




Regular Article

Sensitive periods and other timing hypotheses in developmental psychopathology: A tutorial

LillyBelle K. Deer¹ , Kylie K. Harrall^{2,3}, Deborah H. Glueck^{2,4}, Elysia Poggi Davis^{1,5} , Keith E. Muller³,
Dana Dabelea^{2,4,6} and Jenalee R. Doom¹ 

¹Department of Psychology, University of Denver, Denver, CO, USA, ²Lifecourse Epidemiology of Adiposity and Diabetes (LEAD) Center, University of Colorado Denver, Aurora, CO, USA, ³Department of Health Outcomes & Biomedical Informatics, University of Florida School of Medicine, Gainesville, FL, USA, ⁴Department of Pediatrics, University of Colorado School of Medicine, University of Colorado Denver, Aurora, CO, USA, ⁵Department of Pediatrics, University of California Irvine, Irvine, CA, USA and ⁶Department of Epidemiology, University of Colorado, Aurora, CO, USA

Abstract

Researchers often aim to assess whether repeated measures of an exposure are associated with repeated measures of an outcome. A question of particular interest is how associations between exposures and outcomes may differ over time. In other words, researchers may seek the best form of a temporal model. While several models are possible, researchers often consider a few key models. For example, researchers may hypothesize that an exposure measured during a sensitive period may be associated with repeated measures of the outcome over time. Alternatively, they may hypothesize that the exposure measured immediately before the current time period may be most strongly associated with the outcome at the current time. Finally, they may hypothesize that all prior exposures are important. Many analytic methods cannot compare and evaluate these alternative temporal models, perhaps because they make the restrictive assumption that the associations between exposures and outcomes remains constant over time. Instead, we provide a tutorial describing four temporal models that allow the associations between repeated measures of exposures and outcomes to vary, and showing how to test which temporal model is best supported by the data. By finding the best temporal model, developmental psychopathology researchers can find optimal windows for intervention.

Keywords: Development; longitudinal modeling; methods; sensitive periods; tutorial

(Received 25 August 2023; revised 28 June 2024; accepted 1 July 2024)

Introduction

Developmental psychopathology is a framework that examines risk and resilience processes across development, with the aim of preventing psychopathology and promoting positive development (Cicchetti & Toth, 2009; Cicchetti, 1993; Masten, 2006; Sroufe & Rutter, 1984). This framework is integrative, multidisciplinary, and involves the examination of the dynamic interplay of multilevel influences across development (Cicchetti & Toth, 2009; Masten, 2006). Integral to this approach is the idea that associations between exposures and outcomes often vary across time. As a result, factors during certain developmental periods may have a larger influence on the development of psychopathology and related outcomes (Cicchetti & Toth, 2009). Thus, examining exposure–outcome associations across development, and not just in one developmental period, is critical. Additionally, it is important to consider how these associations may change over time (i.e., strengthen, attenuate, or vary with time since exposure).

While it is known that many associations differ over time, many analytic approaches require the assumption that the association

between an exposure and outcome is constant, no matter what period during development the exposure and outcome are measured. In this tutorial, we present a more general approach to fitting longitudinal models of exposures and outcomes, which allows the association between the exposure and the outcome to change across time. In many cases, the temporal model for how exposures and outcomes are associated is not fully understood. In these cases, an exploratory, hypothesis-generating approach can be used. For these cases, we present a planned testing cascade to select the model which best fits the data. The testing cascade compares the utility of four temporal models, under the assumption that the exposure precedes the outcome.

In this tutorial, we present a general approach to fitting longitudinal models that are relevant to some of the central tenets of the developmental psychopathology framework. One of these tenets is that the developmental psychopathology framework is inherently transdisciplinary and examines multilevel influences on development (Cicchetti & Toth, 2009; Doom & Cicchetti, 2020; Rutter & Sroufe, 2000). In this approach, researchers can test transdisciplinary questions such as whether a child's attachment style might be associated with other domains (e.g., executive functioning; immune system function) across development. Another central tenet of the framework that the current approach is well-suited to explore is developmental cascades, or influences that spread across multiple levels, domains, or systems over time

Corresponding author: LillyBelle K. Deer; Email: lillybelle.deer@du.edu

Cite this article: Deer, L. K., Harrall, K. K., Glueck, D. H., Davis, E. P., Muller, K. E., Dabelea, D., & Doom, J. R. (2024). Sensitive periods and other timing hypotheses in developmental psychopathology: A tutorial. *Development and Psychopathology*, 1–9, <https://doi.org/10.1017/S0954579424001299>

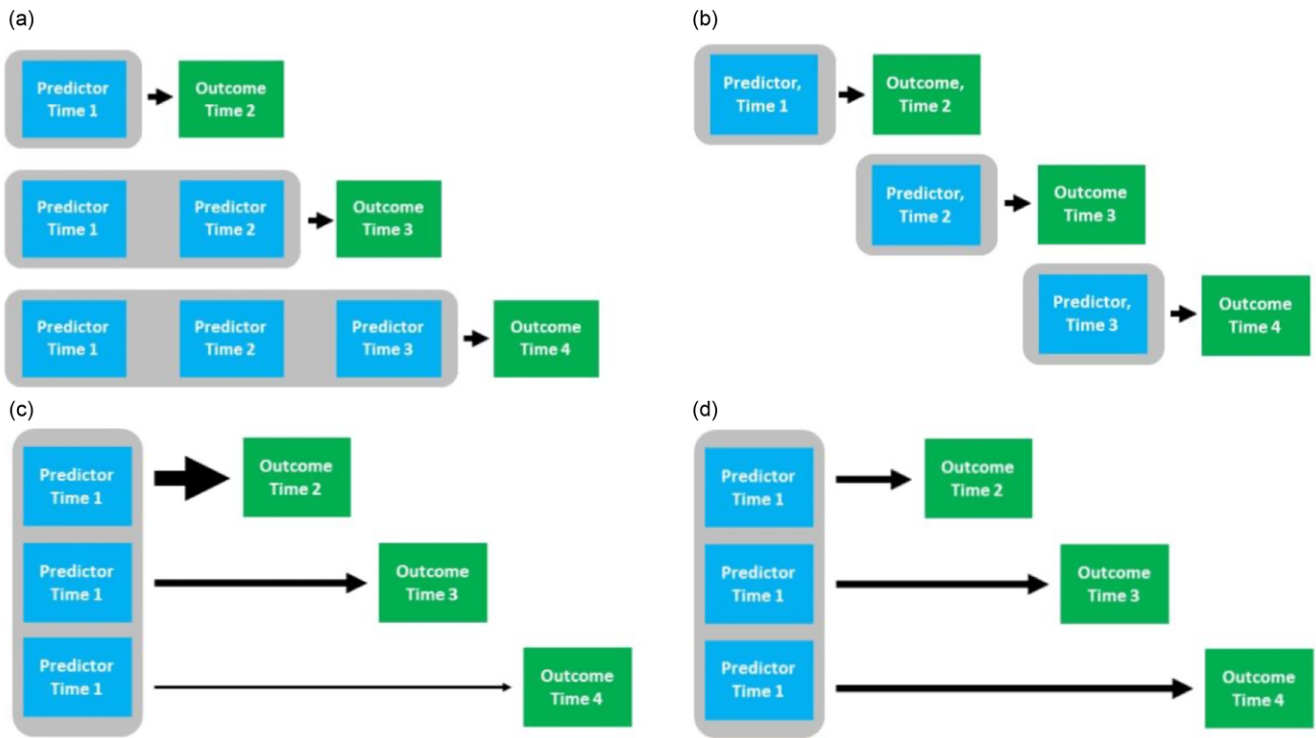


Figure 1. Four models to test hypotheses about different temporal associations between the predictor and the outcome. These include the (a) all-time-before model, (b) immediately-before model, (c) differential-sensitive model, and (d) stable-sensitive-period model.

(Doom & Cicchetti, 2020; Masten & Cicchetti, 2010; Sroufe & Rutter, 1984). Here researchers might be interested in understanding the cascade from parenting stress at one or multiple points in development to children's later externalizing symptoms. Another important consideration in the developmental psychopathology framework is the identification of sensitive periods in development (Cicchetti & Toth, 2009; Cicchetti, 2015b; Pollak, 2015). Researchers might examine whether there are specific developmental periods in which stress exposure has the largest impact on cognition. Finally, this approach allows for the testing of interactions between risk and protective factors over time, which is another central tenet of the developmental psychopathology framework (Doom & Cicchetti, 2020; Masten, 2001, 2006). In this approach, researchers could first assess the relation of maltreatment experiences at one or more points in development with later friendship quality. After establishing this main temporal model, researchers could then test whether having a safe, trusting adult might disrupt the association between maltreatment and friendship quality across development. The statistical approach described in the current manuscript is well suited for examining these central tenets of the developmental psychopathology framework.

Following from this framework, developmentalists often ask questions about how exposures at different time points across development may be associated with trajectories of later outcomes. The following questions are often of interest:

- Are all prior measurements of the predictor are associated with the outcome, with potentially different strengths of association, depending on the timing of the exposure and outcome measures (Figure 1a)?
- Is only the most recently recorded measure of the predictor associated with the outcome (Figure 1b)?

- Is there a *sensitive period* which best predicts the outcome across development? If so:
 - Does the strength of the association between the exposure measured during a sensitive period and the outcome differ depending on when the outcome was assessed (Figure 1c)?
 - Or, does the strength of the association between the exposure measured during a sensitive period and the outcome remain constant no matter when the outcome was assessed (Figure 1d)?

These questions are quite common in developmental research as well as across a number of other fields. For example, a researcher interested in the association between maternal sensitivity and offspring internalizing symptoms over development might hypothesize that all antecedent measurements of maternal sensitivity are associated with offspring internalizing symptoms (Figure 1a). Similarly, a researcher could hypothesize that hypothyroidism at any time point would be associated with expressive language at any point in development (Figure 1a). A researcher interested in the relation between peer deviancy and externalizing problems over time might hypothesize a recency effect, which would be best modeled using measurements of peer deviancy collected immediately before the assessment of externalizing problems (Figure 1b). Similarly, a researcher could hypothesize a recency effect of alcohol intake on risky decision-making where alcohol intake at the preceding time point shows the strongest association with risky decision-making at the following time point (Figure 1b). A researcher interested in examining the relation between iron deficiency and attention over time might hypothesize that there is an early sensitive period in which iron deficiency has the largest impact. They may also be interested in whether iron deficiency in that early sensitive period has the same magnitude of association with attention over time

Table 1. Glossary of key terms used in the manuscript

All-times-before model	For each outcome at each time, values of the predictor collected at any time before the outcome are used as the predictors. The model allows different strengths of associations between each predictor and outcome pair. This model provides evidence of a <i>cumulative effect</i> . The all-times-before model is shown in Figure 1a.
Design matrix	The fixed effect predictor matrix, or X matrix, which contains and organizes the independent variables of interest.
Differential-sensitive-period model	Each outcome at each time is only associated with the measurement of the predictor from time 1 (<i>sensitive period</i>). The model allows different strengths of association between each predictor and outcome pair. The differential-sensitive-period model is shown in Figure 1c.
Immediately-before model	For each outcome at each time, values of the independent variable measured at the time immediately-before the outcome are used as the predictors. The model allows different strengths of associations between each predictor and outcome pair. The immediately-before model is shown in Figure 1b.
Multiple design matrix model	A modeling approach originally described by Srivastava and Giles (1987a). The approach uses a data-driven approach to define the best-fitting design matrix across a series of nested design matrices.
Sensitive period	A sensitive period is a key developmental period during which heightened neural plasticity enables the opportunity for both growth and vulnerability (Knudsen, 2004).
Stable-sensitive-period model	Each outcome at each time is only associated with the measurement of the predictor from time 1 (<i>sensitive period</i>). The model assumes that the strength of the association is the same at each time period. The stable-sensitive-period model is shown in Figure 1d.

(Figure 1d), or whether this magnitude differs over time, and therefore across development (Figure 1c). Researchers interested in teratogenic effects of antenatal Zika infection on offspring cognition might similarly hypothesize that there is a sensitive period during pregnancy where infection may have greater effects relative to exposures outside of this period, though the strength of the association between antenatal Zika infection and the outcome could be the same or differ over time (Figure 1c,d). While each of these researchers might hypothesize certain temporal models for their research questions, it is difficult to know which is the best temporal model without comparing multiple models.

This tutorial describes how to test four hypotheses that are relevant to the psychopathology examples described above. We describe how to fit, compare, and evaluate models that align with each hypothesis. This tutorial has two main aims: (1) to describe how to fit several temporal models and (2) to describe how to compare the models and evaluate the associated hypotheses they represent, with an ordered sequence of tests. While any researcher could benefit from fitting the temporal model that fits their a priori hypothesis, the comparison of multiple models is especially helpful in situations where researchers have a hypothesized model but not enough prior literature to be confident that the proposed temporal model fits the data well.

The organization of the current tutorial is as follows. We first describe four possible temporal hypotheses, which correspond to four models. Mathematical descriptions and example code for the four models appear in the section entitled Supplementary Materials. Second, we show how to construct hypothesis tests to compare the four temporal models, and how a planned sequence of hypothesis tests allows investigators to choose the model that best fits the data. Third, we illustrate the approach by posing the questions in the context of an example longitudinal study of effortful control and body mass index (BMI), each measured at multiple time points. We provide a template for the interpretation of the results. Lastly, we contextualize recommendations for this modeling approach by discussing implications for conclusions in developmental psychopathology. See Table 1 for a glossary of terms used in the tutorial.

Analytic methods

This section details the general linear mixed model. First, we describe the model itself, then the components and assumptions, the hypothesis testing approach, four potential models that can be tested under this approach, and finally, how to compare models.

The general linear mixed model

The general linear mixed model (Laird & Ware, 1982) has several features which allow it to provide a good description of repeated measured data in developmental psychopathology. The mixed model can be used to analyze data with repeated measures of both exposures and outcomes. Additionally, the general linear mixed model can be used for research where some of the planned measures are missing, or do not occur exactly at the time the researchers planned. This model allows describing both the population average response, and participant-specific trajectories. Finally, the model permits modeling the covariance structure for the outcomes, which describes the variance of each measure, and the correlations among the repeated measures. A description of the model appears in the Supplementary Materials section.

Model components

Full descriptions of the components of the general linear mixed model appear in the Supplementary Materials section. The general linear mixed model has a vector of outcomes, a fixed effect exposure matrix, a matrix of fixed effect slopes and intercepts for the population average, a matrix of participant-specific random exposures, and a vector of errors. The fixed effects exposure matrix, also known as a design matrix, or an X matrix, contains known constants and organizes the variables of interest.

Variance assumptions

A complete description of observations with a normal distribution specifies not only the means and slopes but also the covariance matrix, which gives variances and correlations among observations. A general linear mixed model defines the covariance among observations with (1) the random effects exposure matrix, (2) the covariance matrix for random effects, and (3) the covariance matrix for errors.

When fitting repeated measurements from the same individual, the covariance structure among observations must accurately account for within-participant correlation and potentially unequal variances. Choosing the incorrect covariance structure can greatly

Table 2. Four model equations. The paired data for person i is notated as follows: y_{ij} is the outcome for person i at time point j , and x_{ij} is the predictor for person i at time point j

Model					
All-times-before	1				
		1	x_{i1}		
				1	x_{i1} x_{i2}
					1 x_{i1} x_{i2} x_{i3}
Immediately-before	1				
		1	x_{i1}		
				1	x_{i2}
					1 x_{i3}
Differential-sensitive-period	1				
		1	x_{i1}		
				1	x_{i1}
					1 x_{i1}
Stable-sensitive-period	1				
		1	x_{i1}		
				1	x_{i1}
					1 x_{i1}

increase the *Type I error rate* (Gurka et al., 2011), the probability that a hypothesis test incorrectly rejects the null.

Explicit guidance for choosing an appropriate covariance structure for repeated measures appears in multiple manuscripts (e.g., Cheng et al., 2010). Here, we discuss two possible structures. When the true covariance pattern is unknown, a common approach is to fit an unstructured covariance (Gurka et al., 2011; Harrall et al., 2023). The unstructured covariance structure is estimated using the repeated measurements for each study participant. When the variance of the outcome increases with age, as is common for such outcomes as BMI, many authors fit a random regression covariance model. The random regression model allows participant-specific intercepts and slopes and models the variance and covariance as a function of age or duration of time. The random regression approach was used in the four models in this manuscript.

The four models

Graphics are used to describe each model in Figure 1. Model equations are shown for each model in Table 2. The four models are presented in decreasing order of complexity, from most complicated, with many exposures, to least complicated, with fewer exposures.

For tutorial purposes, the models are presented with four repeated measurements. The assumption is made that the exposures are measured before the outcomes. Therefore, we will consider the exposure at time one, time two, and time three. We will consider the outcome at time two, time three, and time four. Exposure matrices for the four models appear in the Supplementary Materials section.

The four models described here can each model developmental processes. The all-times-before model associates the predictor in every prior time period with an outcome and can test different strengths of association at each time point. The immediately-

before model tests the influence of a predictor on an outcome close in time. The two sensitive period models test whether there is a time period where the predictor has greater effects on the outcome. The four models enable assessment of developmental processes and assess which are probable based on the timing of the measurements. We may not know ahead of time which is plausible, especially with new predictor and outcome or in novel time periods, which is why testing all the models is important.

All-times-before model

The all-times-before model is shown in Figure 1a and Table 2. For each outcome at each time, values of the exposure measured at any time before the outcome are used as the exposures. The model allows different strengths of associations between each exposure and outcome pair. For example, in a study with four time points, in this model, the measurement of the exposure at time one is associated with the outcome at time two, the measurements of the exposure at times one and two is associated with measurement of the outcome at time three, and the measurements of the exposure at times one, time two, and time three are associated with the outcome at time four.

Immediately-before model

The immediately-before model is shown in Figure 1b and Table 2. For each outcome at each time, values of the exposure variable measured at the time immediately before the outcome are used as the exposures. The model allows different strengths of associations between each exposure and outcome pair. In this model, in a study with four time points, the measurement of the exposure at time one is only associated with the measurement of the outcome at time two. Similarly, the measurement of the exposure at time two is only associated with the outcome at time three. Lastly, the measurement of the exposure at time three is only associated with the measurement of the outcome at time four.

Differential-sensitive-period model

The differential-sensitive-period model is shown in Figure 1c and Table 2, with the sensitive period assumed to be at time one. The exposure at time one is associated with the outcome at each later time point. The model allows different strengths of association between each exposure and outcome pair. For example, the magnitude of the relation might decrease as more time elapses between the measurements of the exposure and the outcome, or the magnitude of the relation might strengthen with age. In this case, in a study with four time points, the measurement of the exposure at time one is associated with the measurement of the outcome at times two, three, and four.

Stable-sensitive-period model

The stable-sensitive-period model is shown in Figure 1d and Table 2, with the sensitive period assumed to be at time one. The exposure at time one is associated with the outcome at each later time point. The model assumes that the strength of the association is the same at each time period. Like the differential-sensitive-period model, the measurement of the exposure at time one is associated with the measurements of the outcome at times two, three, and four.

Model comparison by hypothesis testing

To determine the importance of specific developmental periods or how an association may change over time, the all-times-before, the

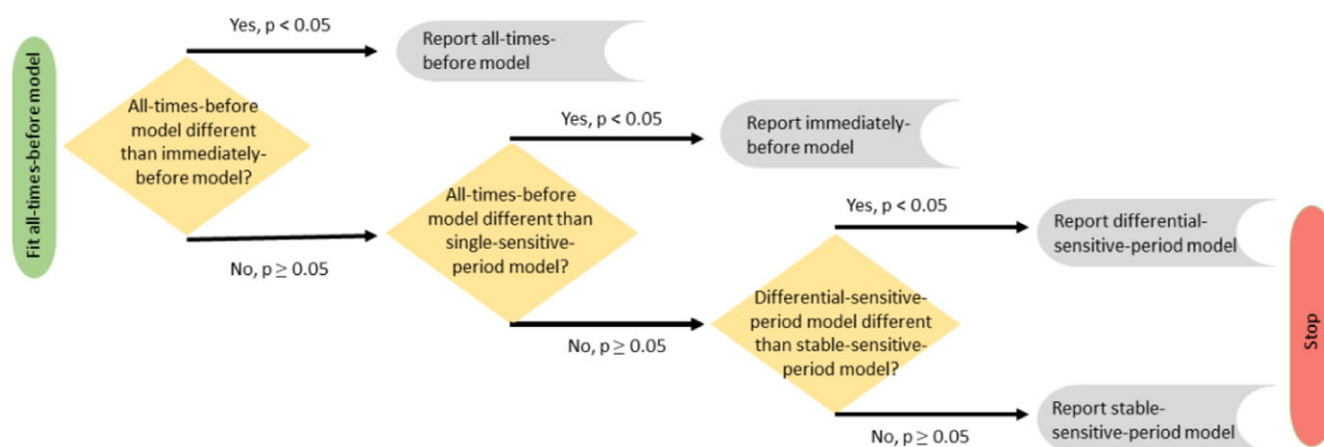


Figure 2. Flow chart for hypothesis tests deciding which model is most appropriate.

immediately-before, the differential-sensitive-period model, and the stable-sensitive-period models can be compared using a series of planned hypothesis tests. This series of tests uses a backward stepdown approach to find the best temporal model. If a sensitive period best predicts the outcome, then a test is used to determine whether the association varies across time or is stable across time. A flowchart for this series of hypothesis tests comparing the four models specified appears in Figure 2. Mathematical specifications of the hypothesis tests appear in the Supplementary Materials section.

Each hypothesis test compares a complex model with a simpler model nested within it. A simple model is considered to be nested within a complex model if the simple model can be obtained from the complex model by setting coefficients equal to zero. The immediately-before model and sensitive period models are nested within the all-times-before model and therefore can be compared. However, the sensitive period models are not nested within the immediately-before model and thus cannot be directly compared. The hypothesis testing sequence does allow arrival at any of the four models as the final model (see Figure 2).

The hypothesis testing sequence starts with the most complex model, the all-times-before model. Hypothesis tests are used to evaluate the utility of exposures. The tests provide comparisons of nested models. The first test compares the all-times-before model to a model nested within it: the immediately-before model. The second test compares the all-times-before model with a different model nested within it: differential-sensitive-period model.

In both tests described, the hypothesis test assesses the multiple-degrees-of-freedom null hypothesis of whether the coefficients which occur in the more complex model, but not the simpler model, are equal to zero. A multiple degrees of freedom hypothesis is a hypothesis which tests multiple simple hypotheses at the same time. A significant p -value (i.e., one less than the nominal level set by the investigator, often 0.05) suggests that the null should be rejected, indicating that at least one of the coefficients is different from zero, and therefore the more complex model is a good fit to the data. A nonsignificant p -value indicates that the null should not be rejected, and that the simple and more complex model seem equally likely. For reasons of parsimony, the hypothesis testing cascade suggests using the simpler model wherever reasonable.

Figure 2 shows one last planned hypothesis test, for the case when a model using a single sensitive period as an exposure is the best fit. The hypothesis test compares two nested models. Testing whether slopes and intercepts are the same across time periods

allows choosing between the differential-sensitive-period model and the stable-sensitive-period model.

Example matrices and code

Organizing the data for fitting the models and testing the hypotheses is key. Commented SAS code (SAS Institute, Cary, North Carolina) appears in the Supplementary Materials section, together with example datasets. The sample code demonstrates how to construct analytic datasets, hypotheses, and the four models presented in this manuscript. For clarity, the example analytic datasets and code assume four repeated measures in both outcome and predictor, and do not include covariates.

Example study

We provide an illustrative example to highlight the strength of this method. Deer et al. (2023) studied the longitudinal association from birth to adolescence between effortful control and BMI using data from a study of pregnant people and their offspring, described in Glynn et al. (2018). We use this example study to highlight how to test associations in the context of both a repeated exposure (in this example, effortful control) and a repeated outcome (in this example, child BMI). Effortful control, or the effortful regulation of thoughts and behaviors such as planning, attention, and inhibiting impulsive behaviors (Rothbart & Bates, 2006), has been identified as a transdiagnostic risk factor for mental and physical health (Anzman-Frasca et al., 2012; Santens et al., 2020; Thamocharan et al., 2013). BMI is a measure of cardiometabolic health and is predictive of weight stigma (Puhl & Lessard, 2020), being a victim of bullying (van Geel et al., 2014), and poorer socioemotional development (Black & Kassenboehmer, 2017), as well as physical and mental health across the lifespan (Jacobs et al., 2022; Sahoo et al., 2015). In the example presented, we operated using the temporality assumption, meaning that there must be temporal delays between causes and effects (Rothman et al., 2012). Here, we posited that alterations in effortful control would precede changes in BMI, but this does not preclude a hypothesis that BMI could predict later effortful control.

Example study methods

Participants in the example study were recruited during pregnancy if they spoke English, did not smoke, were over 18 years of age, were pregnant with a single child, and did not report using drugs or

alcohol during pregnancy. Offspring were included if they were born after 34 weeks of gestation, and had effortful control, height, and weight measured at six months of age. Effortful control was assessed using the Rothbart Temperament Questionnaires, which are a set of developmentally appropriate questionnaires to assess individual differences in reactivity and self-regulation across development (Ellis & Rothbart, 2001; Gartstein & Rothbart, 2003; Putnam *et al.*, 2006; Rothbart *et al.*, 2000; Simonds, 2006). These questionnaires were completed by mothers at 6-months, 1-year, 2-year, 5-year, 6.5-year, 9.5-year, and 11.5-year visits. Weight and length or height were measured by trained research assistants to calculate BMI at the 6-month, 1-year, 2-year, 5-year, 6.5-year, 9.5-year, 11.5-year, and 13-year visits. Infant sex at birth was collected from medical records and used as a moderator in the final model. Mothers reported their income and family size at the 6-month visit for the calculation of income-to-needs ratio (INR), which was used as a covariate in the final model. Further details appear in (Deer, Doom *et al.*, 2023).

Asking the four study questions in the context of the example study

The investigators were interested in assessing whether there were associations between repeated measures of effortful control and BMI. It was assumed that effortful control needed to be measured prior to BMI in order to have a relation with it.

We now rephrase the potential research questions in the context of the example study.

- Are all previous measurements of effortful control associated with the current measurement of BMI?
- Across development, is the measure of effortful control immediately prior to the measure of BMI associated with that measure of BMI?
- Is there a sensitive period when effortful control best predicts BMI over development? If so:
 - Does the strength of the association between effortful control in that sensitive period and BMI differ depending on when BMI was assessed?
 - Or, does the strength of the association between effortful control in that sensitive period and BMI remain constant no matter when BMI was assessed?

Example study analytic methods

All models used general linear mixed models to assess the relation between time-specific or repeated measures of effortful control with BMI trajectories. As prior research has identified a roughly quadratic pattern across this period of development (Wen *et al.*, 2012), each model was fit with a linear and quadratic age term. A random intercept for each participant and random slope for age was fit with an unstructured covariance between the random effects in order to account for within-person correlations between repeated measurements and the known increase in variance of BMI over time (Wen *et al.*, 2012). A series of four models were fit, following the planned hypothesis testing cascade shown in Figure 2, and used in the study (Deer, Doom *et al.*, 2023). First, the all-times-before model was compared to the immediately-before model. Then, the all-times-before model was compared to the differential-sensitive-period model. Finally, the differential-sensitive-period model was compared to the stable-sensitive-period model.

Results from the example study

Contrast tests indicated that the all-times-before model did not explain more variance than the immediately-before model [$p > 0.05$]. The second set of contrast tests indicated that the all-times-before model did not explain more variance than the differential-sensitive-period model [$p > 0.05$]. Finally, comparing the differential-sensitive-period model and the stable-sensitive-period model showed that the magnitude and direction of the association between 6-month effortful control and BMI at each time point did not differ ($p > 0.05$). Therefore, the stable-sensitive-period model was retained as the final model.

This example highlights the benefit of the strategy described in this manuscript. Deer, Doom *et al.* (2023) were able to find a model which best fit their data. Following the model identification stage, sex was added as a moderator of the association between effortful control at six months and BMI over time. Covariates were also added to the model.

Discussion

In this manuscript, we presented four questions that may be of scientific interest in experiments when both the exposure and the outcome were measured on multiple occasions across development. The questions allow an investigator to assess whether the exposure is associated with the outcome, to discern the temporal form of the association, and to test various developmental models. The models that follow from these questions are ideal for testing some of the central tenets of the developmental psychopathology framework, namely the transdisciplinary and multilevel nature of the approach, developmental cascades, understanding sensitive periods of development, and the interaction between risk and resilience factors shaping development over time.

The example study highlighted in this tutorial (Deer, Doom *et al.*, 2023) concluded that the stable-sensitive period model best fit the data. However, in other studies, other temporal models might have been better. Here we discuss the implications of choosing each of the potential models and what this might mean for intervention strategies. If the all-times-before model was the final model, this would indicate that all previous measurements of an exposure are important when explaining the relation between the exposure and the outcome. Thus, it might be best to intervene as early as possible. Alternatively, if the immediately-before model was the final model, this would indicate that timing does not matter as much. Therefore, it would be possible to reduce levels of an exposure at any age in order to have an effect on the outcome. If either of the sensitive period models were the final model, this would indicate that the best time for intervention on the exposure would be during that sensitive period.

The paper fills an important gap by providing statistical models to evaluate sensitive periods and temporal associations between variables in the field of developmental psychopathology. It is possible, or perhaps likely in many instances, that factors at multiple points during development may predict repeated measures of socioemotional or behavioral outcomes, with different magnitudes of association across time. The current paper introduces methods to compare these developmental models.

This manuscript provides analytic methodology for evaluating whether a sensitive period model best fits repeated measures data in contrast to models which include all measurements of the predictor prior to or immediately before measurement of the outcome. Some time periods are considered sensitive periods of development for specific outcomes. Sensitive periods are times

when physiological and behavioral systems are particularly open to inputs from the environment. Thus, different factors (e.g., stress, parenting, sociocultural influences) may play an outsized role in predicting the subsequent development of physical or mental health, in part due to changes in neurobiological functioning during sensitive periods (Cicchetti & Curtis, 2015; Cicchetti, 2016; Knudsen, 2004). Developmental researchers have identified several sensitive periods during which specific factors may predict various outcomes. Frequently studied sensitive periods include the prenatal period, infancy, early childhood, adolescence/puberty, the transition to adulthood, pregnancy, and the transition to parenthood (Davis et al., 2018; Davis & Narayan, 2020; Deer et al., 2023; Doom et al., 2024; Doom & Gunnar, 2013; Glynn et al., 2018; Kaliush et al., 2023; Powers & Casey, 2015; Schulenberg et al., 2004). The method described in the current manuscript allows researchers to test hypotheses about whether these periods are sensitive periods for specific outcomes, or whether models including other temporal patterns of exposures may have a better fit.

The manuscript provides a tutorial on approaches to test hypotheses about temporal issues and sensitive periods. The models described in this tutorial allow for assessment of how associations may change over time. The analytic approach described in this tutorial is a modification of the multiple design matrix approach of Srivastava (Srivastava & Giles, 1987b). For well-understood relations, it is possible that the researcher can choose the correct design matrix *a priori*. In many cases, however, it is not fully understood how the longitudinal predictors and outcomes of interest are associated. In both hypothesis-driven and exploratory cases, it is important that the research consider what plausible relations might exist. While it is possible that hypotheses may be developed with a literature search, comparing the fit of a series of different design matrices provides an empirical and rigorous test of timing questions. As a result, we strongly suggest that researchers test all of the models even if they have an *a priori* hypothesis that one will be the best fit. Additionally, it is important that researchers choose measures that are appropriate and measurable during the timeframe of interest. For example, if a researcher was interested in examining processes that occur during a specific developmental period (e.g., puberty), it would not be appropriate to continue modeling outcomes after this specific period.

When characterizing relations at sensitive periods, questions are often asked about how secondary predictors may moderate or mediate the association. When these types of questions arise and the temporal structure of a relation is unknown, we recommend that the research consider the two-step approach. First, use the modeling framework presented above to define the best-fitting temporal model of the primary predictor and outcome. Then, consider hypotheses about moderation or mediation in a second model, which includes the moderator or mediator of interest. This approach was used in the example study when sex was examined as a moderator following the initial stage of model identification.

Repeated measures of an exposure variable are typically correlated. In practice, correlations among exposures always introduce complexity in selecting a model. Muller and Fetterman (2002) provided extensive discussions of detecting and treating very high correlations among groups of exposures, often described as collinearity, and model selection. Cheng and colleagues (2010) provide some guidance on dealing with the complexity in “building a good enough mixed model.” If scientific considerations lead to exposures with moderate to high collinearity, then extra care must be taken in selecting tests to make correct decisions. One appealing

approach is a backwards-stepdown strategy (Cheng et al., 2010; Muller & Fetterman, 2002), with variables ordered by the principles of chronology and parsimony. Chronological order avoids anachronisms by requiring exposures occurring earlier in time to be listed earlier in the model. Parsimonious order requires, for example, main effects to be listed earlier in the model than interactions. A backwards approach begins by comparing the ultimate, most complex, model to the penultimate model, etc. The approach ensures collinearity is accurately adjusted for statistically and logically.

While the models and associated questions discussed in the current manuscript are quite common in developmental psychopathology and other fields, we do not cover an exhaustive list of potential hypotheses and models. There are a variety of other statistical models which can describe longitudinal series of exposures and outcomes. Clusterbased approaches group participants together with similar exposure patterns providing a categorical exposure which summarizes longitudinal exposure (e.g., Nagin et al., 2018; Nagin, 2014; Nagin & Odgers, 2010; Nagin & Tremblay, 2001). Cumulative approaches model exposure as area-under-the-curve, which is computed as the average product of the length of exposure and the strength of exposure. By contrast, the all-times-before model explicitly models the strength of the association between an exposure and outcome across different points in time. Thus, the all-times-before model cannot be considered a model of cumulative exposure.

This paper provides code written in SAS (SAS Institute Inc., 2016). The programming language was chosen because SAS allows hypothesis testing for the general linear mixed model using the Kenward-Roger corrections (Kenward & Roger, 1997, 2009). The Kenward-Roger corrections ensure accurate type I error rates, and thus the replicability of data analysis. Researchers who prefer other programming languages can use our code as a template to produce similar models in their preferred language.

The methods used in the paper have some limitations. The manuscript considers a restricted class of temporal models. The manuscript describes four models, and a series of hypothesis tests to compare the models identified to test common questions in developmental psychopathology (e.g., sensitive period, all-times-before, and immediately-before). The four models described share a common feature: the exposures are all measured before the outcome. For many pairs of exposures and outcomes, however, a model including both contemporaneous and antecedent measures makes sense. Examples include models of pollution and cardiac autonomic physiology (Parenteau et al., 2022), racism and mental health (Liu et al., 2023), and stress and parenting (Brown et al., 2020). For such pairs of exposures and outcomes, different models than the four considered in this paper apply, including the marching covariate, and the at-or-before models (Srivastava & Giles, 1987b). The marching covariate model uses contemporaneous exposures and outcomes and allows evaluating a hypothesis that only synchronous measures of exposures and outcomes are important. The at-or-before model uses exposures that are measured at the same time as the outcome, or at any time before the outcome. The at-or-before model allows evaluating whether antecedent or contemporaneous exposures are associated with an outcome. However, for the questions discussed in the current manuscript, the restricted class of temporal models are appropriate.

Another limitation of the models described in this manuscript is that they focus on a single sensitive period, rather than multiple sensitive periods. Numerous developmental processes are theorized to have multiple sensitive periods, such as cognitive (Thompson & Steinbeis, 2020) and emotional development

(Woodard & Pollak, 2020). Future work should build on the current models to test multiple sensitive periods.

The strength of the manuscript is that it provides detailed instructions for developmental psychopathology researchers to understand temporal relations. The techniques described in this manuscript could be used to identify the timing for averting adverse sequelae of developmental issues. A thorough understanding of sensitive periods and developmental trajectories is needed in order to create effective prevention and intervention programming (Cicchetti, 2015a; Davis et al., 2018; Deer et al., 2019; Granic, 2005; Hankin et al., 2023; Masten, 2006; Sroufe & Rutter, 1984). The hope is that this tutorial manuscript can add to the analytic repertoire of researchers in the field, and help researchers better understand approaches for repeated measures data.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S0954579424001299>.

Author contributions. LillyBelle K. Deer and Kylie K. Harrall contributed to the manuscript equally.

Funding statement. DHG, DD, KKH and KEM were supported by NIH/NIGMS funded grants 5R01GM121081-08, 3R25GM111901 and 3R25GM111901-04S1. LKD was supported by NHLBI grant F32HL165844. JRD was supported by NHLBI grant K01HL143159. JRD and EPD were supported by NHLBI grant R01HL155744 and EPD was supported by NIMH grant R01MH109662.

Competing interests. None.

References

- Anzman-Frasca, S., Stifter, C. A., & Birch, L. L. (2012). Temperament and childhood obesity risk: A review of the literature. *Journal of Developmental & Behavioral Pediatrics*, 33(9), 732–745. <https://doi.org/10.1097/DBP.0b013e3.1826a119f>
- Black, N., & Kassenboehmer, S. C. (2017). Getting weighed down: The effect of childhood obesity on the development of socioemotional skills. *Journal of Human Capital*, 11(2), 263–295. <https://doi.org/10.1086/692016>
- Brown, S. M., Doom, J. R., Lechuga-Peña, S., Watamura, S. E., & Koppels, T. (2020). Stress and parenting during the global COVID-19 pandemic. *Child Abuse & Neglect*, 110, 104699. <https://doi.org/10.1016/j.chiabu.2020.104699>
- Cheng, J., Edwards, L. J., Maldonado-Molina, M. M., Komro, K. A., & Muller, K. E. (2010). Real longitudinal data analysis for real people: Building a good enough mixed model. *Statistics in Medicine*, 29(4), 504–520. <https://doi.org/10.1002/sim.3775>
- Cicchetti, D. (1993). Developmental psychopathology: Reactions, reflections, projections. *Developmental Review*, 13(4), 471–502. <https://doi.org/10.1006/dev.1993.1021>
- Cicchetti, D. (2015a). Development and psychopathology. In *Developmental psychopathology* (pp. 1–23). John Wiley & Sons, Ltd, <https://doi.org/10.1002/9780470939383.ch1>
- Cicchetti, D. (2015b). Neural plasticity, sensitive periods, and psychopathology. *Development and Psychopathology*, 27(2), 319–320. <https://doi.org/10.1017/S0954579415000012>
- Cicchetti, D. (2016). Socioemotional, personality, and biological development: Illustrations from a multilevel developmental psychopathology perspective on child maltreatment. *Annual Review of Psychology*, 67(1), 187–211. <https://doi.org/10.1146/annurev-psych-12.2414-033259>
- Cicchetti, D., & Curtis, W. J. (2015). The developing brain and neural plasticity: Implications for Normality, psychopathology, and resilience. In *Developmental psychopathology* (pp. 1–64). John Wiley & Sons, Ltd, <https://doi.org/10.1002/9780470939390.ch1>
- Cicchetti, D., & Toth, S. L. (2009). The past achievements and future promises of developmental psychopathology: The coming of age of a discipline. *Journal of Child Psychology and Psychiatry*, 50(1-2), 16–25. <https://doi.org/10.1111/j.1469-7610.2008>
- Davis, E. P., Hankin, B. L., Swales, D. A., & Hoffman, M. C. (2018). An experimental test of the fetal programming hypothesis: Can we reduce child ontogenetic vulnerability to psychopathology by decreasing maternal depression? *Development and Psychopathology*, 30(3), 787–806. <https://doi.org/10.1017/S0954579418000470>
- Davis, E. P., & Narayan, A. J. (2020). Pregnancy as a period of risk, adaptation, and resilience for mothers and infants. *Development and Psychopathology*, 32(5), 1625–1639. <https://doi.org/10.1017/S0954579420001121>
- Deer, L. K., Bernard, K., & Hostinar, C. E. (2019). The influence of maternal stress and child maltreatment on offspring. In L. L. M. Welling, & T. K. Shackelford (Ed.), *The oxford handbook of evolutionary psychology and behavioral endocrinology*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190649739.013.9>
- Deer, L. K., Doom, J. R., Harrall, K. K., Glueck, D. H., Glynn, L. M., Sandman, C. A., & Davis, E. P. (2023). Infant effortful control predicts BMI trajectories from infancy to adolescence. *Pediatric Obesity*, 18(9), e13059. <https://doi.org/10.1111/ijpo.13059>
- Deer, L. K., Su, C., Thwaites, N. A., Davis, E. P., & Doom, J. R. (2023). A framework for testing pathways from prenatal stress-responsive hormones to cardiovascular disease risk. *Frontiers in Endocrinology*, 14(1). <https://doi.org/10.3389/fendo.2023.1111474>
- Doom, J. R., & Cicchetti, D. (2020). The developmental psychopathology of stress exposure in childhood. In K. L. Harkness, & E. P. Hayden (Ed.), *The oxford handbook of stress and mental health* (pp. 8). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190681777.013.12>
- Doom, J. R., Deer, L. K., Dabelea, D., LeBourgeois, M. K., Lumeng, J. C., Martin, C. K., Hankin, B. L., & Davis, E. P. (2024). Biological and behavioral pathways from prenatal depression to offspring cardiometabolic risk: Testing the developmental origins of health and disease hypothesis. *Developmental Psychology*, 60(9), 1620–1638. <https://doi.org/10.1037/dev0001704>
- Doom, J. R., & Gunnar, M. R. (2013). Stress physiology and developmental psychopathology: Past, present, and future. *Development and Psychopathology*, 25(4pt2), 1359–1373. <https://doi.org/10.1017/S0954579413000667>
- Ellis, L. K., & Rothbart, M. K. (2001). *Early adolescent temperament questionnaire-revised*. APA PsycTests. <https://doi.org/10.1037/t07624-000>
- Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the revised infant behavior questionnaire. *Infant Behavior and Development*, 26(1), 64–86. [https://doi.org/10.1016/S0163-6383\(02\)00169-8](https://doi.org/10.1016/S0163-6383(02)00169-8)
- Glynn, L. M., Howland, M. A., & Fox, M. (2018). Maternal programming: Application of a developmental psychopathology perspective. *Development and Psychopathology*, 30(3), 905–919. <https://doi.org/10.1017/S0954579418000524>
- Glynn, L. M., Howland, M. A., Sandman, C. A., Davis, E. P., Phelan, M., Baram, T. Z., & Stern, H. S. (2018). Prenatal maternal mood patterns predict child temperament and adolescent mental health. *Journal of Affective Disorders*, 228, 83–90. <https://doi.org/10.1016/j.jad.2017.11.065>
- Granic, I. (2005). Timing is everything: Developmental psychopathology from a dynamic systems perspective. *Developmental Review*, 25(3), 386–407. <https://doi.org/10.1016/j.dr.2005.10.005>
- Gurka, M. J., Edwards, L. J., & Muller, K. E. (2011). Avoiding bias in mixed model inference for fixed effects. *Statistics in Medicine*, 30(22), 2696–2707. <https://doi.org/10.1002/sim.4293>
- Hankin, B. L., Demers, C. H., Hennessey, E.-M. P., Perzow, S. E. D., Curran, M. C., Gallop, R. J., Hoffman, M. C., & Davis, E. P. (2023). Effect of brief interpersonal therapy on depression during pregnancy. *JAMA Psychiatry*, 80(6), 539–547. <https://doi.org/10.1001/jamapsychiatry.2023.0702>
- Harrall, K. K., Muller, K. E., Starling, A. P., Dabelea, D., Barton, K. E., Adgate, J. L., & Glueck, D. H. (2023). Power and sample size analysis for longitudinal mixed models of health in populations exposed to environmental contaminants: A tutorial. *BMC Medical Research Methodology*, 23(12). <https://doi.org/10.1186/s12874-022-01819-y>
- Jacobs, D. R., Woo, J. G., Sinaiko, A. R., Daniels, S. R., Ikonen, J., Juonala, M., Kartiosuo, N., Lehtimäki, T., Magnussen, C. G., Viikari, J. S. A., Zhang, N., Bazzano, L. A., Burns, T. L., Prineas, R. J., Steinberger, J., Urbina, E. M., Venn, A. J., Raitakari, O. T., & Dwyer, T. (2022). Childhood cardiovascular risk factors and adult cardiovascular events. *New England Journal of Medicine*, 386(20), 1877–1888. <https://doi.org/10.1056/NEJMoa2109191>

- Kaliush, P. R., Conratt, E., Kerig, P. K., Williams, P. G., & Crowell, S. E. (2023). A multilevel developmental psychopathology model of childbirth and the perinatal transition. *Development and Psychopathology*, 36(2), 533–544. <https://doi.org/10.1017/S0954579422001389>
- Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, 53(3), 983–997.
- Kenward, M. G., & Roger, J. H. (2009). An improved approximation to the precision of fixed effects from restricted maximum likelihood. *Computational Statistics & Data Analysis*, 53(7), 2583–2595. <https://doi.org/10.1016/j.csda.2008.12.013>
- Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. *Journal of Cognitive Neuroscience*, 16(8), 1412–1425. <https://doi.org/10.1162/0898929042304796>
- Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 38(4), 963–974. <https://doi.org/10.2307/2529876>
- Liu, S. R., Davis, E. P., Palma, A. M., Stern, H. S., Sandman, C. A., & Glynn, L. M. (2023). Experiences of COVID-19-related racism and impact on depression trajectories among racially/Ethnically minoritized adolescents. *Journal of Adolescent Health*, 72(6), 885–891. <https://doi.org/10.1016/j.jadohealth.2022.12.020>
- Masten, A. S. (2001). Ordinary magic. Resilience processes in development. *The American Psychologist*, 56(3), 227–238. <https://doi.org/10.1037//0003-066x.56.3.227>
- Masten, A. S. (2006). Developmental psychopathology: Pathways to the future. *International Journal of Behavioral Development*, 30(1), 47–54. <https://doi.org/10.1177/0165025406059974>
- Masten, A. S., & Cicchetti, D. (2010). Developmental cascades. *Development and Psychopathology*, 22(3), 491–495. <https://doi.org/10.1017/S0954579410000222>
- Muller, K. E., & Fetterman, B. (2002). *Regression and ANOVA: An integrated approach using SAS software*. SAS Institute Inc.
- Nagin, D. S. (2014). Group-based trajectory modeling: An overview. *Annals of Nutrition and Metabolism*, 65(2-3), 205–210. <https://doi.org/10.1159/000360229>
- Nagin, D. S., Jones, B. L., Passos, V. L., & Tremblay, R. E. (2018). Group-based multi-trajectory modeling. *Statistical Methods in Medical Research*, 27(7), 2015–2023. <https://doi.org/10.1177/0962280216673085>
- Nagin, D. S., & Odgers, C. L. (2010). Group-based trajectory modeling in clinical research. *Annual Review of Clinical Psychology*, 6(1), 109–138. <https://doi.org/10.1146/annurev.clinpsy.121208.131413>
- Nagin, D. S., & Tremblay, R. E. (2001). Analyzing developmental trajectories of distinct but related behaviors: A group-based method. *Psychological Methods*, 6(1), 18–34. <https://doi.org/10.1037/1082-989x.6.1.18>
- Parenteau, A. M., Alen, N. V., La, J., Luck, A. T., Teichrow, D. J., Daang, E. M., Nissen, A. T., Deer, L. K., & Hostinar, C. E. (2022). Associations of air pollution with peripheral inflammation and cardiac autonomic physiology in children. *New Directions for Child and Adolescent Development*, 2022(181-182), 125–154. <https://doi.org/10.1002/cad.20474>
- Pollak, S. D. (2015). Developmental psychopathology: Recent advances and future challenges. *World Psychiatry*, 14(3), 262–269. <https://doi.org/10.1002/wps.20237>
- Powers, A., & Casey, B. J. (2015). The adolescent brain and the emergence and peak of psychopathology. *Journal of Infant, Child, and Adolescent Psychotherapy*, 14(1), 3–15. <https://doi.org/10.1080/15289168.2015.1004889>
- Puhl, R. M., & Lessard, L. M. (2020). Weight stigma in youth: Prevalence, consequences, and considerations for clinical practice. *Current Obesity Reports*, 9(4), 402–411. <https://doi.org/10.1007/s13679-020-00408-8>
- Putnam, S. P., Gartstein, M. A., & Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: The early childhood behavior questionnaire. *Infant Behavior and Development*, 29(3), 386–401. <https://doi.org/10.1016/j.infbeh.2006.01.004>
- Rothbart, M. K., Ahadi, S. A., & Evans, D. E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology*, 78(1), 122–135. <https://doi.org/10.1037//0022-3514.78.1.122>
- Rothbart, M. K., & Bates, J. E. (2006). Temperament. In *Handbook of child psychology: Social, emotional, and personality development*. (vol. 3, 6th ed. pp. 99–166). John Wiley & Sons, Inc.
- Rothman, K. J., Lash, T. L., & Greenland, S. (2012). *Modern epidemiology (third, mid-cycle revision edition)*. LWW.
- Rutter, M., & Sroufe, L. A. (2000). Developmental psychopathology: Concepts and challenges. *Development and Psychopathology*, 12(3), 265–296. <https://doi.org/10.1017/s0954579400003023>
- Sahoo, K., Sahoo, B., Choudhury, A. K., Sofi, N. Y., Kumar, R., & Bhadoria, A. S. (2015). Childhood obesity: Causes and consequences. *Journal of Family Medicine and Primary Care*, 4(2), 187–192. <https://doi.org/10.4103/2249-4863.154628>
- Santens, E., Claes, L., Dierckx, E., & Dom, G. (2020). Effortful control – a transdiagnostic dimension underlying internalizing and externalizing psychopathology. *Neuropsychobiology*, 79(4-5), 255–269. <https://doi.org/10.1159/000506134>
- SAS Institute Inc (2016). *SAS® 9.4 Language Reference: Concepts* (6th ed.). SAS Institute Inc.
- Schulenberg, J. E., Sameroff, A. J., & Cicchetti, D. (2004). The transition to adulthood as a critical juncture in the course of psychopathology and mental health. *Development and Psychopathology*, 16(4), 799–806. <https://doi.org/10.1017/S0954579404040015>
- Simonds, J. (2006). *The role of reward sensitivity and response: Execution in childhood extraversion* (Publication No. 3238469) [Doctoral dissertation]. University of Oregon. ProQuest Dissertations & Theses Global.
- Srivastava, V. K., & Giles, D. E. A. (1987a). *Seemingly unrelated regression equations models: Estimation and inference*. CRC Press.
- Srivastava, V. K., & Giles, D. E. A. (1987b). *Seemingly unrelated regression equations models: Estimation and inference* (1st ed.). CRC Press.
- Sroufe, L. A., & Rutter, M. (1984). The domain of developmental psychopathology. *Child Development*, 55(1), 17–29. <https://doi.org/10.2307/1129832>
- Thamotharan, S., Lange, K., Zale, E. L., Huffhines, L., & Fields, S. (2013). The role of impulsivity in pediatric obesity and weight status: A meta-analytic review. *Clinical Psychology Review*, 33(2), 253–262. <https://doi.org/10.1016/j.cpr.2012.12.001>
- Thompson, A., & Steinbeis, N. (2020). Sensitive periods in executive function development. *Current Opinion in Behavioral Sciences*, 36, 98–105. <https://doi.org/10.1016/j.cobeha.2020.08.001>
- van Geel, M., Vedder, P., & Tanilon, J. (2014). Are overweight and obese youths more often bullied by their peers? A meta-analysis on the relation between weight status and bullying. *International Journal of Obesity*, 38(10), Article–10. <https://doi.org/10.1038/ijo.2014.117>
- Wen, X., Kleinman, K., Gillman, M. W., Rifas-Shiman, S. L., & Taveras, E. M. (2012). Childhood body mass index trajectories: Modeling, characterizing, pairwise correlations and socio-demographic predictors of trajectory characteristics. *BMC Medical Research Methodology*, 12. <https://doi.org/10.1186/1471-2288-12-38>
- Woodard, K., & Pollak, S. D. (2020). Is there evidence for sensitive periods in emotional development? *Current Opinion in Behavioral Sciences*, 36, 1–6. <https://doi.org/10.1016/j.cobeha.2020.05.004>