to revise the restrictions we imposed on the design pool to accommodate such designs), and (3) a design comparable to Design 3 in Mistler and Enders (2012; see Table 3). Design 3 in Mistler and Enders (2012) has six measurement occasions. A comparable design with five measurement occasions that shares the same feature is M22 = {1,2,3,5 | 1,2,4,5 | 1,3,4,5}.

It turns out that the $\widehat{RE}$ of the CD design ranged from .57 to .66, the $\widehat{RE}$ of Graham et al.'s design ranged from .71 to .82, and the $\widehat{RE}$ of Mistler and Enders' design ranged from .69 to .75 for the two fixed effects across the levels of attrition, compared to the corresponding design with the smallest empirical variance for each fix effect. These $\widehat{RE}$ s were low and none of the designs outperformed the efficient designs identified using our procedure. These results indicate that using the efficient designs identified by SEEDMC can save up to 75 % of the research budget [(1/0.57 − 1)×100 %] while achieving the same level of efficiency for target effects, compared to the commonly used or recommended designs in the previous literature.

Example 2: Quadratic GCM

Example 2 aims to search for efficient designs for a quadratic GCM. The same steps in Example 1 were applied here.[5] The only difference lies in the third step in which we simulated data from a quadratic model instead of a linear model. The population parameter values of the quadratic model were selected to match those of a model reported in Singer and Willett (2003; model C on page 221). Time was coded as 0, 1, 2, 3, and 4. The intercept, linear slope, and quadratic term had

---

[5] For convenience, the same design pool is used in both studies. This design pool contains the CD designs with three repeated measures which are not identified for a quadratic model. Including these designs does not affect the search results.

means of 13.97, −1.15, and 0.2, respectively, and variances of 107.08, 24.60, and 1.22, respectively. The mean quadratic curve is shown in Fig. 2. The covariance was −3.69 between the linear slope and the intercept, −1.36 between the quadratic slope and the intercept, and −4.96 the covariance between the quadratic slope and the linear slope. The residual variances were constrained to be equal to 41.98 across time.

Similarly, we included a predictor Z and fit both unconditional and conditional versions of the quadratic models. The predictor Z had a .30 correlation with the intercept, linear slope, and quadratic curvature. In the results report, we focus on describing the efficient designs for four fixed effects: the fixed effects related to the linear slope (i.e., the mean linear slope and the Z effect on the linear slope) and the quadratic curvature (i.e., the mean quadratic curvature and the Z effect on the quadratic term). The results for the other parameters can be found in the online supplementary material.

Example 2: Results

Table 5 lists the efficient designs for each of the four fixed effects under the three levels of attrition.

Again, multiple efficient designs were identified for each effect. These designs are either PM designs with three or four repeated measures per pattern or CD designs with four repeated measures, indicating that researchers generally need one more repeated measures per participant to detect a fixed effect in a quadratic model than in a linear model.

Based on the designs, researchers can select the one that best fits their needs. For example, if a researcher is most interested in detecting the mean quadratic term and the expected attrition level was low, then C4 = {1, 2, 3, 5} ($\widehat{RE}$ = 1) is the most efficient design and M43 = {1,2,3,5 | 1,3,4,5} ($\widehat{RE}$ = .94) and C6 = {1, 3, 4, 5} ($\widehat{RE}$ = .92) are good alternatives. If a researcher wants a design that is efficient for both fixed effects related to the quadratic term and the expected attrition level is high, then C4 = {1,2,3,5} ($\widehat{RE}$ s ≥ .99 for both effects) and M37 = {1,3,5 | 2,3,5 | 3,4,5} ($\widehat{RE}$ s ≥ .99 for both effects)
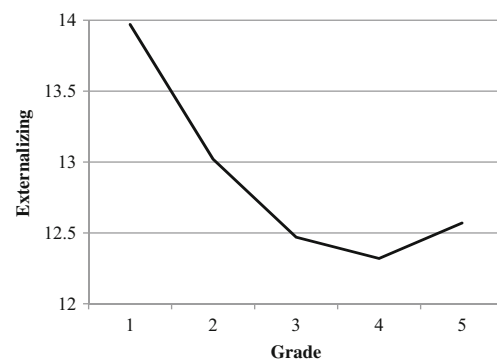

Fig. 2 The mean quadratic curve in Example 2

Table 5 Efficient designs for the target fixed effects in the quadratic model

Effect	Attrition			
	None	Low	High	All levels
Mean linear slope	M30 (1.00), C6 (.97), M43 (.93), M20 (.90)	M30 (1.00), C6 (.99), C4 (.96), M43 (.96), M20 (.92), M39 (.91)	C4 (1.00), M39 (.91), M43 (.91), M29 (.90), M30 (.90)	M30, M43
Mean quadratic term	M43 (1.00), C4 (1.00), C6 (.99), M22 (.92)	C4 (1.00), M43 (.94), C6 (.92)	C4 (1.00), M37 (.99), M11 (.93)	C4
Z on linear slope	M30 (1.00), C6 (1.00), M43 (.95), C4 (.91), M20 (.91)	C6 (1.00), M30 (.97), M43 (.96), C4 (.96), M20 (.92), M39 (.90)	C4 (1.00), M43 (.92), M39 (.91)	M43, C4
Z on quadratic term	M43 (1.00), C4 (.99), C6 (.99), M22 (.92)	C4 (1.00), M43 (.94), C6 (.93)	M37 (1.00), C4 (.99), M11(.92), M27 (.90), M46 (.90)	C4
All fixed effects	C6, M43	C6, C4, M43	C4	NA

Note. The values in the parentheses are $\widehat{RE}$ s

C4 = {1, 2, 3, 5}, C6 = {1, 3, 4, 5}, M11 = {1, 2, 3| 1, 3, 4| 1, 3, 5| 2, 3, 5| 3, 4, 5},

M20 = {1, 2, 3, 4| 1, 2, 3, 5| 1, 3, 4, 5}, M22 = {1, 2, 3, 5| 1, 2, 4, 5| 1, 3, 4, 5},

M27 = {1, 2, 3, 5| 1, 3, 4, 5| 2, 3, 4, 5}, M29 = {1, 2, 3| 1, 2, 4| 1, 2, 5},

M30 = {1, 2, 3| 1, 3, 4| 1, 3, 5}, M37 = {1, 3, 5| 2, 3, 5| 3, 4, 5}, M39 = {1, 2, 3, 4| 1, 2, 3, 5}, M43 = {1, 2, 3, 5| 1, 3, 4, 5}, M46 = {1, 2, 3, 5| 2, 3, 4, 5}

are best choices (see the fourth column of Table 5). In this case, M37 would be preferred if practice effect is a concern as it requires one fewer assessment from each participant than C4.

The designs that are efficient for all four fixed effects at a specific attrition level and the designs that are efficient across all attrition levels for a specific effect are presented in the last row and last column of Table 5, respectively. If all four effects are of interest and the expected attrition level is low, then three designs will lead to highly efficient tests: C6 = {1,3,4,5}, C4 = {1,2,3,5}, and M43 = {1,2,3,5 | 1,3,4,5}. On the other hand, if one is interested in the efficient design for the mean quadratic term but is unsure about the attrition level, then C4 = {1,2,3,5} is a good choice as it is efficient across all levels of attrition. Taking a closer look at the designs in the last row and column of Table 5, one can see that most of the efficient designs shared the following features. First, they all had four assessments per pattern. Second, all patterns in these designs have complete data on three occasions (first, third, and fifth), indicating that the designs with assessments spread out through the length of the study were quite efficient for the target fixed effects in the quadratic model.

Similar to the result for the linear model, the influence of attrition on the efficiency of a design depends on the type of design as well as the level of attrition. For CD designs, attrition tended to influence the allocation of the assessments but not the number of occasions per pattern. However, different to the result for the linear model, attrition shifted the repeated measures toward the occasion with lower instead of higher probability of attrition (e.g., C4 = {1,2,3,5} is preferred to C6 = {1,3,4,5} with high attrition). For PM designs, the most efficient PM designs under no attrition were not much affected by low attrition. However, their REs decreased substantially

by high attrition. For example, for the mean linear slope, M30 = {1,2,3 | 1,3,4 | 1,3,5} ($\widehat{RE}$ = 1) was most efficient under no or low attrition. However, with high attrition, its $\widehat{RE}$ dropped to .90. For the fixed effects related to the quadratic term, M43 = {1,2,3,5 | 1,3,4,5} ($\widehat{RE}$ = 1) was most efficient under no attrition. Its $\widehat{RE}$ dropped to .94 by low attrition and further dropped to .87 by high attrition.

Again, we examined the $\widehat{RE}$ of the CD design with five repeated measures (i.e., C7), Graham et al.'s (2001) design, and Mistler and Enders' (2012) design (i.e., M22) for the target effects in the quadratic model. The $\widehat{RE}$ of the CD design ranged from .80 to .86, the $\widehat{RE}$ of Graham et al.'s design ranged from .71 to .88, and the $\widehat{RE}$ of Mistler and Enders' design ranged from .83 to .92 across the four effects. Although the efficiencies of these designs for the quadratic model were generally higher than those for the linear model, none of them outperformed the most efficient designs identified using SEEDMC except that Mistler and Enders' (2012) design was efficient ($\widehat{RE}$ = .92) for the mean quadratic term and the predictor effect on the quadratic term. These results suggest that using the efficient designs identified by SEEDMC can save up to 41 % of the research budget [(1/0.71 − 1)×100 %] while achieving the same level of efficiency, compared to the commonly used or recommended designs in the previous literature.

Software implementation

We developed an R package "SEEDMC" to implement the proposed procedure (Jia, & Wu, 2015). This package can be downloaded from http://www.people.ku.edu/~wwei/. SEED

MC creates the design pool, and uses functions in the R package MplusAutomation (Hallquist & Wiley, 2014) to automate the Monte Carlo simulation in Mplus (Muthén & Muthén, 1998–2012). The current version of the package (i.e., SEEDMC 1.0.0) can accommodate user-specified unconditional linear and quadratic GCMs, and output the efficient designs for any single effect and multiple effects based on a selected threshold. The conditional models will be added in a later version. The R codes used in the two examples above (for the unconditional models with low attrition condition only) are included in Appendix 2. More examples and guidance on the use of SEEDMC 1.0.0 can be found in the help document of the package and from http://www.people.ku.edu/~wwei/.

Discussion and conclusion

In the article, we proposed a systematic procedure, SEEDMC, to search for efficient PM and CD designs in growth curve modeling using Monte Carlo simulation. SEEDMC allows researchers to identify the designs that are efficient for any single effect of interest in a hypothesized model or for multiple effects simultaneously. Furthermore, it allows researchers to identify efficient designs that are robust to MCAR attrition.

Applying SEEDMC to linear and quadratic GCMs, we found that the efficient designs for the fixed effects related to the slope factor in the linear model were PM designs with two or three repeated measures per pattern and CD designs with three repeated measures per pattern. The efficient designs for the fixed effects related to the slope factors in the quadratic model were PM designs with three or four repeated measures per pattern and CD designs with four repeated measures per pattern. Comparing to the most efficient CD designs, the PM designs generally allow researchers to collect one fewer repeated measure from each participant without sacrificing efficiency. These results echo those from the optimal design literature, indicating that the optimal number of repeated measures per pattern in the efficient designs for fixed effects should be as close as possible to the number of fixed effects. Our study also showed that the efficient designs for the predictor effect on a growth coefficient tended to have a great deal of overlap with those for the mean growth coefficient. This is good news because it means that researchers can resort to one design to maximize the efficiency of tests on both fixed effects.

With this procedure, we are also able to identify designs that are efficient across multiple effects and at the same time robust to MCAR attrition. For the linear model, a PM design with two occasions per pattern (M18 = {1,5 | 2,5 | 3,5 | 4,5}) is especially promising as it maximized the efficiency of the tests for both fixed effects related to the linear slope regardless of level of attrition. For the quadratic model, we failed to find a

design that is efficient across all target fixed effects and all levels of attrition. However, the designs that were efficient across all target effects at a specific level of attrition and the designs that were efficient across all attrition levels for a specific effect were successfully identified.

Furthermore, we found that the influence of attrition on the $\widehat{RE}$ of efficient designs varied depending on the type of design, the level of attrition, and the form of the change trajectory. For CD designs, attrition tended to influence the allocations of the repeated measures instead of the optimal number of repeated measures in the CD designs. For linear model, the repeated measures were shifted toward the occasions with higher probability of attrition, while for the quadratic model, the direction was the opposite. For PM designs, the most efficient PM design for the linear model was robust to MCAR attrition. However, the most efficient PM designs for the quadratic model tended to have a much lower $\widehat{RE}$ when attrition rates were high. The mechanism by which PM designs are influenced by attrition and why attrition favors one PM design over another deserve further investigation.

It is important to note that we examined only one set of parameter values for each of the linear and quadratic models. Thus, the generalizability of the efficient designs we identified in the two examples needs to be investigated. Researchers are encouraged to use the R package we have developed to identify the efficient designs for their own hypothesized models and parameter values.

As alluded to above, the biggest advantage of SEEDMC is its flexibility due to the use of Monte Carlo simulation. It can be extended in several ways to solve design problems in longitudinal studies. First, it can be extended to MNM greater or less than 5. For example, given a fixed length (e.g., 5 years), a smaller unit (e.g., half a year instead of 1 year) results in more possible measurement occasions which will result in a larger pool of CD and PM designs. In this case, the SEEDMC approach may be also used as a tool to identify the appropriate time interval. For example, if the 1-year interval leads to designs with similar efficiency to the half-a-year interval, then it would not be necessary to collect data in a smaller interval.

Second, SEEDMC can be extended to the change trajectories other than linear or quadratic (e.g., spline, exponential, and logistic). To do so, one must simply simulate data from the hypothesized form of the change trajectory and analyze the simulated data using the hypothesized model. This will be a big advantage over the traditional optimal design research because the asymptotic covariance matrix is more difficult to derive and the D-criterion is more difficult to optimize with complex nonlinear models.

Third, SEEDMC can be extended to categorical outcome variables such as ordinal or nominal outcome variables. The

categorical data can be simulated in the third step and analyzed using estimators specifically designed for categorical data such as weighted least squares types of estimators. Efficient designs can be then identified in the same way as described.

Fourth, SEEDMC can be extended to accommodate more types of simulated attrition. The attrition we simulated in the studies is determined by time. In longitudinal studies, the reason a participant misses a planned measurement can be related to his previous measures. In this case, having data on the previous measurement is important to correctly model the missingness mechanism and remove bias. In other words, if the probability of missing an observation is related to a previous measure but one does not have observed data for that measurement, then the missing data mechanism will be missing not at random (MNAR), resulting in biased parameter estimates. In this case, designs should be selected not just based on efficiency but also based on the accuracy of parameter estimates. In addition, attrition itself can be an interesting outcome in some longitudinal studies. For example, in an intervention study, attrition may reflect tolerability of the intervention. In this case, the probability of attrition at each measurement wave is often modeled jointly with the change model. The proposed procedure can also be extended to this case to search for designs that are jointly efficient for estimating effects in the change model as well as predictor effects of attrition.

Note that there are two ways to impose unplanned attrition. One is to impose attrition after computing the sample size for each design, as we do in the current article. With this way of taking into account attrition, the extent to which the designs that are cost comparable with planned missingness are robust to unplanned attrition is examined. However, the designs will no longer be cost comparable after imposing the attrition. The other is to compute the sample size for each design with the expected attrition rates taken into consideration. When some participants fail to attend planned assessment, a certain amount of the budget can be freed up to recruit more participants. In this case, the designs will be still cost comparable after imposing attrition at the design stage. This way of taking into account unplanned attrition may sound better and can be incorporated into the SEEDMC approach; however, it requires researchers to have an accurate prediction of the attrition rates. Otherwise, the actual cost will be different across the designs, which will introduce bias into the predicted efficiency of the designs.

In the current study, we assumed that the cost of collecting a single data point is the same, whether it represents the first data point of a new participant or the fifth data point of an already-enrolled participant. In practice, the cost of an initial measurement (i.e., recruiting a participant) can be much higher than that of an additional repeated measurement (i.e., maintaining a participant) (Tekle et al., 2011). If this cost

discrepancy were high, CD designs would receive an efficiency boost due to their lower average cost-per-data point. The SEEDMC procedure can be extended to account for the different costs. In this case, sample size associated with each design would be determined as a function of the total budget (B), cost of recruiting new sample (C_A), and cost of repeated measure (C_B), and the number of repeated measures in each missing data pattern in the design (T): $N = B/[C_A + (T - 1) \times C_B]$. This extension has been implemented in SEEDMC 1.0.0.

Finally, SEEDMC can be easily adjusted to accommodate other types of resource constraints such as sample size constraints. For instance, studies targeting participants with special characteristics (e.g., autism) may have a limited pool of participants. With a fixed sample size, the CD design with maximum number of repeated measures would be more efficient than both CD designs with a smaller number of repeated measures and all PM designs as it extracts the most information from the fixed set of participants. However, there might be PM designs that are almost as efficient as the CD design but cost less. In addition, when practice effects occur (e.g., in cognitive measurements; see Salthouse, Schroeder, & Ferrer, 2004), PM designs would be preferred because they lead to more valid results even though they are less efficient than the corresponding CD designs. SEEDMC can identify these designs by fixing the sample size instead of budget to be equal across designs. For example, for the quadratic model considered in the article, the efficiency of a PM design with four measures per pattern {1, 2, 3, 5| 1, 3, 4, 5} ($\widehat{RE}$ = 92 %) is an acceptable alternative to the CD design with five measures (the most efficient design) in detecting the mean quadratic rate given the same sample size and low attrition. SEEDMC 1.0.0 allows researchers to specify whether they have a budget or sample size constraint.

In sum, analysis of change is of central interest to many areas of research in social and behavioral sciences. The design issues for analysis of change have received much less attention than the analysis issues. Focusing on an important design issue, how to collect longitudinal data in a way to maximize the efficiency of analysis of change, we developed a flexible procedure to search for efficient CD and PM designs. Combined with power analysis, this procedure will offer researchers a data collection scheme that not only yields a desirable level of power but also does so at a minimized cost. We strongly advocate conducting such a search prior to data collection for any longitudinal studies. We also hope that this study will increase emphasis on design issues in longitudinal studies in social and behavioral sciences.

Appendixes

Appendix 1

Define the design pool

1) Complete data designs

Let MNM represent the maximum number of measurement occasions and T be the number of repeated measures. The total number of complete data designs with $T \geq 3$ and repeated measures at the first and last measurement occasions is given by

$$K_{CD} = \sum_{3 \leq T \leq MNM} \binom{MNM-2}{T-2}.$$

For example, if MNM = 5, there are seven possible complete data designs.

$$K_{CD} = \binom{5-2}{3-2} + \binom{5-2}{4-2} + \binom{5-2}{5-2} = 3 + 3 + 1 = 7$$

2) Missing data designs

Let T_c be the number of measurement occasions that have complete data ($T_c = 0, 1, \ldots, MNM-2$) and T_m be the number of missing observations in a specific response pattern ($1 \leq T_m \leq MNM - T_c - 1$), then the total number of PM designs is given by

$$K_{PM} = \sum_{1 \leq Tm < MNM - Tc -1, Tc=0} \binom{MNM}{Tc}$$
$$+ \sum_{1 \leq Tm \leq MNM - Tc -1, Tc>0} \binom{MNM}{Tc}.$$

For example, if T = 5, there are 48 possible missing data designs.

$$K_{PM} = \binom{5}{0} \times 3 + \binom{5}{1} \times 3 + \binom{5}{2} \times 2 + \binom{5}{3} \times 1$$
$$= 3 + 15 + 20 + 10 = 48$$

Appendix 2

R code used in the two examples

Example 1: Unconditional linear model

```
library(SEEDMC)
### Step 1: Create the design Pool
```

```
pool_linear <- designPool(type ="both",
traj = "linear",
time = 5, budget = 100000, unitcost = 20,
attrition = c(0, 0.075, 0.15, 0.225,
0.3))
summary(pool_linear)
## notations of the complete and missing
data designs
pool_linear$Complete$patterns_notation
pool_linear$Missing$patterns_notation
## Sample sizes of the complete and miss-
ing data designs
pool_linear$Complete$N
pool_linear$Missing$N
### Step 2: Create the Mplus syntax for hy-
pothesized model
fl <- matrix(NA, 5, 2)
fl[,1] <- 1
fl[,2] <- 0:4
latcov <- matrix(NA, 2, 2)
diag(latcov) <- c(28.776, 8.201)
latcov[1, 2] <- 1.56
latcov[2, 1] <- 1.56
latmean <- c(39.457, 8.063)
resvar <- 30
mod_linear <- modelMplus(fl = fl, latcov
= latcov,
latmean = latmean, resvar = resvar)
### Step 3: Monte Carlo simulation using
Mplus
seedmc_linear <- seedmcMplus(pattern =
pool_linear,
model = mod_linear,
nreps = 5000,
seed = 123321)
### Step 4: Compute relative efficiencies
of each design for
### target parameters
eff_linear <- topDesigns(x =
seedmc_linear,
param = c("Means.S"),
out = "mix", maxerr = 0.05,
recut = 0)
```

Example 2: Unconditional quadratic model

```
library(SEEDMC)
pool_quadratic <- designPool(type
="both", traj = "quadratic",
time = 5, budget = 100000, unitcost = 20,
attrition = c(0, 0.075, 0.15, 0.225,
0.3))
fl <- matrix(NA, 5, 3)
```

```
fl[,1] <- 1
fl[,2] <- 0:4
fl[,3] <- fl[,2]^2
latcov <- matrix(NA, 3, 3)
diag(latcov) <- c(107.08, 24.60, 1.22)
latcov[1, 2] <- latcov[2, 1] <- -3.69
latcov[1, 3] <- latcov[3, 1] <- -1.36
latcov[2, 3] <- latcov[3, 2] <- -4.96
latmean <- c(13.97, -1.15, 0.2)
resvar <- 41.98
mod_quadratic <- modelMplus(fl = fl,
latcov = latcov,
latmean = latmean, resvar = resvar)
seedmc_quadratic <- seedmcMplus(pattern
= pool_quadratic,
model = mod_quadratic,
nreps = 5000,
seed = 123321)
eff_quadratic <- topDesigns(x =
seedmc_quadratic,
param = c("Means.S1", "Means.S2"),
out = "mix", maxerr = 0.05,
recut = 0, acrosscut = 0.9)
```

References

Berger, M. P. F., & Wong, W. K. (2009). *An introduction to optimal designs for social and biomedical research*. Chichester, UK: John Wiley & Sons, Ltd. doi:10.1002/9780470746912

Biesanz, J. C., Deeb-Sossa, N., Papadakis, A. A., Bollen, K. A., & Curran, P. J. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods, 9*(1), 30.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation approach*. Hoboken, NJ: Wiley.

Chou, C.-P., Bentler, P. M., & Pentz, M. A. (1998). Comparison of two statistical approaches to study growth curves: The multilevel model and latent curve analysis. *Structural Equation Modeling, 5,* 247–266.

Enders, C. K. (2010). *Applied missing data analysis*. New York: The Guilford Press.

Galbraith, S., Stat, M., & Marschner, I. C. (2002). Guidelines for the design of clinical trials with longitudinal outcomes. *Controlled Clinical Trials, 23*(3), 257–273. doi:10.1016/S0197-2456(02)00205-2

Gan, L. (2011). *Optimal Longitudinal Cohort Designs and Variance Parameter Estimation* (Doctoral dissertation). Retrieved from http://escholarship.org/uc/item/537951j3

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. doi:10.1037/1082-989X.11.4.323

Graham, J. W., Taylor, B. J., & Cumsille, P. E. (2001). Planned missing data designs in the analysis of change. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 335–353). Washington, D. C.: American Psychological Association.

Gustavson, K., von Soest, T., Karevold, E., & Røysamb, E. (2012). Attrition and generalizability in longitudinal studies: Findings from a 15-year population-based study and a Monte Carlo simulation study. *BMC Public Health, 12*(1), 918.

Hallquist, M. & Wiley, J. (2014). MplusAutomation: Automating Mplus model estimation and interpretation. R package version 0.6-3. http://CRAN.R-project.org/package=MplusAutomation

Hancock, G. R., & French, B. F. (2013). Power analysis in covariance structure models. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 117–159). Charlotte, NC: Information Age Publishing, Inc.

Harel, O., Stratton, J., & Aseltine, R. (2012). Designed missingness to better estimate efficacy of behavioral studies. Technical Report (11–15), The Department of Statistics, University of Connecticut.

Hogue, C. M., Pornprasertmanit, S., Fry, M. D., Rhemtulla, M., & Little, T. D. (2013). Planned missing data designs for spline growth models in salivary cortisol research. *Measurement in Physical Education and Exercise Science, 17*(4), 310–325.

Jia, F., & Wu, W. (2015). SEEDMC: SEarch for Efficient Designs using Monte Carlo Simulation. R package version 1.0.0.

Jorgensen, T. D., Rhemtulla, M., Schoemann, A., McPherson, B., Wu, W., & Little, T. D. (2014). Optimal assignment methods in three-form planned missing data designs for longitudinal panel studies. *International Journal of Behavior Development. Advance Online Publication*. doi:10.1177/0165025414531094

Little, T. D., & Rhemtulla, M. (2013). Planned missing data designs for developmental researchers. *Child Development Perspectives, 7*(4), 199–204. doi:10.1111/cdep.12043

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*(1), 107–122. doi:10.1007/BF02294746

Mistler, S. A., & Enders, C. K. (2012). Planned missing data designs for developmental research. In B. Laursen, T. D. Little, & N. A. Card (Eds.), *Handbook of developmental research methods*. New York, NY: Guilford Press.

Moerbeek, M. (2005). Robustness properties of A-, D-, and E-optimal designs for polynomial growth models with autocorrelated errors. *Computational Statistics and Data Analysis, 48*(4), 765–778. doi:10.1016/j.csda.2004.03.014

Muthén, L.K. and Muthén, B.O. (1998–2012). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(4), 599–620. doi:10.1207/s15328007sem0904_8

Ortega-Azurduy, S., Tan, F., & Berger, M. (2008). The effect of dropout on the efficiency of D-optimal designs of linear mixed models. *Statistics in Medicine, 27*(14), 2601–2617.

Ouwens, M. J., Tan, P. E., & Berger, M. P. (2002). Maximin D-optimal designs for longitudinal mixed effects models. *Biometrics, 58*(4), 735–741.

Rhemtulla, M., Jia, F., Wu, W., & Little, T. D. (2014). Planned missing designs to optimize the efficiency of latent growth parameter estimates. *International Journal of Behavior Development. Advance online publication*. doi:10.1177/0165025413514324

Rubin, D. B. (1987). The calculation of posterior distributions by data augmentation: Comment: A noniterative sampling/importance re-sampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: The SIR algorithm. *Journal of the American Statistical Association, 82*(398), 543–546.

Salthouse, T. A., Schroeder, D. H., & Ferrer, E. (2004). Estimating retest effects in longitudinal assessments of cognitive functioning in adults between 18 and 60 years of age. *Developmental Psychology, 40*(5), 813–822. doi:10.1037/0012-1649.40.5.813

Singer, J., & Willett, J. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. USA: Oxford University Press.

Tekle, F. B., Tan, F. E., & Berger, M. P. (2008). D-optimal cohort designs for linear mixed-effects models. *Statistics in Medicine, 27*(14), 2586–2600.

Tekle, F. B., Tan, F. E., & Berger, M. P. (2011). Too many cohorts and repeated measurements are a waste of resources. *Journal of Clinical Epidemiology, 64*(12), 1383–1390.

Willett, J. B., Singer, J. D., & Martin, N. C. (1998). The design and analysis of longitudinal studies of development and psychopathology in context: Statistical models and methodological recommendations. *Development and Psychopathology, 10*(02), 395–426.

Wu, W., Selig, J. P., & Little, T. D. (2012). Longitudinal models. In T. D. Little (Ed.), *Oxford handbook of quantitative methods* (Vol. 2, pp. 387–410). New York, NY: Oxford University press.