Longitudinal Analysis Manuscript: Working Draft
Longitudinal Analysis using data from the ABCD© Study

Samuel Hawes[a,*], Andrew K. Littlefield[b], Daniel A. Lopez[c], Kenneth J. Sher[d], Erin L. Thompson[e], Raul Gonzalez[e], Additional Co-authors[f], Wesley K. Thompson[g]

[a]*Florida International University, Center for Children & Families, 11200 SW 8th St, AHC-IV, Miami, USA, 33199*
[b]*Texas Tech University, Psychological Sciences, Lubbock, USA,*
[c]*University of Rochester Medical Center, Department of Public Health, Rochester, USA,*
[d]*University of Missouri, Psychological Sciences, Columbia, USA,*
[e]*Florida International University, Center for Children and Families, Miami, USA,*
[f]*xxxxxx University, xxxxxx Dept, xxxx, xxxxx,*
[g]*Laureate Institute for Brain Research, Center for Population Neuroscience and Genetics, Tulsa, USA,*

Abstract

The Adolescent Brain Cognitive Development^SM (ABCD) Study provides a unique opportunity for researchers to investigate developmental processes in a large, diverse cohort of youths, aged 9-10 at baseline and assessed annually for 10 years. Given the size and complexity of the ABCD Study©, researchers analyzing its data will encounter a myriad of methodological and analytical considerations. This review provides an examination of key concepts and techniques related to longitudinal analyses of the ABCD Study data. We discuss the importance of longitudinal data, focusing on the types of inferences that are possible when one assesses individuals across multiple time points, including: 1) characterization of the factors associated with variation in developmental trajectories; 2) assessment of how level and timing of exposures may impact subsequent development; 3) quantification of how variation in developmental domains may be associated with outcomes, including mediation models and reciprocal relationships. We emphasize the importance of selecting appropriate statistical models to address these research questions, e.g., accounting for correlation in repeated measurements, employing linear or non-linear models as indicated by the data, and using link functions that adequately model outcome distributions. By presenting the advantages and potential challenges of longitudinal analyses in the ABCD Study, this review seeks to equip researchers with foundational knowledge and tools to make informed decisions as they navigate and effectively analyze and interpret the multi-dimensional longitudinal data currently available.

Keywords: ABCD Study, longitudinal analyses, development

1. Introduction

The Adolescent Brain Cognitive Development (ABCD) Study® is the largest long-term investigation of neurodevelopment and child health in the United States. Conceived and initiated by the National Institutes of Health (NIH), this landmark prospective longitudinal study aims to transform our understanding of the genetic and environmental factors impacting neurodevelopment and their roles in behavioral and health outcomes across ten years of adolescence (Volkow et al. 2018). At its heart, the study is designed to chart the course of human development across multiple interacting domains from late childhood to early adulthood and to identify factors that lead to both positive and negative outcomes. Central to achieving these goals is the commitment of the ABCD Study® and its NIH funders to an open science framework, intended to facilitate sharing of data and analytical methods, by espousing practices that increase access, integrity,

---

[*]Corresponding author

Email addresses: `shawes@fiu.edu` (Samuel Hawes), `andrew.littlefield@ttu.edu` (Andrew K. Littlefield), `Daniel_Lopez2@URMC.Rochester.edu` (Daniel A. Lopez), `SherK@missouri.edu` (Kenneth J. Sher), `erthomps@fiu.edu` (Erin L. Thompson), `gonzara@fiu.edu` (Raul Gonzalez), `xxxxx@fiu.edu` (Additional Co-authors), `wes.stat@gmail.com` (Wesley K. Thompson)

and reproducibility of scientific research. In this sense, the ABCD Study is a collaboration with the larger research community.

The size and scope of the ABCD Study data allow the research community to perform a large variety of developmental analyses of both substantive and methodological interest, presenting a unique opportunity to significantly advance our understanding of how a multitude of biopsychosocial processes unfold across critical periods of development. In this paper, we describe models and methods for longitudinal analysis of ABCD Study data that can address these fundamental scientific aims, including: 1) characterization of the genetic and environmental factors associated with variation in developmental trajectories; 2) assessment of how the level and timing of exposures may impact subsequent neurodevelopment; 3) quantification of how variation in developmental domains may be associated with outcomes, including mediation models and reciprocal relationships. We instantiate these longitudinal analyses in worked examples using the ABCD Release 5.0 data with accompanying R scripts. Worked examples are available in Quarto files, accessible at **XXXX**.

1.1. The ABCD Study Data

The ABCD Study enrolled a cohort of n=11,880 participants born between 2006-2008 and aged between 9-11 years at baseline, as well as their parents/guardians. The study sample was recruited from household populations in defined catchment areas for each of the 21 (originally 22) study sites across the United States. Information regarding funding agencies, recruitment sites, investigators, and project organization can be obtained at https://abcdstudy.org/. The ABCD Study design is described in more detail in [1] and [2].

The ABCD Study is currently collecting longitudinal data on a rich variety of outcomes that will enable the construction of complex statistical models potentially incorporating factors from many domains. Each new wave of data collection provides another building block for characterizing developmental trajectories and implementing longitudinal analyses that allow researchers to characterize normative development, to identify variables that presage deviations from normative development, and to assess a range of variables associated with biopsychosocial outcomes of interest. These data include: 1) a neurocognitive battery

[3, 4]; 2) mental and physical health assessments [5]; 3) measures of culture and environment [6, 7]; 4) substance use [8]; 5) gender identity and sexual health [9]; 6) biospecimens [10]; 7) structural and functional brain imaging [11, 12]; 8) geolocation-based environmental exposure data [13]; 9) wearables, and mobile technology [14]; and 10) whole genome genotyping [15]. Many of these measures are collected at in-person annual visits, with brain imaging collected at baseline and every other year going forward. A limited number of assessments are collected in semi-annual telephone interviews between in-person visits.

Data are publicly released roughly annually, currently through the NIMH Data Archive (NDA). By necessity, the study's earliest data releases consisted primarily of one or two visits per participant. However, the most recent public release as of the writing of this paper (Release 5.0) contains data collected across five annual visits, including three brain imaging assessments (baseline, year 2 follow-up, and year 4 follow-up visits) for at least a subset of the cohort. Hence, starting with Release 5.0, it is feasible for researchers to begin focusing on the estimation of neurodevelopmental and other trajectories.

1.2.

> **Organization and Aims**
>
> - Part I. Longitudinal Research
>   - Identify fundamental concepts
> - Part II. Longitudinal Data
>   - Highlight key challenges
> - Part III. Longitudinal Analysis
>   - Methods & Analysis
> - Part IV. Supplemental materials
>   - Linked open-source resources

2. Developmental Research

2.1. Basic Concepts and Considerations

There are several important concepts to consider when conducting longitudinal analyses in a developmental context. These include different ways of thinking about developmental course, whether certain periods of development are relatively sensitive or insensitive to various types of insults or stressors, whether some time periods or situations inhibit the expression of individual differences due to extreme

environmental pressures, and whether the same behavior manifested at different times represents the same or different phenomena.

Moreover, in the case of developmentally-focused longitudinal research, each new measurement occasion not only provides a more extended portrait of the child's life course but also brings with it greater methodological opportunities to make use of statistical models that distinguish within- from between-person effects and that loosen constraints that need to be imposed on the furtherance of critical scientific questions.

For example, collecting two or more within-person observations on the same construct at different times enables estimation of individual rates of change (slopes); more observations allow for more precise estimates of individual slopes, as well as characterization of non-linear development. Rate of change or other trajectory characteristics may be more informative about individuals than the simple snapshots of level differences that cross-sectional data are limited to informing. Cross-sectional age-related differences across individuals are poor substitutes for longitudinal trajectory estimates except under highly restrictive assumptions, e.g., parallel trajectories and lack of age cohort and experience effects [16]. Appreciation of these and other issues can help to guide the analysis and interpretation of data and aid translation to clinical and public health applications.

2.1.1. Vulnerable periods.

Adolescent development progresses normatively from less mature to more mature levels of functioning. However, unique epochs and experiences can alter the course of this idealized form of development. Consider research that shows cannabis use during adolescence is associated with later psychosis to a greater degree than cannabis use initiated later in development [17, 18, 19, 20]; or similarly, experimental research on rodents that shows rodent brains to be especially sensitive to the neurotoxic effects of alcohol on brain structure and learning early in development, corresponding to early adolescence in humans [21, 22, 23]. In another example, longitudinal data from the National Consortium on Alcohol in Adolescence (NCANDA) show that binge drinking is associated more strongly with decrements in gray matter volume early in adolescence compared to later [24]. These examples highlight the importance of considering the role of vul-

nerable periods – e.g., temporal windows of rapid brain development or remodeling during which the effects of environmental stimuli on the developing brain may be particularly pronounced– when trying to establish an accurate understanding of the association between exposures and outcomes.

2.1.2. Developmental disturbances.

Whereas vulnerable periods heighten neurobiological susceptibility to environmental influences, at other times, environmental exposures will tend to suppress stability and disrupt the orderly stochastic process of normative development [e.g., 25]. This situation reflects a developmental disturbance in that the normal course of development is "altered" for a time by some time-limited process. In such cases, we might find that prediction of behavior in the period of the disturbance is reduced and/or, similarly, the behavior exhibited during the disturbance might have less predictive power with respect to distal outcomes compared to the behavior exhibited before and following the disrupted period. That is, once the environmental presses are removed (or the individual is removed from the environment), patterns of individual differences (and autoregressive effects) recover to levels similar to those prior to entering the environment. For example, in [24], recent binge drinking appears to be most predictive of gray matter volume trajectories, as opposed to more distal prior binge drinking or cumulative number of binge drinks, suggesting the potential for recovery of gray matter trajectories to prior levels of growth if binge drinking levels subside.

2.1.3. Developmental snares and cascade effects.

Normative development can also be upended by experiences (e.g., drug use) that, through various mechanisms, disrupt the normal flow of development wherein each stage establishes a platform for the next. For instance, substance use could lead to association with deviant peers, precluding opportunities for learning various adaptive skills and prosocial behaviors, in effect creating a "snare" that delays psychosocial development, such as maturing out of adolescent antisocial behavior [26]. Relatedly, the consequences of these types of events can cascade (e.g., school dropout, involvement in the criminal justice system) so that the effects of the snare are amplified [e.g., 27, 28]. Although conceptually distinct from vulnerable periods, both types

of developmental considerations highlight the importance of viewing behavior in the context of development and attempting to determine how various developmental pathways unfold. Longitudinal data are crucial in this context to assess individual levels of development prior to and following onset of experiences or other environmental factors; e.g., the ABCD Study collected data starting at ages 9-10 and hence before the onset of substance use for the vast majority of participants.

2.1.4 Intermediate processes and feedback loops Something here on the concept of mediation and cross-lagged relationships?

3. Longitudinal Data

3.1. Interpretation / Issues / Pitfalls & Assumption

The hallmark characteristic of longitudinal data analysis is the administration of repeated measurements of the same constructs on assessment targets (e.g., individuals, families) across time. The primary rationale for collecting longitudinal data is to assess within-person change over time, allowing researchers to estimate individual developmental trajectories and the genetic and/or environmental factors that may impact these trajectories. Administering repeated measurements more frequently or over longer time spans enables researchers to ask more nuanced questions and to make stronger inferences.

3.1.1. Two Time Points versus Three or More.

Although the clear leap from cross-sectional to the realm of longitudinal data involves going from one assessment to two or more assessments, there are also notable distinctions in designs based on two-assessment points versus three or more measurement occasions. Just as cross-sectional data can be informative in some situations, two waves of data can be beneficial in contexts such as when experimental manipulation is involved (e.g., pre/post tests), or if the central goal is prediction (e.g., trying to predict scores on Variable A at time T as a function of prior scores on Variable A and Variable B at time T-1). At the same time, data based on two assessments are inherently limited on multiple fronts. As [29] noted over forty years ago, "Two waves of data are better than one, but maybe not much better" (p. 744).

These sentiments are reflected in more contemporary recommendations regarding best-practice guidelines for prospective data, which increasingly emphasize the benefits of additional measurement occasions for trajectory estimation, model identification and accurate parameter inferences. It is also consistent with research recommending that developmental studies include three or more assessment points, given it is impossible for data based on two-time points to determine the shape of development (given that linear change is the only estimable form for two assessments waves; [see 30]). Research designs that include three (but preferably more) time points allow for non-linear trajectory estimation and increasingly nuanced analyses that more adequately tease apart sources of variation and covariation among the repeated assessments [31]– a key aspect of developmental research.

To illustrate, developmental theories are useful for understanding patterns of within-individual change over time (discussed in further detail, below); however, two data points provide meager information on change at the person level. This point is further underscored in a recent review of statistical models commonly touted as distinguishing within-individual vs between-individual sources of variance in which the study authors concluded "… researchers are limited when attempting to differentiate these sources of variation in psychological phenomenon when using two waves of data" and perhaps more concerning, "…the models discussed here do not offer a feasible way to overcome these inherent limitations" [32]. It is important to note, however, that despite the current focus on two-wave designs versus three or more assessment waves, garnering three assessment points is not a panacea for longitudinal modeling. Indeed, several contemporary longitudinal models designed to isolate within-individual variability (e.g., the Latent Curve Model with Structured Residuals; [33]) require at least four assessments to parameterize fully and, more generally, increasingly accurate and nuanced parameter estimates are obtained as more assessment occasions are used [30].

3.1.2. Types of stability and change

If one were to try to sum up what developmental trajectories in a living organism are exactly, one could plausibly argue they are the patterns of stability and change in its phenotypes as the organism traverses the life course. Symbolically, developmen-

tal trajectories can be expressed as fi(t), a possibly multivariate function of time t, specific to the ith individual and typically taking values in the real numbers for continuous phenotypes and the integers for discrete phenotypes. Ideally, t is a biologically meaningful temporal index (e.g., calendar age) as opposed to an exogenous progression of events (e.g., study visit number). Properties of interest might include rate of change over time, degree of smoothness (e.g., continuously differentiable), shape (e.g., polynomial or asymptotic behavior), how and how much fi(t) differs across individuals, and what factors predict either within-individual variation (at different times) or between-individual variation (either overall or at specific times).

There are a few different ways to think about patterns of stability and change (see Figure 1). Consider measuring school disengagement at the start of middle school and the end of middle school . A common first step may be to compare sixth graders' average disengagement values and eighth graders' disengagement values. This comparison of the average scores for the same group of individuals at multiple time points is referred to as "mean-level", as it provides information about change over time (or lack thereof) for an outcome of interest aggregated across members of a group. In contrast, "between-individual" stability could be assessed, e.g., by calculating the Spearman correlation between the values obtained at different time points (e.g., 'disengagement in sixth grade' with 'disengagement in eighth grade). This analysis focuses on the degree to which individuals retain their relative placement in a group across time. Consider someone who reported the lowest frequencies of disengagement in 6th grade and may report significantly higher disengagement over middle school (i.e., exhibit high levels of change), but report the lowest frequencies of disengagement in eighth grade. That is, the individual is manifesting rank-order stability, even in the context of high mean-level change.

Both types of stability and change are important. Mean-level change in certain traits might help to explain why, in general, populations of individuals tend to be particularly vulnerable to the effects of environmental factors in specific age ranges; rank-order stability might help to quantify the extent to which certain characteristics of the individual are more or less trait-like compared to others. For example, in some areas of development, considerable mean-level change occurs over time (e.g.,

changes in Big 5 personality traits; [34]), but exhibit relatively high rank-order stability, at least over shorter measurement intervals [34, 35, 36]. Despite the useful information afforded by examining mean-level and rank-order stability and change, these approaches are limited in that they provide little information about the overall patterns of within-individual change (i.e., trajectories) and, in turn, can result in fundamental misinterpretations about substantial or meaningful changes in an outcome of interest [37]. For example, questions related to the impact of early-onset substance use on brain development focus on changes within a given individual (i.e., intraindividual differences). The longitudinal nature of the ABCD Study, which will provide researchers with over ten time points for some constructs (e.g., substance use) across a ten-year period, allows for the study of within-person processes.
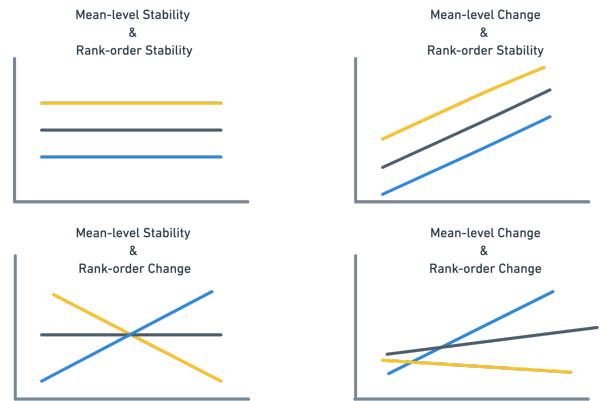

Figure 1: Types of Stability and Change

3.1.3. Use of appropriate longitudinal models

There is growing recognition that statistical models commonly applied to longitudinal data often fail to align with the developmental theory they are being used to assess [e.g., 37, 38, 32].

First, developmental studies typically involve the use of prospective data to inform theories that are concerned with clear within-person processes (e.g., how phenotypes change or remain stable within individuals over time, [e.g., 37]). Despite this, methods generally unsuited for disaggregating between- and within-person effects (e.g., cross-lagged panel models [CLPM]) remain common within various extant literatures. Fortunately, there exists a range

of models that have been proposed to tease apart between- and within-person sources of variance across time [see 32, 39]. Most of these contemporary alternatives incorporate time-specific latent variables to capture between-person sources of variance and model within-person deviations around an individual's mean (or trait) level across time (e.g., random-intercept cross-lagged panel model [RI-CLPM], [40]; latent curve models with structured residuals [LCM-SR], [33]). It is important to note however that these models require multiple assessments waves (e.g., four or more to fully specify the LCM-SR), additional expertise to overcome issues with model convergence, and appreciation of modeling assumptions when attempting to adjudicate among potential models in each research context [see 32, for further discussion].

Second, many statistical models assume certain characteristics about the data to which they are being applied. Common assumptions of parametric statistical models (e.g., linear mixed-effects models) include normality and equality of variances. These assumptions must be carefully considered before finalizing analytical approaches, so that valid inferences can be made from the data, as violation of a model's assumptions can substantively invalidate the interpretation of results. For example, longitudinal data can exhibit heterogeneous variability (i.e., the variance of the response changes over the duration of the study) that may need to be accounted for within a model. Another pertinent modeling assumption is whether trajectories are linear or non-linear. With 2-3 assessments per individual, only a linear model of within-person change is usually feasible; with more time points, higher level polynomials or models with more flexible trajectory shapes (e.g., curve smoothing via cubic splines) becomes possible [REFs]. Note, however, baseline age in the ABCD Study ranges over two full years; for some outcomes it may be desirable to include a (possibly nonlinear) effect of baseline age along with a linear effect of within-person change in age with only 2-3 assessment times [41]. As the study progresses and more times points are assessed, nonparametric curve estimation at the mean [GAMMs] and at the individual level may also become useful [42].

3.1.4. Continuous and Discrete Outcomes

Implementing valid and efficient statistical models requires an understanding of the type of data being analyzed. For example, repeated assessments within the ABCD Study can be based on continuous or discrete measures. Examples of discrete measures include repeated assessments of binary variables (e.g., past 12-month alcohol use disorder status measured across ten years), ordinal variables (e.g., caregiver-reported items measuring emotional and behavioral concerns via the Child Behavior Checklist including the categories of "Not True", "Somewhat True", and "Very True"), and count variables (e.g., number of cigarettes smoked per day). In many ways, the distributional assumptions of indicators used in longitudinal designs mirror the decision points and considerations when delineating across different types of discrete outcome variables, a topic that spans entire textbooks [e.g., see 43]. For example, the Mplus manual [44] includes examples of a) censored and censored-inflated models, b) linear growth models for binary or ordinal variables, c) linear growth models for a count outcome assuming a Poisson model, and d) linear growth models for a count outcome assuming a zero-inflated Poisson model. Beyond these highlighted examples, other distributions (e.g., negative binomial) can be assumed for the indicators when modeling longitudinal data [45]. These models can account for issues that can occur when working with discrete outcomes, including overdispersion [when the variance is higher than would be expected based on a given distribution; see [43]]. Given the sheer breadth of issues relevant to determining better models for discrete outcomes, it is not uncommon for texts on longitudinal data analysis to only cover models and approaches that assume continuous variables [e.g., 46]. However, some textbooks on categorical data analysis provide more detailed coverage of the myriad issues and modeling choices to consider when working with discrete outcomes [e.g., [43], Chapter 11 for matched pair/two-assessment designs; Chapter 12 for marginal and transitional models for repeated designs, such as generalized estimating equations, and Chapter 13 for random effects models for discrete outcomes].

3.1.5. Issues in attributing longitudinal change to development

One can observe systematic changes over time in a variable of interest and assume this change is attributable to development. However, various pitfalls with longitudinal data can complicate or even invalidate this conclusion. For example, if data missingness or participant dropout are related to the values of the outcome, changing sample compo-

sition as the study progresses can bias mean trajectory estimates (we describe this in more detail in Section 3.1.7 below). Another prerequisite for valid developmental interpretations of longitudinal data is to establish whether a construct is measured consistently over time, i.e., longitudinal measurement invariance [47, 48, 49]. Establishing longitudinal measurement invariance provides researchers with greater confidence that any change over time identified for a construct is attributable to individual change rather than a measurement artifact. For example, one study using data from the ABCD Study [[50]) found differential item functioning in two items from a brief delinquency measure, revealing significant bias in an arrest item across Black and White youth. More specifically, Black youth were more likely to report being arrested compared to White youth with similar levels of delinquency. Prevalence rates of delinquent behavior would have been severely biased if measurement invariance had not been tested.

Observed patterns of growth and decline often differ between cross-sectional vs. longitudinal effects [51] where subjects gain increasing experience with the assessment with each successive measurement occasion. Such experience effects on cognitive functioning have been demonstrated in adolescent longitudinal samples similar to ABCD [52] and highlight the need to consider these effects and address them analytically. In the case of performance-based measures [e.g., matrix reasoning related to neurocognitive functioning; see [51]], this can be due to "learning" the task from previous test administrations (e.g., someone taking the test a second time performs better than they did the first time simply as a function of having taken it before). Even in the case of non-performance-based measures (e.g., levels of depression), where one cannot easily make the argument that one has acquired some task-specific skill through learning, it has been observed that respondents tend to endorse lower levels on subsequent assessments [e.g., 53, 54] and this phenomenon has been well documented in research using structured diagnostic interviews [55]. While it is typically assumed that individuals are rescinding or telling us less information on follow-up interviews, there is reason to suspect that in some cases the initial assessment may be artefactually elevated [see 56].

Some longitudinal studies, e.g., accelerated longitudinal designs [ALDs; [16]] are especially well suited for discovering these effects and modeling

them. While ABCD is not an ALD, the variability in age (and grade in school) at the time of baseline recruitment (9 years, 0 months to 10 years, 11 months) allows some measures, collected every year, to be conceptualized as an ALD (e.g., substance use; prosocial behavior; family conflict; screen time). It is possible that in later waves, analyses will allow for disaggregating the confounded effects of age and the number of prior assessments. However, ABCD is fundamentally a single-cohort, longitudinal design, wherein number of prior assessments and age are mostly confounded, and for, perhaps, most analyses, the possible influence of experience effects needs to be kept in mind.

3.1.6. Covariance Structures

A central issue for repeated measurements on an individual is how to account for the correlated nature of the data. Traditional techniques, such as a standard regression or ANOVA model, assume residuals are independent and thus are inappropriate for designs that assess the same individuals across time. That is, given that residuals are no longer independent, the standard errors from the models are biased and can produce misleading inferential results. Although there are formal tests of independence for time series data (e.g., the Durbin-Watson statistic [57], more commonly, independence is assumed to be violated in study designs with repeated assessments. Therefore, an initial question to be addressed by a researcher analyzing prospective data is how to best model the covariance structure of said data.

Statistical models for longitudinal data include two main components to account for assumptions that are commonly violated when working with repeated measures data: a model for the covariance among repeated measures (involving the covariance among pairs of repeated outcomes on an individual), coupled with a model for the mean response and its dependence on covariates (e.g., sex at birth). There exists a range of methods to model covariance structures, each with its own set of tradeoffs between model fit and parsimony and which may be more or less appropriate for each specific application [e.g., see 58].

There are various approaches to adjust models for the lack of independence of residuals across time, be it for data with repeated assessments or other situations with so-called nested data (e.g., children nested within schools; siblings nested within fami-

lies). The most common approach is to use random effects. Essentially, random effects allow for variance estimates around fixed effects. A classic example (from [59]; [60]) involves math achievement measured among students nested within schools. In a basic, intercept-only model with no covariates, there would be one fixed effect (the grand mean, or intercept, of math achievement) and two random effects (one representing variation in the intercept, or the variation between schools, and another for the within-school residual, or how much variance in math achievement is left over after taking into account nesting in school). From this framework, each student's score would be a function of the fixed effect (the overall mean), the variation in scores that exists between schools, and the variation that exists among students within the same school. Assumptions about the variance and covariance components of this model dictate the form of the variance/covariance structure. For example, if we assume the random effects are independent (i.e., uncorrelated), the implied structure would be compound symmetry, where it is assumed the covariance of any two students in a single class is represented by a random effect (the variance between school means) and the covariance of any two students in different classes is zero. The assumptions of this relatively simple covariance structure can be relaxed, resulting in different covariance structures with additional parameters (see [60]).

In a similar fashion, individual growth models could be fit to a series of repeated assessments, where assessments are nested within individuals. For example, an unconditional linear growth model would involve two fixed effects – one for the intercept (an estimate of the average score when time is coded zero) and one for the linear slope (an estimate of the change in scores for each unit increase in time). Random effects could include a random effect for intercept (or an estimate of the variation in scores when time is coded as zero), a random effect for the linear slope (or an estimate of the variation in linear change across time), and a random effect for the within-person residual (or an estimate of the left over variance, or residual, of a given assessment when accounting for the intercept and linear slope). Assumptions regarding the covariation among the random effects also indicate different covariance structures. For example, it is typical to assume the random intercept and random slope components covary (e.g., if time is coded such

that the first time point = zero, the covariance estimate suggests that an individual's score at baseline relates to the amount of change exhibited across time). Further, particularly in structural equation model forms of this model, it is typically assumed the random effect for the within-person residual varies across assessments ([61]).

One alternative structure that attempts to handle the reality that correlations between repeated assessments tend to diminish across time is the autoregressive design. As the name implies, the structure assumes a subsequent measurement occasion (e.g., assessment at Wave 2) is regressed onto (that is, is predicted by) a prior measurement occasion (e.g., assessment at Wave 1). The most common type of autoregressive design is the AR(1), where assessments at time $T + 1$ are regressed on assessments at Time T. Identical to compound symmetry, this model assumes the variances are homogenous across time. Diverting from compound symmetry, this model assumes the correlations between repeated assessments decline exponentially across measurement occasions rather than remaining constant. That is, we can think of the underlying process as a stochastic one that wears itself out over time. For example, per the AR(1) structure, if the correlation between Time 1 and Time 2 data is thought to be .5, then the correlation between Time 1 and Time 3 data would be assumed to be $.5 \times .5 = .25$, and the correlation between Time 1 and Time 4 data would be assumed to be $.5 \times .5 \times .5 = .125$. As with compound symmetry, the basic AR(1) model is parsimonious in that it only requires two parameters (the variance of the assessments and the autoregressive coefficient). Notably, the assumption of constant autoregressive relations between assessments is often relaxed in commonly employed designs that use autoregressive modeling (e.g., cross-lagged panel models [CLPM]). These designs still typically assume an AR(1) process (e.g., it is sufficient to regress the Time 3 assessment onto the Time 2 assessment and is not necessary to also regress the Time 3 assessment onto the Time 1 assessment, which would result in an AR(2) process). However, the magnitude of these relations is often allowed to differ across different AR(1) pairs of assessment (e.g., the relation between Time 1 and Time 2 can be different from the relation between Time 2 and Time 3). These more commonly employed models also often relax the assumption of equal variances of the repeated assessments. Al-

though the AR(1) structure may involve a more realistic set of assumptions compared to compound symmetry, in that the AR(1) model allows for diminishing correlations across time, the basic AR(1) model, as well as autoregressive models more generally, can also suffer from several limitations in contexts that are common in prospective designs. In particular, recent work demonstrates that if a construct being assessed prospectively across time is trait-like in nature, then a simple AR(1) process fail to adequately account for this trait-like structure, with the downstream consequence that estimates derived from models based on AR structures (such as the CLPM) can be misleading and fail to adequately demarcate between- vs. within-person sources of variance [40]. Finally, it is also worth noting that in many longitudinal contexts, the time intervals between assessments are not equidistant and researchers need to consider carefully how to appropriately model time in their model and what this model implies.

3.1.7. Missing Data/Attrition

Attrition from a longitudinal panel study such as ABCD is inevitable and represents a potential threat to the validity of longitudinal analyses and cross-sectional analyses conducted at later time points, especially since attrition can only be expected to grow over time[62]. The ABCD Retention Workgroup (RW) employs a data-driven approach to examine, track, and intervene in these issues and while preliminary findings show participant race and parent education level to be associated with late and missing visits, although to date, attrition in ABCD has been minimal [63]. Ideally, one tries to minimize attrition through good retention practices from the outset via strategies designed to maintain engagement in the project [64, 65, 66]. However, even the best-executed studies need to anticipate growing attrition over the length of the study and implement analytic strategies designed to provide the most valid inferences. Perhaps the most key concern when dealing with data that is missing due to attrition is determining the degree of bias in retained variables that is a consequence of attrition. Such bias can attenuate generalizability, particularly if the pattern of missingness is not random (e.g., certain subsets of the population are more likely to drop out/not attend a visit). Assuming that the data are not missing completely at random, attention to the nature of the missingness and employing techniques designed to mitigate

attrition-related biases need to be considered in all longitudinal analyses.

Three types of missingness are considered in the literature [see 67], namely: a) missing completely at random (MCAR), b) missing at random (MAR), and c) missing not at random (MNAR). Data that are MCAR are a simple random sample of all data in a given dataset. MAR implies missing data are a random sample (i.e., does not hinge on some unmeasured variables) within strata of the measured covariates in a dataset (e.g., biological sex). Data that are MNAR are missing as a function of unobserved variables and may bias assocations even after conditioning on the observed covariates. [68] provides an excellent and easy-to-digest overview of further details involving missing data considerations.

Modern approaches for handling missing data, such as full information maximum likelihood, propensity weighting, auxiliary variables and multiple imputation avoid the biases of older approaches [see 69, 68]. [68] noted several "myths" regarding missing data. For example, Graham notes many assume the data must be minimally MAR to permit estimating procedures (such as maximum likelihood or multiple imputation) compared to other, more traditional approaches (e.g., using only complete case data). Violations of MAR impact both traditional and more modern data estimation procedures, though as noted by Graham, violations of MAR tend to have a greater effect on older methods. Graham thus suggests that imputing missing data is a better approach compared to listwise deletion in most circumstances, regardless of the model of missingness [i.e., MCAR, MAR, MNAR; see [68]].

3.1.8. Quantifying effect sizes longitudinally

Given that longitudinal data involve multiple sources of variance, quantifying effect sizes longitudinally is a more difficult task compared to deriving such estimates from cross-sectional data. An effect size can be defined as, "a population parameter (estimated in a sample) encapsulating the practical or clinical importance of a phenomenon under study." [70]. Common effect size metrics include the correlation r between two variables and the standardized difference between two means, Cohen's d [71]. An extensive discussion of effect sizes and their relevance for ABCD is given in [2]. Adjustments to common effect size calculations, such as Cohen's d, are required even when only two time points are con-

sidered (e.g., [72]]. [73] note there are multiple approaches to obtaining standardized within-person effects, and that commonly suggested approaches (e.g., global standardization) can be problematic [see 73, for more details]. Thus, obtaining effect size metrics based on standardized estimates that are relatively simple in cross-sectional data (such as r) becomes more complex in the context of prospective longitudinal data. [74] noted that equations for effects sizes used in studies involving growth modeling analysis (e.g., latent growth curve modeling) were not mathematically equivalent, and the effect sizes were not in the same metric as effect sizes from traditional analysis [see 74, for more details]. Given this issue, there have been various proposals for adjusting effect size measures in repeated assessments. [75] reviews the approach for effect size metrics for analyses based on growth modeling, including when considering linear and non-linear (i.e., quadratic) growth factors. [72] review various equations for effect size calculations relevant to combining estimates in meta-analysis with repeated measures and independent-groups designs. Other approaches to quantifying effect sizes longitudinally may be based on standardized estimates from models that more optimally disentangle between- and within-person sources of variance. As an example, within a random-intercept cross-lagged panel model (RI-CLPM) framework, standardized estimates between random intercepts (i.e., the correlation between two random intercepts for two different constructs assessed repeatedly) could be used to index the between-person relation, whereas standardized estimates among the structured residuals could be used as informing the effect sizes of within-person relations.

3.1.9. Longitudinal Data Structures

An ideal longitudinal study integrates (a) a well-articulated theoretical model, (b) an appropriate longitudinal data structure, and (c) a statistical model that is an operationalization of the theoretical model [76]. To accommodate various research questions and contexts, different types of longitudinal data and data structures have emerged (see Figure **X**). An understanding of these data structures is helpful, as they can warrant different types of longitudinal data analysis (LDA). Given that identifying a starting point for making comparisons is somewhat arbitrary, Curran and Bauer [2019; [77]] provide a nice on-ramp in first distinguishing between the use of "time-to-event" and "repeated mea-

sures" data. Although both model time, the former is concerned with whether and when an event occurs, whereas the later is focused on growth and change [77]. Time-to-event structures measure time from a well-defined origin point up to the occurrence of an event of interest. This data structure is most often analyzed using survival analysis methods (e.g., hazard rate models, event history analysis, failure-time models) and the time-to-event data can be based on a single assessment or include multiple recurrent or competing events. While much has been written about "time-to-event" data (**see xxx; xxx; xxxx**), including a recent analysis examining exclusionary discipline in schools using data from the ABCD Study ([50]), our emphasis will be given to the modeling of "repeated measures" data.
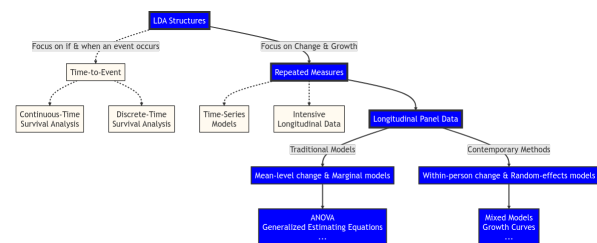
Figure 2: Longitudinal Data Analysis Structural Diagram

When discussing longitudinal analysis, we are most often talking about data collected on the same unit (e.g., individuals) across multiple measurement occasions. However, repeated-measures analysis is not a monolith, and it will serve us well to distinguish between a few of the most common types. One such approach to repeated measures analysis is the use of time-series models. These models generally consist of a long sequence of repeated measurements (50-100 measurements) on a single (or small number of) variable of interest. Time-series analysis is often used to predict/forecast temporal trends and cyclic patterns and is geared toward making inferences about prospective outcomes within a population (with relatively with less focus on inferring individual-level mechanisms and risk factors). A related type of repeated measures analysis is Intensive Longitudinal Data (ILD). Similar to time-series analysis, ILD models involve frequent measurements ($\sim$ 30-40 measurements) of the same individuals in a relatively circumspect period of time (e.g., experience sampling to obtain time series on many individuals). Although ILD models may include slightly fewer measurement occasions than time-series data, ILD models tend to have more

subjects than time series models ($\sim$ 50-100 subjects). This allows ILD models to examine short-term patterns by incorporating a time series model that can sometimes fit parameter estimates to each individual's data in order to model individual difference outcomes. The final type of repeated measures analysis that we will primarily focus on is the longitudinal panel model. These models follow a group of individuals— a panel (also referred to as a cohort) — across relatively fewer measurement occasions ($\sim$ 5-15), and are often interested in examining change across both, within-individuals and between-individuals.

While other longitudinal designs have their own unique strengths and applications, the longitudinal panel design is particularly well-suited for investigating developmental processes in the context of the ABCD Study. In the following sections, we will discuss various analytic methods commonly used to analyze longitudinal panel data, including growth models, mixed models, and a number of additional trajectory models. These methods provide valuable insights into within- and between-individual differences and are highly relevant for researchers working with the ABCD Study dataset. By focusing on these methods, we aim to equip readers with the knowledge necessary to conduct longitudinal research and perform analyses using the rich, longitudinal, and publicly available data from the ABCD Study.
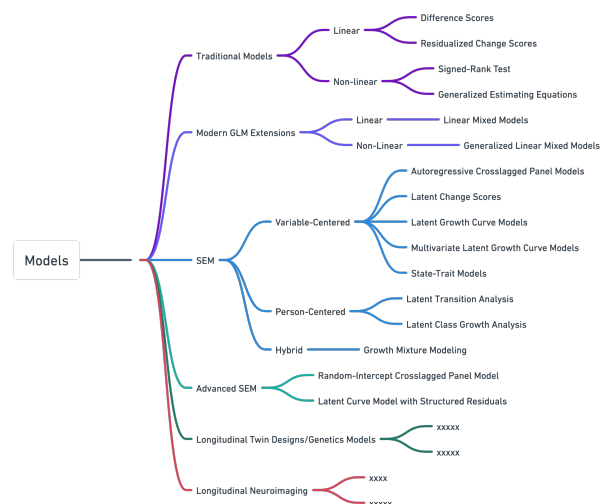
4. Longitudinal Analysis

4.0.1. Types of longitudinal panel models

With the large and continually expanding body of research on statistical methods for longitudinal analyses, determining which longitudinal model to implement can be challenging. The aim of this section is to help researchers navigate these many options to identify the statistical approach most appropriate to their unique research question when deciding on how to measure change over time. Notably, there are a myriad of viable ways one can go about grouping various types of longitudinal models for presentation.

Common examples include grouping by linear vs nonlinear models [76], the number of measurement occasions [31], and statistical equivalency [e.g., change scores vs. residualized change; see [78]]. The organization we use below overlaps in a number of ways with these examples, and in particular

with [77]. However, it is important to note that in each case, the chosen way of grouping is primarily intended to allow the reader to compare and contrast various analytical approaches. In the following sections, we briefly summarize the advantages/disadvantages of a series of longitudinal models organized into the following groupings: Traditional Models, Modern GLM Extensions, SEM, and Advanced SEM. We note that this is not an exhaustive review of each of these methods, and for more in-depth detail we do provide the reader with relevant resources. As aptly summarized by [77], "there are many exceptions, alternatives, nuances, 'what ifs', and 'but couldn't you's that aren't addressed here."

Traditional Models.

Traditional methods for longitudinal analysis primarily focus on modeling mean-level change, and how these changes may differ across groups or levels of some other variable. For example, is there a difference in average internalizing symptoms obtained across multiple assessments between boys and girls? Longitudinal models that focus on mean-level change are also referred to as marginal models and examples of specific methods include repeated measures ANOVA (and MANOVA), ANCOVA, and Generalized Estimating Equations (GEEs). Mean-level change models are commonly used when data is only available from 2 measurement occasions. For example, computing a difference score (e.g., mean internalizing scores at time 2 - mean internalizing scores at time 1) that can be used as an outcome in a subsequent GLM analysis (e.g., paired-samples t-test, repeated measures ANOVA) to test for differ-

ences in patterns of change over time and between groups. Additionally, the longitudinal signed-rank test, a nonparametric alternative to the paired t-test, can be a useful tool for analyzing non-normal paired data. Another common approach, often used in pre-/post-design studies but can be used with ABCD Study data, is to use residualized change score analysis to assess the degree of change in a variable, while controlling for its initial level [78].

For example, to examine change in cortico-limbic connectivity among ABCD participants, [79] regressed cortico-limbic connectivity at the year 2 follow-up on baseline cortico-limbic connectivity, which allowed the authors to examine the associations between negative life events and the variance of cortico-limbic connectivity unexplained by baseline connectivity. Similarly, [80] used a residualized-change model to examine the bidirectional influences of executive functioning and a general psychopathology factor 'p' across the first two years of the ABCD Study. Both studies were able to conclude associations between their constructs of interest that could not be accounted for by prior frequencies at baseline.

Traditional longitudinal, such as residualized chance score models, can be useful in some contexts (e.g., two measurement occasions), but overall, their practical utility for answering questions about developmental processes is limited. Perhaps most notably, these models do not allow for characterizing patterns of within-person change. This is a particularly important limitation since most psychological theories posit within-person processes (i.e., what will happen within a given individual). As such, traditional approaches often correspond poorly with most theoretical models of change and a failure to disaggregate between-person and within-person effects can result in consequential errors of inference [e.g., ecological fallacy, 37]. Moreover, even determining which of these procedures to use for comparing change over two time points across groups can be surprisingly complicated. A particularly vexing example is that of imbalanced baseline scores (i.e., when baseline scores are correlated with a covariate of interest), which can produce different conclusions across methods [e.g., **X versus X**; see [81], for a review]. Given these shortcomings, and the complexity of the issues surrounding some of these methods, it is typically recommended that researchers make use of more modern approaches for analyzing longitudinal data (and preferably make use of

data collected across three or more time points, as is currently true for many ABCD assessments).

Modern GLM Extensions.

Modern approaches to longitudinal data analysis have advanced beyond traditional methods by offering greater flexibility and a more in-depth understanding of within-person and between-person variability. Generalized Estimating Equations (GEE), Linear Mixed Models (LMM), Generalized Linear Mixed Models (GLMM), and Autoregressive Cross-Lagged Panel Models (ARCL) are examples of such contemporary techniques. GEE, an extension of Generalized Linear Models, combines the generalized linear model for non-normal outcomes with repeated measures (marginal) model and is suitable for analyzing correlated longitudinal data and modeling population-averaged effects. For example, [82] used GEE to obtain relative risks for psychiatric diagnoses among children in the ABCD Study with a family history of depression and used the ABCD Study sampling weights to generalize prevalence rates among 9 and 10 year-olds across the US. LMMs, also known as multilevel or hierarchical linear models, facilitate the simultaneous analysis of within-person and between-person variability, making them ideal for nested data structures or repeated measures. Within the ABCD Study, researchers may want to consider nesting by individual, family (i.e., siblings or twins), school or district, and/or site. GLMMs further extend the LMM framework to accommodate non-normal response variables, such as binary, count, or ordinal data, such as the use of ABCD data on substance use [e.g., 83] screen media use [84], and microstructure of the brain [85]. Finally, ARCL models are used to investigate reciprocal relationships between variables over time, as they estimate both autoregressive and cross-lagged effects [although, ARCL models are relatively less useful for teasing apart between-person and within-person sources of variances; see [86]].

The strengths of these modern methods lie in their ability to account for individual differences, within-person change, and time-varying predictors, thereby providing a more comprehensive understanding of complex relationships in longitudinal data. Despite these advantages, modern approaches may require more complex modeling assumptions and higher computational demands compared to traditional methods. Additionally, proper

model specification and the interpretation of results can be more challenging, especially in cases of high multicollinearity or missing data. However, modern longitudinal analysis methods have generally surpassed traditional methods in addressing a wider range of research questions, accommodating diverse data structures, and elucidating the intricate dynamics of developmental processes.

SEM.

Structural equation modeling (SEM) approaches have gained prominence in longitudinal data analysis due to their ability to estimate complex relationships among observed and latent variables while accounting for measurement error. SEM techniques can be categorized into variable-centered, person-centered, and hybrid approaches, each with unique strengths and limitations. The choice of method depends on the research question, data structure, and underlying assumptions.

Variable-centered approaches, such as latent change scores, latent growth curve models, multivariate (parallel process) latent growth curve models, and latent state-trait models, focus on examining relationships among variables and population-level patterns. These models offer a powerful means to estimate individual change trajectories and the relationships between growth parameters of different variables, while also decomposing observed measurements into latent state and trait components. However, these approaches may not adequately capture distinct subgroups of individuals who share similar patterns of change over time, which can be crucial for understanding heterogeneous developmental processes.

Person-centered approaches, including latent transition analysis and latent class growth analysis, address this limitation by identifying subgroups of individuals who share similar patterns of change. These models can reveal meaningful subpopulations and help researchers understand the factors that contribute to differences in developmental trajectories. For example, taking advantage of the large sample size of the ABCD Study, [87] found evidence of four subrgoups of youth with unique longitudinal patterns of depressive symptoms over time and identified risk factors that were differentially associated with the different trajectories. The use of such models allows for a more nuanced understanding of the associations between risk factors and change in

symptomatology as opposed to a snapshot of symptomatology at one time point. Nevertheless, person-centered methods may overlook important relationships among variables, which can be essential for understanding the dynamics of change over time.

Person-centered approaches, including latent transition analysis and latent class growth analysis, address this limitation by identifying subgroups of individuals who share similar patterns of change. These models can reveal meaningful subpopulations and help researchers understand the factors that contribute to differences in developmental trajectories. For example, taking advantage of the large sample size of the ABCD Study, [87] found evidence of four subgroups of youth with unique longitudinal patterns of depressive symptoms over time and identified risk factors that were differentially associated with the various trajectories. The use of such models allows for a more nuanced understanding of the associations between risk factors and change in symptomatology, as opposed to a snapshot of symptomatology at a single time point. Despite a range of potential model specifications for longitudinal mixture modeling, person-centered approaches tend to use parameterizations that default to settings found in popular software packages (e.g., Mplus). It has recently been demonstrated (see [88]) that the use of such specifications tends to identify the so-called "cat's cradle" solution (see [89]) that consists of "…(a) a consistently 'low' group, (b) an 'increase' group, (c) a 'decrease' group, and (d) a consistently 'high' group" ([89], p. 322). Indeed, [87] describe their four-group solution as follows: "Of all participants, 536 (10.80%) were classified as increasing, 269 (5.42%) as persistently high, 433 (8.73%) as decreasing, and 3724 (75.05%) as persistently low" ([87], p. 162). Although [89] cautioned that groups from these trajectory-based approaches should not be over-reified, this practice also remains common (e.g., [90]; [91]). Thus, though person-centered approaches can, in theory, help researchers understand the factors that contribute to differences in developmental trajectories, researchers should more thoughtfully consider alternative specifications (see [92], as an example) and be especially skeptical when default specifications identify these four prototypic groups.

Hybrid approaches, such as growth mixture modeling, combine aspects of both variable-centered and person-centered models, allowing for the identification of latent subgroups while also modeling rela-

tionships among growth parameters. This combination provides a more comprehensive understanding of longitudinal data by capturing both within- and between-person variability. However, hybrid models can be more complex, necessitating careful model specification, selection, and interpretation. Additionally, these methods may require larger sample sizes to ensure the stability and accuracy of results.

In summary, SEM approaches offer powerful tools for longitudinal data analysis, enabling researchers to investigate complex relationships, individual differences, and change dynamics over time. The choice between variable-centered, person-centered, and hybrid approaches depends on the research objectives and the nature of the data. Despite their limitations, these models have greatly advanced our understanding of developmental processes and the factors that contribute to individual differences in change trajectories.

Advanced SEM.

Advanced structural equation modeling (SEM) approaches, such as the RI-CLPM and LCM-SR models, have emerged to address more complex research questions and data structures in longitudinal analysis. These advanced models extend traditional SEM techniques, enabling researchers to disentangle within-person and between-person effects, as well as capture additional time-specific dependencies and associations that may not be accounted for by the latent growth factors.

The RI-CLPM enhances the traditional cross-lagged panel model by incorporating random intercepts, which allow for the separation of stable individual differences from the dynamic within-person associations between variables over time. Within-person variance in these models are captured by a series of latent variables that reflect time specific variance (i.e., the residual variance from the random intercept). These time-specific variables are referred to as structured residuals. Distinguishing between-person variance subsumed by the random intercept from the structured residuals is particularly valuable for understanding the time-specific effects of one variable on another, while accounting for the influence of individual differences. However, RI-CLPM may require larger sample sizes to ensure stability and accuracy of the estimates and can be computationally demanding. Using three

waves of ABCD Study data, [93] found a prospective association between psychopathology and childhood obesity as well as between childhood obesity and later eating behavior. The authors also showed that reciprocal associations were overestimated when stable, interindividual trait differences was not included in the model (i.e., via the random intercept).

LCM-SR, on the other hand, extends the RI-CLPM by including additional growth factors, such as a random linear slope. That is, the LCM-SR is a hybrid between a latent growth model and CLPM. This approach allows for a more comprehensive understanding of within-person change dynamics and factors influencing change over time. By including structured residuals, LCM-SR can capture additional time-specific relationships that are not explained by the latent growth factors. However, even more so than the RI-CLPM, LCM-SR comes with increased model complexity and requires careful specification and interpretation.

In conclusion, advanced SEM approaches for longitudinal data analysis provide valuable tools for addressing complex research questions and data structures. While they offer more nuanced insights into within-person change dynamics and the influence of individual differences, these models also come with certain limitations, such as the necessity of multiple assessments (e.g., four or more for LCM-SR), increased complexity, computational demands, and the need for careful model specification and interpretation. As with any statistical method, researchers should carefully consider their research objectives, data characteristics, and the assumptions of each model when selecting the most appropriate advanced SEM approach for longitudinal analysis. Given that these modeling approaches necessitate more waves of data, they are not yet commonly used with ABCD Study data. We anticipate that as more waves of ABCD data are publically released, these models can be used to address some of the pitfalls of the more traditional methods.

5. Discussion

As we enter the era of large-scale longitudinal investigations, it is essential to critically examine the various analytical methods that can be employed to glean insights from these rich datasets. The complex nature of longitudinal data demands sophisticated and well-suited methodologies to accu-
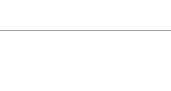
Figure 3: LDA Table

rately address research questions and minimize biases. This paper aimed to provide an overview of diverse longitudinal analysis techniques, with a particular emphasis on their application to extensive longitudinal studies such as the ABCD Study. Beyond contributing to the ever-growing body of knowledge on longitudinal data analysis, we hope this manuscript also serves as a valuable resource for researchers seeking to optimize the use of large-scale longitudinal investigations in advancing our understanding of human development and behavior. In this discussion, we will focus on the key findings and recommendations of our review and discuss potential innovations that can further enhance the utility of these methods.

We began by addressing fundamental concepts and considerations in longitudinal research that are essential for generating accurate and meaningful insights into developmental processes. Concepts such as vulnerable periods, developmental disturbances and snares, or cascade and experience effects (among many others), are instrumental in shaping the design, analysis, and interpretation of longitudinal studies. Together, these concepts provide a framework for understanding the mechanisms underlying the course of development, while also accounting for the complex interplay between individual development and the influence of environmental factors. By considering the intricate relationships among these factors, researchers can better identify the critical time periods, situations, and contexts that contribute to individual differences in developmental outcomes. This awareness enables more precise inferences regarding the causal relationships between exposures and outcomes, ultimately leading to more robust and meaningful findings that can help facilitate the translation of research findings into practical applications in clinical and public health settings.

We also discussed some of the opportunities, challenges, and pitfalls that arise when working with longitudinal data. Key issues include selecting appropriate methods to account for the intricacies of longitudinal data, addressing missing data in a way that minimizes biases, and determining suitable longitudinal data structures that align with research questions and context. To address these challenges, researchers should carefully consider issues such as study design, selection of methods that account for both within- and between-person sources of variance, and employing modern techniques, (e.g., FIML, multiple imputation) for handling missing data. By adhering to best practices in longitudinal research and remaining vigilant of potential pitfalls, researchers can effectively harness the power of longitudinal data to maximize the potential of their investigations and gain valuable insights into complex developmental processes, individual differences, and the underlying mechanisms that drive change over time.

The final section, along with associated code and additional resources made available as online supplements (), aims to serve as a resource for researchers seeking to understand and implement various longitudinal panel models. By providing an overview of different approaches, their strengths and limitations, and key considerations for their use, we hope to facilitate the selection of appropriate models tailored to specific research questions and data structures. It is essential for researchers to consider their research objectives, the characteristics of their data, and the assumptions underlying each model when choosing the most suitable approach for longitudinal analysis. We encourage researchers to consult the cited literature and online supplements for further guidance in selecting and implementing longitudinal models when using the ABCD Study dataset. As the field continues to advance, we anticipate the emergence of new methods and refinements to existing approaches, further expanding the toolkit available to researchers for the analysis of longitudinal data. By staying informed about developments in this area and critically evaluating the appropriateness of different models for their research questions, researchers can ensure that their longitudinal analyses are both rigorous and informative. Notably, in this vast and continually evolving field, with numerous models and approaches available to address a wide range of research questions, no single model is universally applicable or without

15

limitations. The diversity of methods ensures that researchers can find an appropriate tool for their specific needs. By familiarizing themselves with the various types of longitudinal models, researchers can more effectively navigate the complexities of longitudinal data and contribute valuable insights into the developmental processes and individual differences that shape human experience.

Part V: References

[1] H. Garavan, H. Bartsch, K. Conway, A. Decastro, R. Goldstein, S. Heeringa, T. Jernigan, A. Potter, W. Thompson, D. Zahs, Recruiting the abcd sample: Design considerations and procedures, Developmental cognitive neuroscience 32 (2018) 16–22.

[2] A. S. Dick, D. A. Lopez, A. L. Watts, S. Heeringa, C. Reuter, H. Bartsch, C. C. Fan, D. N. Kennedy, C. Palmer, A. Marshall, et al., Meaningful associations in the adolescent brain cognitive development study, NeuroImage 239 (2021) 118262.

[3] M. Luciana, J. M. Bjork, B. J. Nagel, D. M. Barch, R. Gonzalez, S. J. Nixon, M. T. Banich, Adolescent neurocognitive development and impacts of substance use: Overview of the adolescent brain cognitive development (ABCD) baseline neurocognition battery, Developmental cognitive neuroscience 32 (2018) 67–79, 00000 Publisher: Elsevier.

[4] W. K. Thompson, D. M. Barch, J. M. Bjork, R. Gonzalez, B. J. Nagel, S. J. Nixon, M. Luciana, The structure of cognition in 9 and 10 year-old children and associations with problem behaviors: Findings from the ABCD study's baseline neurocognitive battery, Developmental cognitive neuroscience 36 (2019) 100606, 00000 Publisher: Elsevier.

[5] D. M. Barch, M. D. Albaugh, S. Avenevoli, L. Chang, D. B. Clark, M. D. Glantz, J. J. Hudziak, T. L. Jernigan, S. F. Tapert, D. Yurgelun-Todd, Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: Rationale and description, Developmental cognitive neuroscience 32 (2018) 55–66, 00192 Publisher: Elsevier.

[6] R. Gonzalez, E. L. Thompson, M. Sanchez, A. Morris, M. R. Gonzalez, S. W. F. Ewing, M. J. Mason, J. Arroyo, K. Howlett, S. F. Tapert, et al., An update on the assessment of culture and environment in the abcd study®: Emerging literature and protocol updates over three measurement waves, Developmental cognitive neuroscience 52 (2021) 101021.

[7] R. A. Zucker, R. Gonzalez, S. W. F. Ewing, M. P. Paulus, J. Arroyo, A. Fuligni, A. S. Morris, M. Sanchez, T. Wills, Assessment of culture and environment in the Adolescent Brain and Cognitive Development Study: Rationale, description of measures, and early data, Developmental cognitive neuroscience 32 (2018) 107–120, 00064 Publisher: Elsevier.

[8] K. M. Lisdahl, S. Tapert, K. J. Sher, R. Gonzalez, S. J. Nixon, S. W. F. Ewing, K. P. Conway, A. Wallace, R. Sullivan, K. Hatcher, et al., Substance use patterns in 9-10 year olds: Baseline findings from the adolescent brain cognitive development (abcd) study, Drug and alcohol dependence 227 (2021) 108946.

[9] A. S. Potter, S. L. Dube, L. C. Barrios, S. Bookheimer, A. Espinoza, S. W. F. Ewing, E. G. Freedman, E. A. Hoffman, M. Ivanova, H. Jefferys, et al., Measurement of gender and sexuality in the adolescent brain cognitive development (abcd) study, Developmental Cognitive Neuroscience 53 (2022) 101057.

[10] K. A. Uban, M. K. Horton, J. Jacobus, C. Heyser, W. K. Thompson, S. F. Tapert, P. A. F. Madden, E. R. Sowell, Biospecimens and the ABCD study: Rationale, methods of collection, measurement and early data, Developmental Cognitive Neuroscience 32 (2018) 97–106, 00034. doi:10.1016/j.dcn.2018.03.005.

URL https://www.sciencedirect.com/science/article/pii/S1878929317301822

[11] B. J. Casey, T. Cannonier, M. I. Conley, A. O. Cohen, D. M. Barch, M. M. Heitzeg, M. E. Soules, T. Teslovich, D. V. Dellarco, H. Garavan, C. A. Orr, T. D. Wager, M. T. Banich, N. K. Speer, M. T. Sutherland, M. C. Riedel, A. S. Dick, J. M. Bjork, K. M. Thomas, B. Chaarani, M. H. Mejia, D. J. Hagler, M. Daniela Cornejo, C. S. Sicat, M. P. Harms, N. U. F. Dosenbach, M. Rosenberg, E. Earl, H. Bartsch, R. Watts, J. R. Polimeni, J. M. Kuperman, D. A. Fair, A. M. Dale, The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites, Developmental Cognitive Neuroscience 32 (2018) 43–54, 00232. doi:10.1016/j.dcn.2018.03.001.
URL http://www.sciencedirect.com/science/article/pii/S1878929317301214

[12] D. J. Hagler, S. Hatton, M. D. Cornejo, C. Makowski, D. A. Fair, A. S. Dick, M. T. Sutherland, B. J. Casey, D. M. Barch, M. P. Harms, R. Watts, J. M. Bjork, H. P. Garavan, L. Hilmer, C. J. Pung, C. S. Sicat, J. Kuperman, H. Bartsch, F. Xue, M. M. Heitzeg, A. R. Laird, T. T. Trinh, R. Gonzalez, S. F. Tapert, M. C. Riedel, L. M. Squeglia, L. W. Hyde, M. D. Rosenberg, E. A. Earl, K. D. Howlett, F. C. Baker, M. Soules, J. Diaz, O. R. de Leon, W. K. Thompson, M. C. Neale, M. Herting, E. R. Sowell, R. P. Alvarez, S. W. Hawes, M. Sanchez, J. Bodurka, F. J. Breslin, A. S. Morris, M. P. Paulus, W. K. Simmons, J. R. Polimeni, A. van der Kouwe, A. S. Nencka, K. M. Gray, C. Pierpaoli, J. A. Matochik, A. Noronha, W. M. Aklin, K. Conway, M. Glantz, E. Hoffman, R. Little, M. Lopez, V. Pariyadath, S. R. Weiss, D. L. Wolff-Hughes, R. DelCarmen-Wiggins, S. W. Feldstein Ewing, O. Miranda-Dominguez, B. J. Nagel, A. J. Perrone, D. T. Sturgeon, A. Goldstone, A. Pfefferbaum, K. M. Pohl, D. Prouty, K. Uban, S. Y. Bookheimer, M. Dapretto, A. Galvan, K. Bagot, J. Giedd, M. A. Infante, J. Jacobus, K. Patrick, P. D. Shilling, R. Desikan, Y. Li, L. Sugrue, M. T. Banich, N. Friedman, J. K. Hewitt, C. Hopfer, J. Sakai, J. Tanabe, L. B. Cottler, S. J. Nixon, L. Chang, C. Cloak, T. Ernst, G. Reeves, D. N. Kennedy, S. Heeringa, S. Peltier, J. Schulenberg, C. Sripada, R. A. Zucker, W. G. Iacono, M. Luciana, F. J. Calabro, D. B. Clark, D. A. Lewis, B. Luna, C. Schirda, T. Brima, J. J. Foxe, E. G. Freedman, D. W. Mruzek, M. J. Mason, R. Huber, E. McGlade, A. Prescot, P. F. Renshaw, D. A. Yurgelun-Todd, N. A. Allgaier, J. A. Dumas, M. Ivanova, A. Potter, P. Florsheim, C. Larson, K. Lisdahl, M. E. Charness, B. Fuemmeler, J. M. Hettema, H. H. Maes, J. Steinberg, A. P. Anokhin, P. Glaser, A. C. Heath, P. A. Madden, A. Baskin-Sommers, R. T. Constable, S. J. Grant, G. J. Dowling, S. A. Brown, T. L. Jernigan, A. M. Dale, Image processing and analysis methods for the Adolescent Brain Cognitive Development Study, NeuroImage 202 (2019) 116091, 00000. doi:10.1016/j.neuroimage.2019.116091.
URL http://www.sciencedirect.com/science/article/pii/S1053811919306822

[13] C. C. Fan, A. Marshall, H. Smolker, M. R. Gonzalez, S. F. Tapert, D. M. Barch, E. Sowell, G. J. Dowling, C. Cardenas-Iniguez, J. Ross, et al., Adolescent brain cognitive development (abcd) study linked external data (led): Protocol and practices for geocoding and assignment of environmental data, Developmental cognitive neuroscience 52 (2021) 101030.

[14] K. S. Bagot, S. A. Matthews, M. Mason, L. M. Squeglia, J. Fowler, K. Gray, M. Herting, A. May, I. Colrain, J. Godino, Current, future and potential use of mobile and wearable technologies and social media data in the ABCD study to increase understanding of contributors to child health, Developmental cognitive neuroscience 32 (2018) 121–129, 00046 Publisher: Elsevier.

[15] R. J. Loughnan, C. E. Palmer, W. K. Thompson, A. M. Dale, T. L. Jernigan, C. C. Fan, Polygenic score of intelligence is more predictive of crystallized than fluid performance among children, bioRxiv (2020) 63751200002 Publisher: Cold Spring Harbor Laboratory.

[16] W. K. Thompson, J. Hallmayer, R. O'Hara, A. D. N. Initiative, Design considerations for characterizing psychiatric trajectories across the lifespan: application to effects of apoe-$\varepsilon 4$ on cerebral cortical thickness in alzheimer's disease, American Journal of Psychiatry 168 (9) (2011) 894–903.

[17] L. Arseneault, M. Cannon, R. Poulton, R. Murray, A. Caspi, T. E. Moffitt, Cannabis use in adolescence and risk for adult psychosis: longitudinal prospective study, Bmj 325 (7374) (2002) 1212–1213.

[18] J. Bechtold, A. Hipwell, D. A. Lewis, R. Loeber, D. Pardini, Concurrent and sustained cumulative effects of adolescent marijuana use on subclinical psychotic symptoms, American Journal of Psychiatry 173 (8) (2016) 781–789.

[19] A. Hasan, R. von Keller, C. M. Friemel, W. Hall, M. Schneider, D. Koethe, F. M. Leweke, W. Strube, E. Hoch, Cannabis use and psychosis: a review of reviews, European archives of psychiatry and clinical neuroscience 270 (2020) 403–412.

[20] D. M. Semple, A. M. McIntosh, S. M. Lawrie, Cannabis as a risk factor for psychosis: systematic review, Journal of psychopharmacology 19 (2) (2005) 187–194.

[21] L. P. Spear, Consequences of adolescent use of alcohol and other drugs: Studies using rodent models, Neuroscience & Biobehavioral Reviews 70 (2016) 228–243.

[22] F. T. Crews, C. J. Braun, B. Hoplight, R. C. Switzer III, D. J. Knapp, Binge ethanol consumption causes differential brain damage in young adolescent rats compared with adult rats, Alcoholism: Clinical and Experimental Research 24 (11) (2000) 1712–1723.

[23] Z. Ji, L. Yuan, X. Lu, H. Ding, J. Luo, Z.-J. Ke, Binge alcohol exposure causes neurobehavioral deficits and gsk3β activation in the hippocampus of adolescent rats, Scientific Reports 8 (1) (2018) 3088.

[24] M. Infante, S. Eberson, Y. Zhang, T. Brumback, S. Brown, I. Colrain, F. Baker, D. Clark, M. De Bellis, D. Goldston, et al., Adolescent binge drinking is associated with accelerated decline of gray matter volume, Cerebral Cortex 32 (12) (2022) 2611–2620.

[25] J. Schulenberg, J. Maslowsky, M. E. Patrick, M. Martz, Substance use in the context of adolescent development, in: R. A. Zucker, S. A. Brown (Eds.), The Oxford handbook of adolescent substance abuse, Oxford University Press, 2019, pp. 19–36.

[26] T. E. Moffitt, Life-course-persistent versus adolescence-limited antisocial behavior, Developmental Psychopathology: Volume Three: Risk, Disorder, and Adaptation (2015) 570–598.

[27] A. S. Masten, G. I. Roisman, J. D. Long, K. B. Burt, J. Obradović, J. R. Riley, K. Boelcke-Stennes, A. Telle-

gen, Developmental cascades: linking academic achievement and externalizing and internalizing symptoms over 20 years., Developmental psychology 41 (5) (2005) 733.

[28] F. A. Rogosch, A. Oshri, D. Cicchetti, From child maltreatment to adolescent cannabis abuse and dependence: A developmental cascade model, Development and psychopathology 22 (4) (2010) 883–897.

[29] D. Rogosa, D. Brandt, M. Zimowski, A growth curve approach to the measurement of change., Psychological bulletin 92 (3) (1982) 726, 01525 Publisher: American Psychological Association.

[30] T. E. Duncan, S. C. Duncan, The ABC's of LGM: An Introductory Guide to Latent Variable Growth Curve Modeling, Social and Personality Psychology Compass 3 (6) (2009) 979–991, 00000 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1751-9004.2009.00224.x. doi:10.1111/j.1751-9004.2009.00224.x.
URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1751-9004.2009.00224.x

[31] K. M. King, A. K. Littlefield, C. J. McCabe, K. L. Mills, J. Flournoy, L. Chassin, Longitudinal modeling in developmental neuroimaging research: Common challenges, and solutions from developmental psychology, Developmental Cognitive Neuroscience 33 (2018) 54–72, 00044. doi:10.1016/j.dcn.2017.11.009.
URL http://www.sciencedirect.com/science/article/pii/S1878929317300300

[32] A. K. Littlefield, K. M. King, S. F. Acuff, K. T. Foster, J. G. Murphy, K. Witkiewitz, Limitations of cross-lagged panel models in addiction research and alternative models: An empirical example using project MATCH., Psychology of Addictive Behaviors00001 Publisher: US: American Psychological Association (2021). doi:10.1037/adb0000750.
URL https://psycnet.apa.org/fulltext/2021-47508-001.pdf

[33] P. J. Curran, A. L. Howard, S. Bainter, S. T. Lane, J. S. McGinley, The Separation of Between-person and Within-person Components of Individual Change Over Time: A Latent Curve Model with Structured Residuals, J Consult Clin Psychol 82 (5) (2014) 879–894, 00194. doi:10.1037/a0035297.
URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4067471/

[34] W. Bleidorn, T. Schwaba, A. Zheng, C. J. Hopwood, S. S. Sosa, B. W. Roberts, D. Briley, Personality stability and change: A meta-analysis of longitudinal studies., Psychological bulletin 148 (7-8) (2022) 588.

[35] B. W. Roberts, W. F. DelVecchio, The rank-order consistency of personality traits from childhood to old age: a quantitative review of longitudinal studies., Psychological bulletin 126 (1) (2000) 3.

[36] B. W. Roberts, K. E. Walton, W. Viechtbauer, Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies., Psychological bulletin 132 (1) (2006) 1.

[37] P. J. Curran, D. J. Bauer, The disaggregation of within-person and between-person effects in longitudinal models of change, Annual review of psychology 62 (2011) 583–619, 01308 Publisher: Annual Reviews.

[38] L. Hoffman, Longitudinal analysis: Modeling within-person fluctuation and change, Routledge, 2015, 00524.

[39] U. Orth, D. A. Clark, M. B. Donnellan, R. W. Robins, Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models., Journal of Personality and Social Psychology 120 (4) (2021) 1013, 00047 Publisher: American Psychological Association.

[40] E. L. Hamaker, R. M. Kuiper, R. P. P. P. Grasman, A critique of the cross-lagged panel model, Psychological Methods 20 (1) (2015) 102–116, 00848 Place: US Publisher: American Psychological Association. doi:10.1037/a0038889.

[41] W. K. Thompson, G. N. Savla, I. V. Vahia, C. A. Depp, R. O'Hara, D. V. Jeste, B. W. Palmer, Characterizing trajectories of cognitive functioning in older adults with schizophrenia: does method matter?, Schizophrenia research 143 (1) (2013) 90–96.

[42] J. O. Ramsay, B. W. Silverman, Applied functional data analysis: methods and case studies, Springer, 2002.

[43] S. T. Lenz, Alan Agresti (2013): Categorical data analysis, Statistical Papers 57 (3) (2016) 849, 00002 Publisher: Springer Nature BV.

[44] L. K. Muthén, Mplus user's guide. Los Angeles: Muthén & Muthén; 1998, 00000 (2017).

[45] J. Ren, S. Tapert, C. C. Fan, W. K. Thompson, A semi-parametric bayesian model for semi-continuous longitudinal data, Statistics in Medicine 41 (13) (2022) 2354–2374.

[46] T. D. Little, The Oxford Handbook of Quantitative Methods, Vol. 2: Statistical Analysis, Oxford University Press, 2013, 00006 Google-Books-ID: bCFwAgAAQBAJ.

[47] Y. Liu, R. E. Millsap, S. G. West, J.-Y. Tein, R. Tanaka, K. J. Grimm, Testing measurement invariance in longitudinal data with ordered-categorical measures., Psychological methods 22 (3) (2017) 486.

[48] R. Van De Schoot, P. Schmidt, A. De Beuckelaer, K. Lek, M. Zondervan-Zwijnenburg, Measurement invariance (2015).

[49] M. T. Willoughby, R. Wirth, C. B. Blair, Executive function in early childhood: longitudinal measurement invariance and developmental change., Psychological assessment 24 (2) (2012) 418.

[50] S. J. Brislin, D. A. Clark, D. B. Clark, C. E. Durbin, A. C. Parr, L. Ahonen, K. D. Anderson-Carpenter, M. M. Heitzeg, B. Luna, C. Sripada, et al., Differential item functioning in reports of delinquent behavior between black and white youth: evidence of measurement bias in self-reports of arrest in the adolescent brain cognitive development study, Assessment (2023) 10731911231164627.

[51] T. A. Salthouse, Why are there different age relations in cross-sectional and longitudinal comparisons of cognitive functioning?, Current directions in psychological science 23 (4) (2014) 252–256, 00069 Publisher: Sage Publications Sage CA: Los Angeles, CA.

[52] E. V. Sullivan, T. Brumback, S. F. Tapert, D. Prouty, R. Fama, W. K. Thompson, S. A. Brown, K. Cummins, I. M. Colrain, F. C. Baker, Effects of prior testing lasting a full year in NCANDA adolescents: contributions from age, sex, socioeconomic status, ethnicity, site, family history of alcohol or drug abuse, and baseline performance, Developmental cognitive neuroscience 24 (2017) 72–83, 00008 Publisher: Elsevier.

[53] A. T. Beck, C. H. Ward, M. Mendelson, J. Mock, J. Erbaugh, An inventory for measuring depression, Archives of general psychiatry 4 (6) (1961) 561–571, 45810 Pub-

lisher: American Medical Association.

[54] D. P. French, S. Sutton, Reactivity of measurement in health psychology: how much of a problem is it? What can be done about it?, British journal of health psychology 15 (3) (2010) 453–468, 00185 Publisher: Wiley Online Library.

[55] L. Robins, Epidemiology: Reflections on Testing the Validity of Psychiatric Interviews | JAMA Psychiatry | JAMA Network, 00479 (1985).
URL https://jamanetwork.com/journals/jamapsychiatry/article-abstract/493658

[56] P. E. Shrout, G. Stadler, S. P. Lane, M. J. McClure, G. L. Jackson, F. D. Clavél, M. Iida, M. E. Gleason, J. H. Xu, N. Bolger, Initial elevation bias in subjective reports, Proceedings of the National Academy of Sciences 115 (1) (2018) E15–E23.

[57] J. Durbin, G. S. Watson, Testing for serial correlation in least squares regression: I, Biometrika 37 (3/4) (1950) 409–428, 03276 Publisher: JSTOR.

[58] C. Kincaid, Guidelines for selecting the covariance structure in mixed model analysis, paper 198-30 in Proceedings of the Thirtieth Annual SAS Users Group Conference, Inc., Cary, North Carolina00013 (2005).

[59] A. S. Bryk, S. W. Raudenbush, Hierarchical linear models: applications and data analysis methods., Sage Publications, Inc, 1992.

[60] J. D. Singer, Using sas proc mixed to fit multilevel models, hierarchical models, and individual growth models, Journal of educational and behavioral statistics 23 (4) (1998) 323–355.

[61] P. J. Curran, Have multilevel models been structural equation models all along?, Multivariate behavioral research 38 (4) (2003) 529–569.

[62] A. K. Littlefield, K. M. King, S. F. Acuff, K. T. Foster, J. G. Murphy, K. Witkiewitz, Limitations of cross-lagged panel models in addiction research and alternative models: An empirical example using project match., Psychology of Addictive Behaviors 36 (3) (2022) 271.

[63] S. W. F. Ewing, G. F. Dash, W. K. Thompson, C. Reuter, V. G. Diaz, A. Anokhin, L. Chang, L. B. Cottler, G. J. Dowling, K. LeBlanc, et al., Measuring retention within the adolescent brain cognitive development (abcd) sm study, Developmental Cognitive Neuroscience 54 (2022) 101081.

[64] R. B. Cotter, J. D. Burke, M. Stouthamer-Loeber, R. Loeber, Contacting participants for follow-up: how much effort is required to retain participants in longitudinal studies?, Evaluation and Program Planning 28 (1) (2005) 15–21, 00105 Publisher: Elsevier.

[65] K. G. Hill, D. Woodward, T. Woelfel, J. D. Hawkins, S. Green, Planning for long-term follow-up: Strategies learned from longitudinal studies, Prevention Science 17 (7) (2016) 806–818, 00036 Publisher: Springer.

[66] N. Watson, E. Leissou, H. Guyer, M. Wooden, Best Practices for Panel Maintenance and Retention, in: Advances in Comparative Survey Methods, John Wiley & Sons, Ltd, 2018, pp. 597–622, 00006 Section: 29 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118884997.ch29. doi:10.1002/9781118884997.ch29.
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118884997.ch29

[67] R. J. Little, D. B. Rubin, The Analysis of Social Science Data with Missing Values, Sociological Methods & Re-

search 18 (2-3) (1989) 292–326, 01328 Publisher: SAGE Publications Inc. doi:10.1177/0049124189018002004.
URL https://doi.org/10.1177/0049124189018002004

[68] J. W. Graham, Missing Data Analysis: Making It Work in the Real World, Annual Review of Psychology 60 (1) (2009) 549–576, 05760 _eprint: https://doi.org/10.1146/annurev.psych.58.110405.085530. doi:10.1146/annurev.psych.58.110405.085530.
URL https://doi.org/10.1146/annurev.psych.58.110405.085530

[69] C. K. Enders, Applied Missing Data Analysis, Guilford Press, 2010, 06821 Google-Books-ID: MN8ruJd2tvgC.

[70] H. C. Kraemer, Effect size, The encyclopedia of clinical psychology (2014) 1–300178 Publisher: Wiley Online Library.

[71] J. Cohen, Statistical power, Analysis for the behavioral sciences (1988) 273–40603369.

[72] S. B. Morris, R. P. DeShon, Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs., Psychological methods 7 (1) (2002) 105, 02272 Publisher: American Psychological Association.

[73] L. Wang, Q. Zhang, S. E. Maxwell, C. Bergeman, On standardizing within-person effects: Potential problems of global standardization, Multivariate Behavioral Research 54 (3) (2019) 382–403.

[74] A. Feingold, Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis., Psychological methods 14 (1) (2009) 43, 00633 Publisher: American Psychological Association.

[75] A. Feingold, Time-varying effect sizes for quadratic growth models in multilevel and latent growth modeling, Structural equation modeling: a multidisciplinary journal 26 (3) (2019) 418–429, 00008 Publisher: Taylor & Francis.

[76] L. M. Collins, Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model, Annu. Rev. Psychol. 57 (2006) 505–528.

[77] D. J. Bauer, P. J. Curran, Conducting longitudinal data analysis: Knowing what to do and learning how to do it. (2019).
URL https://centerstat.org/wp-content/uploads/2019/03/SRCD2019-2up.pdf

[78] L. Castro-Schilo, K. J. Grimm, Using residualized change versus difference scores for longitudinal research, Journal of Social and Personal Relationships 35 (1) (2018) 32–58, 00161 Publisher: SAGE Publications Ltd. doi:10.1177/0265407517718387.
URL https://doi.org/10.1177/0265407517718387

[79] A. E. Brieant, L. M. Sisk, D. G. Gee, Associations among negative life events, changes in corticolimbic connectivity, and psychopathology in the abcd study, Developmental Cognitive Neuroscience 52 (2021) 101022.

[80] A. L. Romer, D. A. Pizzagalli, Is executive dysfunction a risk marker or consequence of psychopathology? a test of executive function as a prospective predictor and outcome of general psychopathology in the adolescent brain cognitive development study®, Developmental Cognitive Neuroscience 51 (2021) 100994.

[81] A. K. Littlefield, Longitudinal data analysis, in: H. E. Cooper, M. N. Coutanche, L. M. McMullen, A. T. Panter, D. E. Rindskopf, K. J. Sher (Eds.), APA handbook of research methods in psychology: Data analysis and

research publication, Vol. 3, American Psychological Association, 2023, pp. 245–268.

[82] M. T. Van Dijk, E. Murphy, J. E. Posner, A. Talati, M. M. Weissman, Association of multigenerational family history of depression with lifetime depressive and other psychiatric disorders in children: Results from the adolescent brain cognitive development (abcd) study, JAMA psychiatry 78 (7) (2021) 778–787.

[83] M. E. Martz, M. M. Heitzeg, K. M. Lisdahl, C. C. Cloak, S. W. F. Ewing, R. Gonzalez, F. Haist, K. H. LeBlanc, P. A. Madden, J. M. Ross, et al., Individual-, peer-, and parent-level substance use-related factors among 9-and 10-year-olds from the abcd study: prevalence rates and sociodemographic differences, Drug and alcohol dependence reports 3 (2022) 100037.

[84] B. Lees, L. M. Squeglia, F. J. Breslin, W. K. Thompson, S. F. Tapert, M. P. Paulus, Screen media activity does not displace other recreational activities among 9–10 year-old youth: A cross-sectional abcd study®, BMC public health 20 (1) (2020) 1–11.

[85] C. E. Palmer, D. Pecheva, J. R. Iversen, D. J. Hagler Jr, L. Sugrue, P. Nedelec, C. C. Fan, W. K. Thompson, T. L. Jernigan, A. M. Dale, Microstructural development from 9 to 14 years: Evidence from the abcd study, Developmental cognitive neuroscience 53 (2022) 101044.

[86] P. J. Curran, G. R. Hancock, The challenge of modeling co-developmental processes over time, Child development perspectives 15 (2) (2021) 67–75.

[87] Q. Xiang, K. Chen, L. Peng, J. Luo, J. Jiang, Y. Chen, L. Lan, H. Song, X. Zhou, Prediction of the trajectories of depressive symptoms among children in the adolescent brain cognitive development (abcd) study using machine learning approach, Journal of Affective Disorders 310 (2022) 162–171.

[88] D. McNeish, J. R. Harring, Improving convergence in growth mixture models without covariance structure constraints, Statistical Methods in Medical Research 30 (4) (2021) 994–1012.

[89] K. J. Sher, K. M. Jackson, D. Steinley, Alcohol use trajectories and the ubiquitous cat's cradle: cause for concern?, Journal of abnormal psychology 120 (2) (2011) 322.

[90] S. W. Hawes, S. B. Perlman, A. L. Byrd, A. Raine, R. Loeber, D. A. Pardini, Chronic anger as a precursor to adult antisocial personality features: The moderating influence of cognitive control., Journal of Abnormal Psychology 125 (1) (2016) 64.

[91] S. W. Hawes, A. L. Byrd, R. Gonzalez, C. Cavanagh, J. Bechtold, D. R. Lynam, D. A. Pardini, The developmental course of psychopathic features: Investigating stability, change, and long-term outcomes, Journal of Research in Personality 77 (2018) 83–89.

[92] A. K. Littlefield, K. J. Sher, D. Steinley, Developmental trajectories of impulsivity and their association with alcohol use and related outcomes during emerging and young adulthood i, Alcoholism: Clinical and Experimental Research 34 (8) (2010) 1409–1416.

[93] L. K. Kulisch, K. Arumäe, D. A. Briley, U. Vainik, Triangulating causality between childhood obesity and neurobehavior: Behavioral genetic and longitudinal evidence, Developmental Science (2023) e13392.