

# MODELING LONGITUDINAL MEDIATION OF WITHIN-PERSON CHANGES

Corinne M. Henk

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Psychology.

Chapel Hill  
2016

Approved by:

Laura Castro-Schilo

Daniel J. Bauer

Patrick J. Curran

© 2016  
Corinne M. Henk  
ALL RIGHTS RESERVED

## **ABSTRACT**

Corinne M. Henk: Longitudinal Mediation of Within-Person Changes  
(Under the direction of Laura Castro-Schilo)

Mediation analysis concerns the discovery of mechanisms that transmit an effect from one variable to another. At times, social scientists propose hypotheses involving mediation of within-person changes (e.g., do within-person changes in cognitive ability mediate the relation between changes in depression and changes in disability?) The current manuscript introduces an approach within the latent change score (LCS) modeling framework to test such hypotheses. This approach is compared to existing methods, including the cross-lagged panel model (CLPM; Cole & Maxwell, 2003) and the LCS mediation model (Selig & Preacher, 2009). Results from (1) a single-replication simulation and (2) an empirical example utilizing data from the Health and Retirement study are presented.

## TABLE OF CONTENTS

LIST OF TABLES.....	vi
LIST OF FIGURES.....	vii
CHAPTER 1: INTRODUCTION.....	1
The Cross-Lagged Panel Model.....	5
Advantages of the CLPM.....	8
The Latent Change Score Mediation Model.....	13
The Cross-Lagged Latent Change Score Model.....	17
Relations Among Structural Parameters in the CLPM and CL-LCS Model.....	20
Limitations of the CLPM.....	21
CHAPTER 2: METHOD.....	23
Proof of Concept.....	23
Participants and Procedures.....	24
Measures.....	25
Data Analysis.....	27
CHAPTER 3: RESULTS.....	30
Proof of Concept.....	30
Illustration: Descriptive Statistics.....	31
Illustration: CL-LCS Model Results.....	32
CHAPTER 4: DISCUSSION.....	35

APPENDIX A: R CODE FOR SIMULATING CL-LCS DATA.....	58
APPENDIX B: EXAMPLE <i>MPLUS</i> CODE FOR FITTING CL-LCS MODEL.....	62
ENDNOTES.....	65
REFERENCES.....	66

## **LIST OF TABLES**

Table 1. Unstandardized Results: CL-LCS Model, No Linear Trend Simulation.....	40
Table 2. Unstandardized Results: CL-LCS Model, Linear Trend Simulation.....	42
Table 3. Descriptive Statistics: Indicators of Depression, Cognitive Ability, and Disability.....	44
Table 4. Results: Final Selected Model.....	45

## LIST OF FIGURES

Figure 1. Path diagram of one specification of Cole and Maxwell's (2003) CLPM with latent variables.....	47
Figure 2(a). Causal steps cross-sectional model. ....	48
Figure 2(b). Causal steps model with temporal separation.....	48
Figure 3(a). Basic LCS building block with the initial level-to-change association specified as a regression.....	49
Figure 3(b). Basic LCS building block with the initial level-to-change association specified as a correlation. ....	49
Figure 4. Path diagram of Selig and Preacher's (2009) LCS-MM.....	50
Figure 5. Path diagram of a fully mediated process using the CL-LCS.....	51
Figure 6. Conceptual diagram of the CL-LCS in the case of full mediation.....	52
Figure 7. Random selection of 100 simulated trajectories governed by no trend across 4 waves.....	53
Figure 8. Random selection of 100 simulated trajectories governed by a linear trend across 4 waves. ....	54
Figure 9. Individual trajectories of a random sample of 100 HRS participants' factor scores on depression over time. ....	55
Figure 10. Individual trajectories of a random sample of 100 HRS participants' factor scores on cognitive ability over time. ....	56
Figure 11. Individual trajectories of a random sample of 100 HRS participants' factor scores on disability over time. ....	57

## CHAPTER 1: INTRODUCTION

Mediation analyses are ubiquitous in the social sciences; at the time of writing, the term *mediation* generated 1.43 million results in a Google scholar search. Increasingly, psychologists seek to understand not just *whether* two constructs are related, but *why* the relation exists. While cross-sectional data were once considered sufficient for conducting sound mediation analyses, researchers now recognize that longitudinal data are preferable for gleaning insight into unfolding mediation processes. Indeed, repeated measures afford researchers the opportunity to draw stronger inferences about the nature of relations over time (e.g., Shadish, Cook, & Campbell, 2002).

When working with longitudinal data, hypotheses about change predicting future change often emerge (Baltes & Nesselroade, 1979). As an example, researchers hypothesize that *increases* in lateral ventricle size predict subsequent *declines* in memory performance among older adults (Grimm, An, McArdle, Zonderman, & Resnick, 2012). Research questions involving associations among changes are central to longitudinal research; in fact, Grimm and colleagues (2012, p. 268) posit that a key goal of developmental science is to “[understand] the dynamic interplay between two (or more) constantly changing constructs.” Hypotheses involving change—as a predictor, mediator, and/or outcome—are becoming more common in psychological research. Indeed, researchers interested in modeling change as both a predictor and an outcome will find that the “change-to-change” literature is expanding (e.g., Grimm et al., 2012; Henk & Castro-Schilo, 2015; Selig & Preacher, 2009; Smith et al., 2013).

Importantly, these change-to-change hypotheses are qualitatively and quantitatively distinct from other types of research questions, such as those that involve the level or status on a variable as a predictor of change in another variable (i.e., level-to-change hypotheses). To clarify the qualitative differences, consider two examples. First, a level-to-change hypothesis may posit that older adults with the *highest levels* of depression will show the *most declines* in cognitive ability later in life. In contrast, a change-to-change hypothesis may be that as individuals become *more depressed* (increase on depression), they will show either *gains* or less rapid *declines* in cognitive ability. In this case, change is the mechanism by which the relation operates. One does not have to search far and wide in psychology to find unique examples of level-to-change and change-to-change hypotheses. Each type of hypothesis demands a quantitative strategy that will appropriately test the associations of interest, yet strategizing is not always straightforward.

Compounding the task of selecting an optimal modeling strategy is the issue of whether the motivating hypothesis entails tests of within-person effects, between-person effects, or both (e.g., Curran & Bauer, 2011). Somewhat vexingly, level-to-change and change-to-change hypotheses can be focused on any combination of within- or between-person associations. A researcher's analytic choices determine which types of effects are tested. For example, consider again the level-to-change hypothesis that individuals who are more depressed have more drastic declines in cognitive ability. If interest lies in the between-person effect, one appropriate modeling strategy could be to use a parallel process latent growth model (Cheong, MacKinnon, & Khoo, 2003), with the intercept of depression predicting the slope of cognitive ability. Similarly, if interest lies in the within-person effect when considering the change-to-change hypothesis that changes in depression are linked to changes in cognitive ability, a parallel process growth model could be parameterized with a regression among the slopes. Using the

means of the slopes to aid interpretation, a researcher could conclude that on average, increases in depression are associated with decreases in cognitive ability (assuming the means were positive and negative for depression and cognitive ability, respectively). Results would inform the degree to which relations held between-persons.

However, when interest lies in within-person change-to-change effects, a correlation or regression among parallel process slopes will not be informative, and it becomes necessary to consider alternate modeling strategies. Researchers may formulate a within-person change-to-change research question to understand, for example, whether a particular individual who has decreasing severity of depression will subsequently experience changes in her cognitive ability. These hypotheses have major implications and ramifications for individual-targeted interventions and individualized medicine, where interest lies in whether changes in one variable will produce changes in another variable *at the individual level*. A key goal of the current work is to clarify best practices for testing hypotheses regarding within-person change-to-change relations, as these types of hypotheses have received relatively little attention in the quantitative psychology literature.

Although many mediation hypotheses fundamentally posit within-person change as a predictor and an outcome, the most common approach for testing longitudinal mediation—the cross-lagged panel model (CLPM; Cole & Maxwell, 2003)—focuses on within-person level-to-change associations (Preacher, 2015). Unfortunately, within-person change-to-change hypotheses cannot be tested using the CLPM, as this model can only accommodate change as an outcome—not as a predictor. Testing within-person change-to-change hypotheses requires explicit specification of change in statistical models. Belonging to the structural equation modeling (SEM) family, the latent change score (LCS) framework is apt for investigating

patterns, causes, and consequences of intraindividual change (e.g., Castro-Schilo, Ferrer, Hernández, & Conger, 2015; Ferrer & McArdle, 2010; Grimm, Castro-Schilo, & Davoudzadeh, 2013; McArdle, 1994; McArdle, 2001; McArdle, 2009; McArdle & Hamagami, 2001, McArdle & Nesselroade, 1994). In this framework, within-person changes are explicitly represented as latent variables, a feature that enables change to be specified as a predictor, criterion, or both. Selig and Preacher (2009) capitalized on the LCS framework and introduced the LCS mediation model (LCS-MM) for testing mediational hypotheses rooted in within-person changes. However, a major setback of the LCS-MM is that its setup precludes testing competing patterns of causality.

To illustrate this point and to foreshadow an empirical example that will be presented later in this project, consider a hypothesis positing that for an aging adult, the relation between his/her changes in depression (X) and changes in disability (Y) is mediated by changes in cognitive ability (M). A competing causal hypothesis would consist of switching the roles of these variables, treating changes in disability as the exogenous predictor (X) and changes in depression as the outcome (Y), yet testing this hypothesis with the LCS-MM would result in two non-nested models that are not easily comparable. Another competing hypothesis could include bidirectional effects (e.g., changes in depression and changes in disability having reciprocal relations)—but such effects cannot be specified in the LCS-MM. As such, researchers relying on the LCS-MM may inadvertently engage in confirmation bias because this approach is designed to test only the a priori hypothesis. Thus, the goal of this project is to introduce a model, which I term the cross-lagged LCS (CL-LCS) model, that allows for more rigorous tests of within-person change-to-change mediation.

The remainder of this paper is organized as follows: First, I describe the CLPM in detail

and explicate its key advantages relative to cross-sectional approaches to assessing mediation. I then present the LCS-MM and CL-LCS models, and compare their appropriateness for assessing longitudinal mediation among intraindividual changes. I show analytically that results and significance tests from the CL-LCS model will necessarily differ from the CLPM approach. I also enumerate some of the limitations of the CLPM approach and provide further rationale for using the CL-LCS model in applications in which data exhibit systematic mean changes over time. Next, I present a small simulation study as a proof of concept; that is, when data are generated to follow a CL-LCS process, the model appears identified and can be estimated with reasonable parameter estimates. Then, using an empirical example of data from the Health and Retirement Study (HRS), I fit the CL-LCS model to test the hypothesis that within-person changes in depression lead to changes in cognitive ability, which in turn lead to changes in disability. Henk and Castro-Schilo (2015) found support for this hypothesis using just two waves of data; here I test the hypothesis using four waves of data, which provides potential for stronger internal validity. I conclude by enumerating strengths and limitations of the suggested CL-LCS approach and provide future directions for research on this newly developed model.

### **The Cross-Lagged Panel Model**

The CLPM can be specified with observed or latent variables in the SEM framework; here, I consider the latter as doing so will facilitate discussion of other models that rely on latent variables. Figure 1 depicts a specification of the CLPM with three constructs,  $x$ ,  $m$ , and  $y$ . For simplicity, each construct is measured with just one indicator across four time-points (i.e.,  $X_1, \dots, X_4, M_1, \dots, M_4, Y_1, \dots, Y_4$ ), although multiple-indicators could be included to define common factors.

With at least four repeated measures, single-indicator latent variables can be specified by

fixing factor loadings to 1, such that measurement error can be estimated and parsed from true score variability (Jöreskog, 1978),

$$X_{it} = x_{it} + \delta_{xit} \quad (1)$$

$$x_{it} = X_{it} - \delta_{xit} \quad (2)$$

where  $X_{it}$  is the observed variable for individual  $i$  at time  $t$ ,  $x_{it}$  is the latent true score for individual  $i$  at time  $t$ , and  $\delta_{xit}$  is the measurement error for individual  $i$  at time  $t$  and is  $\sim N(0, \theta_x)$  and assumed to be uncorrelated with  $X_{it}$ . For identification, these error variances are fixed to equality across time (although with enough waves of data, some equality constraints in error variances may be relaxed; see Jöreskog, 1978). Equations (1) and (2) operate identically for the  $m$  and  $y$  process. The latent variables at the first wave of assessment are exogenous and freely allowed to correlate. All other latent variables (i.e., after Wave 1) are described by an autoregressive (AR) process such that a given variable is a function of its true score at the previous time-point, plus a disturbance:

$$x_{it} = \beta_x (x_{it-1}) + \xi_{xit} \quad \text{for } t > 1 \quad (3)$$

where  $\beta_x$  is the AR parameter quantifying the rank-order stability over time for  $x_{it}$ ,  $x_{it-1}$  is individual  $i$ 's latent factor score at the previous point in time, and  $\xi_{xit}$  is a time-specific residual assumed to be  $\sim N(0, \Theta_x)$  and uncorrelated with the initial score  $x_{it-1}$ , and all other terms are as defined above. In Figure 1, the  $m$  process is defined to be a function of itself and  $x$  at the prior time-point,

$$m_{it} = \beta_m (m_{it-1}) + \beta_{mx} (x_{it-1}) + \xi_{mit} \quad \text{for } t > 1 \quad (4)$$

where  $\beta_m$  is the AR parameter,  $m_{it-1}$  is individual  $i$ 's latent factor score at the previous point in time,  $\beta_{mx}$  is the cross-lagged effect capturing the effect of  $x$  on  $m$ , and  $\xi_{mit}$  is a time-specific residual assumed to be  $\sim N(0, \Theta_m)$  and uncorrelated with the initial score  $m_{it-1}$  as well as  $x_{it-1}$ .

Similarly, the  $y$  process is governed by an AR and cross-lagged process:

$$y_{it} = \beta_y (y_{it-1}) + \beta_{ym} (m_{it-1}) + \xi_{yit} \quad \text{for } t > 1 \quad (5)$$

where  $\beta_y$  is the AR parameter,  $y_{it-1}$  is individual  $i$ 's latent factor score at the previous point in time,  $\beta_{ym}$  is the cross-lagged parameter, and  $\xi_{yit}$  is a time-specific residual assumed to be  $\sim N(0, \Theta_y)$  and uncorrelated with the initial score  $y_{it-1}$ ,  $m_{it-1}$ , and  $x_{it-1}$ . All other terms are as defined above. Importantly, Figure 1 represents just one potential specification of the CLPM; other specifications are important for systematic assessment of mediation, such as the inclusion of direct effects from  $x$  to  $y$ , bidirectional effects, or competing models of causality (e.g.,  $y$  predicts  $m$ , which predicts  $x$ ). Moreover, it is worth noting that Equations 3-5 do not include an intercept because the CLPM assumes data are centered; that is, no stable between-person differences exist (Hamaker, Kuiper, & Grasman, 2015).

Interpretation of the AR parameters (e.g.,  $\beta_x$  and  $\beta_y$  in Figure 1) is straightforward; these represent rank-order stability in the construct over time. An AR coefficient of 1 indicates the construct under study is perfectly stable over time. Here, the term *stable* refers to a lack of interindividual differences in within-person change; either no one is changing, or every individual experiences the same amount of change with respect to magnitude and direction. An AR parameter estimated at 1 suggests the rank ordering of individuals remains unchanged over time (e.g., Person A has the highest depression at all waves of assessment; Person B has the lowest scores on depression at all waves of assessment), assuming the disturbances do not disrupt the rankings. Generally, however, AR coefficients are estimated at some value less than 1.

The cross-lagged coefficients (e.g.,  $\beta_{mx}$  and  $\beta_{ym}$  in Figure 1), also called coupling parameters, are interpreted as level-to-change effects. Strictly speaking,  $\beta_{ym}$  represents the

expected shift in  $y_{it}$  given a 1-unit increase in  $m_{it-1}$ , *controlling for the prior level of  $y$*  (i.e.,  $y_{it-1}$ ). Because the effect of  $y_{i1}$  on  $y_{i2}$  is partialled out, the coupling parameter can be interpreted as a prediction of change in  $y$ . Thus, the CLPM is equipped to model change as an outcome, but not as a predictor. At best, relations among changes in different constructs can be captured in the CLPM through estimation of within-time correlations among residual variances, although this specification would not be informative of structured (e.g., lead-lag) relations among changes. In sum, specification of the CLPM is most informative for hypotheses about lead-lag relations among the level (or status) of variables, but cannot speak to change-to-change hypotheses.

### **Advantages of the CLPM**

The CLPM is widely used for evaluating longitudinal mediation in the social sciences; according to a Google Scholar search, it had been cited over 1,200 times at the time of writing this paper. Given that this model was explicitly developed to address shortcomings of the cross-sectional causal steps approach (Baron & Kenny, 1986), a brief description of that model follows. Then, the advantages of the CLPM are discussed in detail.

The causal steps approach involves testing three regression models: (1) the regression of the dependent variable ( $Y$ ) on the independent variable ( $X$ )<sup>1</sup>; (2) the regression of the mediator ( $M$ ) on  $X$ ; and (3) the multiple regression of  $Y$  on both  $X$  and  $M$ . If a mediation process is underway, the first two models will yield significant regression coefficients, signaling that  $X$  has a direct effect on  $Y$ , and that  $X$  has a direct effect on  $M$ . Moreover, the effect of  $X$  on  $Y$  should diminish once the effect of  $M$  is accounted for (i.e., when comparing the relevant coefficients from the first and third regression models); this is termed partial mediation. In the case of full mediation, the effect of  $X$  on  $Y$  will drop to zero when  $M$  is accounted for.

Notably, the causal steps approach is often used for testing cross-sectional mediation

among variables. However, Sobel (1990) argued that for true mediation to hold, M “must truly be a dependent variable relative to X, which implies that X must precede [M] in time; and [M] must be a truly independent variable relative to Y, implying that [M] precedes Y in time” (Cole & Maxwell, 2003, p. 561). As such, researchers have emphasized the importance of using longitudinal data to conduct more rigorous tests of mediation. Indeed, Cole and Maxwell (2003) made a critical contribution to this literature when they expanded the causal steps approach with the cross-lagged panel model (CLPM), which can accommodate tests of longitudinal mediation.

There are four key advantages of the CLPM as compared to the causal steps approach. (1) Accounting for prior values of endogenous variables leads to unbiased estimates of cross-lagged effects. (2) The CLPM allows estimation and testing of overall indirect effects. (3) Alternate lagged effects may be tested. (4) Alternate models of causal direction can be tested (e.g., perhaps Y has an effect on X that is mediated by M, rather than X affecting Y through M). In this section, I will explain each of these advantages in detail.

The key difference between the CLPM and Baron and Kenny’s (1986) causal steps approach of testing mediation is that the former accounts for prior values of each endogenous variable. Cole and Maxwell (2003, p. 560) call these prior values an “almost ubiquitous third variable confound” because, as is well known within the context of the general linear modeling (GLM) framework, regression coefficients will be biased if important predictors are omitted from the model. The reason is that these coefficients are partial regression weights; thus, their values change depending on the other predictors in the model and the degree to which those predictors correlate (i.e., share variance) with each other. Two variables measured at the same time-point (e.g., X1 and M1) are likely to be correlated, if for no other reason than by virtue of being assessed at the same occasion. Thus, to obtain an unbiased estimate of the cross-lagged

effect of X1 to M2, M1 must be included as a covariate. Failing to control for prior values of an endogenous variable—as in the causal steps approach—will almost always lead to inflated estimates of the direct and indirect effects. Figure 2 displays two models that could be assessed with the causal steps approach for testing mediation: (a) a cross-sectional mediation model and (b) a mediation model with temporal separation. Notably, although Figure 2b incorporates longitudinal data, it too will lead to biased estimates of  $\beta'_{mx}$  and  $\beta'_{ym}$ . Temporal separation of the independent variable, mediator, and outcome is not sufficient for obtaining accurate coefficients; it is critical to control for prior values of each endogenous variable (Cole & Maxwell, 2003).

In addition to controlling for prior values of endogenous variables—thus yielding more accurate estimates—the CLPM includes a simultaneous test of overall indirect effects, whereas the causal steps approach tests each time-specific indirect effect. The overall indirect effect captures the total extent to which the independent variable transmits an effect on the outcome through the mediator, across the entire duration of the study. Consider the specification of Figure 1. The overall indirect effect that  $x$  has on  $y$  is equal to the sum of all indirect effects across waves:

$$(a) x_1 \rightarrow x_2 \rightarrow m_3 \rightarrow y_4 = (\beta_x \cdot \beta_{mx} \cdot \beta_{ym})$$

$$(b) x_1 \rightarrow m_2 \rightarrow m_3 \rightarrow y_4 = (\beta_{mx} \cdot \beta_m \cdot \beta_{ym})$$

$$(c) x_1 \rightarrow m_2 \rightarrow y_3 \rightarrow y_4 = (\beta_{mx} \cdot \beta_{ym} \cdot \beta_y)$$

$$\begin{aligned} \text{Overall Indirect Effect} &= \beta_x \cdot \beta_{mx} \cdot \beta_{ym} + \beta_{mx} \cdot \beta_m \cdot \beta_{ym} + \beta_{mx} \cdot \beta_{ym} \cdot \beta_y \\ &= \beta_{mx}(\beta_x \cdot \beta_{ym} + \beta_m \cdot \beta_{ym} + \beta_{ym} \cdot \beta_y) \\ &= \beta_{mx}(\beta_{ym}(\beta_x + \beta_m + \beta_y)) \end{aligned} \tag{6}$$

Note that any single indirect effect will not be representative of the overall indirect effect.<sup>2</sup> This

point is particularly important because mediation is usually conceived of as an ongoing, unfolding process as opposed to a one-time occurrence; thus, differentiating time-specific indirect effects from the overall indirect effect is key (Cole & Maxwell, 2003). Moreover, if a researcher has gone through the process of collecting longitudinal data, it behooves him or her to make full use of every wave of data. Ultimately, more waves of data provide a more accurate estimate of the overall indirect effect.

The third advantage of the CLPM is that it allows for tests of alternate autoregressive relations. Figures 1 and 2 only have lag-1 relations (i.e., each construct is predicted by itself at the prior time-point), but in some research settings it may be appropriate to include lag-2 paths. Cole and Maxwell (2003, p. 571) suggest that these “wave-skipping paths” may indicate “interesting nonlinear relations: The system may not be stationary, causal relations may be accelerating or decelerating, or the selected time lag between waves might not be optimal to represent the full causal effect of one variable on another.” Testing the need for additional lagged effects can be done using a series of likelihood ratio tests (LRT) as the reduced set (e.g., lag-1) is nested within the larger model. If there is a significant increment in fit when additional lags are included, these paths should be retained.

Lastly, and perhaps most importantly, the CLPM allows for tests of alternate patterns of causality, which is a critically important step in SEM underscored by MacCallum, Wegender, Uchino, and Fabrigar (1993). In their seminal paper, the notion of equivalent models was introduced. That is, there may exist alternative models that are mathematically equivalent (i.e., provide identical model fit) to the proposed/hypothesized model. The authors conducted a literature review and showed that many applications of SEM neglect to consider alternate, equivalent models. As a result, MacCallum and colleagues concluded “the existence of any (let

alone many) equivalent models presents a serious challenge to the inferences typically made by researchers using [SEM]” (p. 196). Potential solutions of this serious issue include experimentally manipulating key variables and/or using longitudinal designs. Both of these approaches allow researchers to safely rule out at least some alternative models. In the case of longitudinal designs, equivalent models that involve effects moving backward in time (e.g.,  $X$  at Time 2 predicting  $M$  at Time 1) are not plausible and can thus be ruled out.

Most applications of the causal steps approach involve testing mediation with cross-sectional data (see Figure 2a), which precludes testing the directionality of effects. Theory may suggest a model wherein  $x$  predicts  $m$ , which then predicts  $y$ , but the “true” model may be such that  $y$  is exogenous with respect to  $m$ , and  $m$  is exogenous with respect to  $x$ . These two models are indistinguishable using the causal steps method,<sup>3</sup> but can be tested in the CLPM. Namely, LRTs can be conducted to distinguish the directionality of effects based on model fit. A full CLPM can be specified to include bidirectional effects, such that (a)  $x$  predicts  $m$ , which predicts  $y$ , and (b)  $y$  predicts  $m$ , which predicts  $x$ . Removing paths in part (a) or (b) allows for LRTs; if significant, the LRT suggests the fit of the model significantly worsens with the removal of these paths, and thus relevant effects should be retained in the model. Even in longitudinal applications of the causal steps approach, however, testing alternate patterns of causality is not straightforward, as each construct is represented at one and only one point in time (see Figure 2b). Thus, respecifying the model to test an alternate causal process (e.g.,  $Y$  predicting  $M$ , predicting  $X$ ) requires a different set of data wherein  $Y$  is measured at Time 1,  $M$  at Time 2, and  $X$  at Time 3.

Recall the motivating hypothesis that the relation between changes in depression and changes in disability is mediated by changes in cognitive health. At issue here are two

possibilities: equivalent models (i.e., that provide identical model fit) and alternate models that posit a different structure of causality but do not provide identical fit (i.e., Y predicts M, which predicts X, instead of the a priori hypothesis that X leads to M, which leads to Y). Fitting the CLPM is not a panacea for the problem of equivalent models; equivalent models will still be present. However, the setup of this model does allow researchers to test alternate structures of causality. In fact, Cole and Maxwell (2003) encouraged researchers to evaluate the tenability of different models of causality, and termed the pathways in these alternative models “theoretically backward effects.” To avoid confusion with backward-in-time effects, I use the term “alternative models of causality.” Ultimately, testing alternative models of causality helps researchers to avoid engaging in confirmation bias. To the extent that one alternate model of causality provides better fit to the data than the a priori model—and makes sense from a theoretical standpoint—it may be warranted to conduct future studies and/or revise the original hypothesis.

In sum, the CLPM offers many advantages over the traditional causal steps approach, and is held in high esteem for this reason. However, when a researcher’s motivating theory includes within-person change as a predictor of subsequent change, it is necessary to move to an alternate framework.

### **The Latent Change Score Mediation Model**

The LCS framework affords great flexibility in using intraindividual change as an outcome, predictor, or both. Before delving into the LCS-MM, I begin by presenting some basic latent change score concepts. First, to specify a latent change score, “true” scores are separated from their imperfectly measured observed variable counterparts, exactly as was done in Equations (1) and (2). Then, as implied in the following equation,

$$x_{it} = 1(x_{it-1}) + \Delta x_{it} \quad \text{for } t > 1 \quad (7)$$

creation of latent change scores relies on the imposition of unit-weighted autoregressive effects, which imply that a latent score (i.e., level) at time  $t$  is exactly equal to itself at time  $t - 1$  plus the change that occurred between time  $t - 1$  and  $t$ . (Note that the subscripts of the change scores begin at 2 because at least two waves of data must be collected before changes can be modeled). Contrast this expression with that of Equation (3), which expresses the “level” equation for  $x$  in the CLPM. At first glance, the LCS formulation may appear to be a subset of the CLPM with the autoregressive effect fixed to 1, yet these models are not nested. The “residual” in the LCS framework is the within-person change ( $\Delta x_{it}$ ) and is allowed to correlate with or be regressed on  $x_{it}$ . The rationale for estimating this correlation/regression is that in many research contexts, initial levels of a construct are associated with how much change takes place. In contrast, the residual in the CLPM (and in any traditional regression model)—by definition—is assumed to have zero correlation with the predictors (see Figure 3 for a path diagram of the latent change score basic building block that maps onto Equation 6).

Importantly, the imposition of perfect autoregressive effects allows for the explicit representation of within-person change as a latent variable, which lends itself to testing different structural relations among the changes. Indeed, a variety of structures can be imposed on the latent change scores (see McArdle, 2009 for a description of common LCS model specifications). Most LCS specifications include level-to-change predictive paths, within- and across-constructs. To my knowledge, Selig and Preacher (2009) introduced the first change-to-change specification with the LCS-MM, although others had been exploring change-to-change models in substantive fields prior to that (e.g., Hertzog et al., 2003).

A path diagram of the LCS-MM is displayed in Figure 4. To achieve temporal separation, Selig and Preacher (2009) modeled each construct in the mediation chain at staggered time-

points such that  $x$  is modeled only at Time 1 and 2,  $m$  is modeled at Time 2 and 3, and  $y$  is modeled at Time 3 and 4. I uphold this convention for clarity, although the spacing of assessments could vary in practice. Because the latent changes are the outcomes of primary interest, I present these equations below:

$$\Delta x_{i2} = \beta_x (x_{i1}) + \xi_{\Delta x i2} \quad (8)$$

$$\Delta m_{i3} = \beta_m (m_{i2}) + \beta_{mx} (x_{i1}) + \beta_{\Delta m \Delta x} (\Delta x_{i2}) + \xi_{\Delta m i3} \quad (9)$$

$$\begin{aligned} \Delta y_{i4} = & \beta_y (y_{i3}) + \beta_{ym} (m_{i2}) + \beta_{yx} (x_{i1}) + \beta_{\Delta y \Delta m} (\Delta m_{i3}) \\ & + \beta_{\Delta y \Delta x} (\Delta x_{i2}) + \xi_{\Delta y i4} \end{aligned} \quad (10)$$

In the LCS-MM, changes in  $x$  in the first interval ( $\Delta x_{i2}$ ) for person  $i$  are a function only of the initial level of  $x$  plus a disturbance,  $\xi_{\Delta x i2}$ . Changes in  $m$  are modeled in the second interval ( $\Delta m_{i3}$ ), and are structured as a function of the level of  $m$  at Time 2, the level of  $x$  at Time 1, the change in  $x$  between Time 1 and 2, and a disturbance  $\xi_{\Delta m i3}$ . Finally, changes in  $y$  are modeled in the third interval ( $\Delta y_{i4}$ ), and are a function of the level of  $y$  at Time 3, the level of  $m$  at Time 2, the level of  $x$  at Time 1, changes in  $m$  that occurred between Time 2 and 3, changes in  $x$  that occurred between Time 1 and 2, and a disturbance  $\xi_{\Delta y i4}$ .

When comparing aspects of the LCS-MM to the standards put forth by Cole and Maxwell (2003), there are several discrepancies worth noting. First, Cole and Maxwell (2003) established the need to control for prior values of endogenous variables through use of autoregressive effects. In the LCS-MM, interest primarily lies in the cross-construct change-to-change effects, which suggests that *within-construct prior changes should be accounted for*. Stated differently, if a researcher is interested in predicting changes in  $m$  between Time 2 and 3 from changes in  $x$  between Time 1 and 2, an unbiased estimate of this effect will only be obtained if the effect of changes in  $m$  between Time 1 and 2 is partialled out. Again, by virtue of occurring over the same

interval (i.e., between Time 1 and 2), it is likely that changes in  $x$  and changes in  $m$  will be correlated; thus, it is requisite to include prior changes in  $m$  as a predictor. However, given that the LCS-MM has only one change score per construct, this specification is not possible.

Importantly, the LCS-MM *does* control for prior *levels* of  $x$ ,  $m$ , and  $y$ . Consider again the case in which the motivating hypothesis is that changes in  $x$  ( $\Delta x$ ) predict changes in  $m$  ( $\Delta m$ ), which subsequently predict changes in  $y$  ( $\Delta y$ ). The partial regression weights  $\beta_{\Delta m \Delta x}$  and  $\beta_{\Delta y \Delta m}$  will inform this hypothesis. Namely,  $\beta_{\Delta m \Delta x}$  will represent the effect  $\Delta x$  has on  $\Delta m$ , *accounting for prior levels of  $m$* . Whether prior levels should be controlled for is the concern of Lord's Paradox (Lord, 1967). Lord's Paradox was introduced in the context of an experimental design with two groups. The paradox refers to the fact that researchers draw discrepant results and conclusions depending on whether they (a) use the grouping variable (i.e., control versus experimental) to predict the Time 2 post-test, while controlling for the Time 1 pre-test (e.g., in an ANCOVA), *or* (b) predict change (i.e., Time 2 post-test minus Time 1 pre-test) from the grouping variable (e.g., in a difference/change score model). Recent work has explored the issue of Lord's Paradox and clarified a number of aspects: (1) it is not limited to experimental designs with a categorical grouping variable, but rather pertains in the context of a continuous predictor as well; (2) ANCOVA and difference/change models make different assumptions about the data, and these assumptions are untestable; and (3) difference/change score models are preferable over ANCOVA models in the context of observational studies, whereas ANCOVA is preferable when randomization is part of the study design (van Breukelen, 2013; Castro-Schilo & Grimm, in press). Analogous to ANCOVA, the LCS-MM controls for previous levels of  $x$ ,  $m$ , and  $y$ , but this practice might only be ideal in experimental studies. Instead, controlling for prior changes in  $x$ ,  $m$ , and  $y$  might be preferable to obtain unbiased estimates of the change-to-change effects.

There is another aspect of the LCS-MM that will be troublesome in most if not all applications. First, recall that one of the benefits of the CLPM is the ability to test alternate structural relations among processes (e.g., lag-2). Unfortunately, the variety of structures that can be tested using the LCS-MM is limited by the fact that there is just one change score per construct. As a result, this setup prohibits testing different autoregressive structures among the changes. Similarly, it is not possible to test different patterns of causality (e.g.,  $\Delta y$  predicting  $\Delta m$ , predicting  $\Delta x$ ) in the LCS-MM because each downstream latent change is modeled at a later point in time than the upstream change scores. In the context of the depression, cognitive ability, and disability hypothesis, the LCS-MM would require changes in depression to be modeled between Time 1 and 2, changes in cognitive health between Time 2 and 3, and changes in depression between Time 3 and 4. Thus, it would be impossible to test a model wherein changes in disability lead to changes in cognitive health, which then lead to changes in depression. Such a model would test temporally backward effects, unless a different set of data was used. Thus, the setup of the LCS-MM may lead researchers to engage in confirmation bias with increased frequency. This unfortunate consequence of the LCS-MM is the major impetus for the development of an expanded LCS model suited for testing mediation among changes.

### **The Cross-Lagged Latent Change Score Model**

At each time-point and for each construct, the measurement model of the CL-LCS model is expressed as in Equation (1). For maximum benefits, each level/status of the constructs across time should be modeled as multiple indicator latent factors so that the factors capture common variance and can be purged from measurement error. However, for simplicity I consider the case of only one indicator per construct, fixing the residuals of the manifest variables to equality as was shown in the CLPM section. Using latent variables with single indicators (and estimating the

residuals of the indicators) is preferable over taking the differences of observed variables because the factors are corrected for unreliability of the manifest variables. Moreover, a fundamental advantage of the LCS modeling framework is to be able to purge difference/change scores from measurement error, as traditional criticisms of difference/change scores no longer apply when these are perfectly reliable (Castro-Schilo & Grimm, under review; Henk & Castro-Schilo, 2015).

For the structural specification of the CL-LCS model, each process has two exogenous variables: the initial level (i.e., at Time 1) and the first latent change score. After the first time-point, the level for each construct is defined as a function of the previous time-point plus the within-person change, exactly as was done in Equation (6).

The hallmark of the CL-LCS model is its ability to flexibly incorporate various structural associations among the change scores. Here, I consider one potential specification; that is, the case whereby the effect that  $\Delta x$  has on  $\Delta y$  is fully mediated through  $\Delta m$  (see Figure 5). For the  $x$  process, each change score after the first interval is specified as a function of the prior change score:

$$\Delta x_{it} = \gamma_{\Delta x} (\Delta x_{it-1}) + \zeta_{\Delta xit} \quad \text{for } t > 1 \quad (11)$$

where  $\Delta x_{it}$  is the latent change that occurred between time  $t - 1$  and  $t$ ,  $\gamma_{\Delta x}$  is the autoprotection effect capturing the influence of prior change on current change,  $\Delta x_{it-1}$  is the prior change in  $x$ , and  $\zeta_{\Delta xit}$  is a random disturbance for individual  $i$  at time  $t$ . In this specification, I set  $\gamma_{\Delta x}$  to be invariant across time, although this assumption may be relaxed. The  $m$  and  $y$  processes operate similarly, except each has an additional term in its equation to capture the cross-construct change-to-change effects:

$$\Delta m_{it} = \gamma_{\Delta m} (\Delta m_{it-1}) + \gamma_{\Delta m \Delta x} (\Delta x_{it-1}) + \zeta_{\Delta mit} \quad \text{for } t > 1 \quad (12)$$

$$\Delta y_{it} = \gamma_{\Delta y} (\Delta y_{it-1}) + \gamma_{\Delta y \Delta m} (\Delta m_{it-1}) + \zeta_{\Delta y it} \quad \text{for } t > 1 \quad (13)$$

where  $\gamma_{\Delta m \Delta x}$  is the change-to-change effect of prior  $\Delta x_{it}$  on  $\Delta m_{it}$ ,  $\zeta_{\Delta x it}$  is a random disturbance for individual  $i$  at time  $t$ ,  $\gamma_{\Delta y \Delta m}$  is the change-to-change effect of  $\Delta m$  at the previous time-point on  $\Delta y$ , and  $\zeta_{\Delta y it}$  is a random disturbance for individual  $i$  at time  $t$ . All disturbances are uncorrelated with one another and with all other latent variables in the model. However, this assumption can be relaxed to accommodate within-time covariances among the residuals of the change scores. Incorporation of these covariances implies there are omitted causal effects that explain shared residual variance in the change scores. Notice that unlike Equations (8) and (9), the change score equations here do not include levels as predictors.

In many ways, the CL-LCS model can be considered an extension of the LCS-MM. Specifically, the CL-LCS expands upon the LCS-MM so that changes in each construct can be modeled at every time-point, rather than only at pre-selected waves. Contrast Figure 5, which contains a path diagram of the CL-LCS model, with Figure 4; with four time-points of data, the CL-LCS model contains nine latent change score variables—three per construct—as opposed to the LCS-MM, which includes just one latent change per construct. Importantly, inclusion of multiple latent changes for each construct in the CL-LCS model allows researchers to conduct more rigorous tests of longitudinal mediation (Cole & Maxwell, 2003).

In sum, the CL-LCS model specification has several key advantages over the LCS-MM approach. The CL-LCS controls for prior values of each endogenous latent change, which yields unbiased cross-construct change-to-change effects. Moreover, the CL-LCS allows for tests of overall change-to-change indirect effects (provided there are four or more waves of data), whereas the LCS-MM allows for tests of overall indirect effects, but the change-to-change indirect effect is purely time-specific (e.g., changes in  $x$  between Time 1 to 2, changes in  $m$

between Time 2 and 3, changes in  $y$  between Time 3 and 4). Finally, the CL-LCS allows researchers to test different structures of causality. A full CL-LCS model can be specified so that (a) changes in  $x$  predict changes in  $m$ , which predict changes in  $y$  and (b) changes in  $y$  predict changes in  $m$ , which predict changes in  $x$ . Removal of paths in (a) or (b) will yield a nested model that can be compared to the full model using an LRT. Importantly, this specification represents a fully mediated process, and in practice it would be advantageous to test for the direct effect from changes in  $x$  to changes in  $y$ .

Ultimately, the CL-LCS model allows researchers to make full use of their data, which results in increased accuracy of model estimates, whereas the LCS-MM requires researchers to truncate their data such that each construct is represented by just two measurement occasions. In the next section I briefly highlight differences between the CLPM and CL-LCS model to further argue for consideration of relations among changes.

### **Relations Among Structural Parameters in the CLPM and CL-LCS Model**

At first glance, the CLPM and CL-LCS appear quite similar; the former concerns relations among levels of variables while the latter concerns relations among changes. A comparison of Figure 1 with Figure 6 highlights the conceptual similarities between these two approaches. However, the following example shows how these models differ with respect to the structure placed on observed variables. Consider the mediator process and, more specifically, the CLPM model-predicted value of  $m$  at Time 3 for individual  $i$ . Using Equation (4), the level of  $m$  is

$$m_{i3} = \beta_m(m_{i2}) + \beta_{mx}(x_{i2}) + \xi_{mi3} \quad (14)$$

In contrast, Equation (13) of the CL-LCS model dictates what the structure of the change in  $m$  should be at Time 3 for individual  $i$ :

$$\Delta m_{i3} = \gamma_{\Delta m} (\Delta m_{i2}) + \gamma_{\Delta m \Delta x} (\Delta x_{i2}) + \zeta_{\Delta m i3} \quad (15)$$

In this example, the changes can be re-written as the differences between adjacent levels. Solving for  $m_{i3}$  yields

$$\begin{aligned} m_{i3} - m_{i2} &= \gamma_{\Delta m} (m_{i2} - m_{i1}) + \gamma_{\Delta m \Delta x} (x_{i2} - x_{i1}) + \zeta_{\Delta m i3} \\ m_{i3} &= \gamma_{\Delta m} (m_{i2} - m_{i1}) + m_{i2} + \gamma_{\Delta m \Delta x} (x_{i2} - x_{i1}) + \zeta_{\Delta m i3} \\ m_{i3} &= m_{i2} (\gamma_{\Delta m} + 1) - \gamma_{\Delta m} (m_{i1}) + \gamma_{\Delta m \Delta x} (x_{i2} - x_{i1}) + \zeta_{\Delta m i3} \end{aligned} \quad (16)$$

Examination of Equations (15) and (17) make it clear that the CL-LCS model implicitly places a structure on the levels of the variables that is quite different from that of the CLPM. As variables further downstream are considered (e.g., the  $y$  process at Time 4), the equations become more complicated. However, the general pattern remains: the CL-LCS model imposes a different structure on a set of data and indeed tests a different theory as compared to the CLPM.

Moreover, the CL-LCS model does not inherit all the methodological limitations of the CLPM. This is particularly important in light of recent work by Hamaker and colleagues (2015) and is the topic of the following section.

### **Limitations of the CLPM**

Although the CLPM is “the most often used mediational model for longitudinal data” (Preacher, 2015, p. 831), Hamaker and colleagues (2015) recently pointed out some of its limitations. Notably, the model requires a stationary process and is not suited for constructs that exhibit patterns of stable, trait-like individual differences. The requirement of stationarity limits the utility of the CLPM if systematic mean trends over time are expected. In contrast, the CL-LCS model should be able to model processes with linear trends over time, because an advantage of taking first-order differences –which the latent change scores in the CL-LCS model represent– is that some forms of nonstationary data (e.g., data following a linear trend) become stationary

(Shumway & Stoffer, 2010). For example, consider a process that follows a linear trajectory:

$$y_{it} = \beta_0 + \beta_1(time) + \varepsilon_{it} \quad (17)$$

where  $y_{it}$  is the score of individual  $i$  at time  $t$ , and time is coded as  $0, \dots, t$ . Taking the difference of two adjacent time-points (often referred to as “differencing”) yields

$$\begin{aligned} y_{it} - y_{it-1} &= \beta_0 + \beta_1(time) + \varepsilon_{it} - [\beta_0 + \beta_1(time - 1) + \varepsilon_{it-1}] \\ &= \beta_0 + \beta_1(time) + \varepsilon_{it} - \beta_0 - \beta_1(time - 1) - \varepsilon_{it-1} \\ &= \beta_0 + \beta_1(time) + \varepsilon_{it} - \beta_0 - \beta_1(time) + \beta_1 - \varepsilon_{it-1} \\ &= \beta_1 + \varepsilon_{it} - \varepsilon_{it-1} \end{aligned} \quad (18)$$

Thus, the linear effect of time drops out, and we are left with a stationary process.

Another assumption of the CLPM is that it assumes that stability of constructs over time is not due to trait-like individual differences (Hamaker et al., 2015). If there is between-person variance that is attributable to stable traits, the CLPM will yield biased estimates of the cross-lagged effects. This limitation is particularly noteworthy since many—if not most—psychological constructs are likely characterized in part by individual differences. For example, although a person’s anxiety levels may fluctuate, some individuals in general are more anxious than others. Again, differencing provides a solution to this problem; by definition the effect of stable between-person differences are removed from a change score. In the above example, the stable part of anxiety falls out when differencing two adjacent time points. As such, the CL-LCS model provides pure estimates of within-person effects, whereas the CLPM may lead to estimates that conflate within- and between-person variation.

In sum, there are two major motivations for the development of the CL-LCS model: (1) to provide a method for making stronger change-to-change causal inferences than can be drawn from existing methods and (2) to adhere to rigorous standards for testing longitudinal mediation.

## CHAPTER 2: METHOD

### Proof of Concept

To support the notion of identifiability of the CL-LCS model, and to demonstrate recovery of population parameters, I simulated two datasets with  $N = 100,000$  and fit the CL-LCS model to each<sup>4</sup>. In each dataset, I simulated four waves of yearly data. Single indicators were generated to represent three processes—depression, cognitive ability, and disability—over time, drawing upon existing literature (e.g., Henk & Castro-Schilo, 2015). The original correlation matrix of the exogenous variables and all autoregressive and change-to-change  $\beta$ s remained invariant across the two conditions. The differentiating characteristic of the two datasets was that one was generated with a stable mean across time and the other had a linear trend.

All population parameters used to generate the trend and no-trend data are listed in the first column of Tables 1 and 2, respectively. In the no-trend dataset, all intercepts were zero (see Table 2), whereas in the trend dataset, the intercepts changed over time. Namely, depression was specified in the population to increase on average by 2 units per year, cognitive ability was specified to decrease by 1 unit per year, and disability was specified to increase by 1.5 units per year.

The rationale for simulating datasets with and without trends comes from the similarities between the CL-LCS model and the CLPM. In other words, it is intuitive to assume the CL-LCS model would require data that do not exhibit trends over time, as does the CLPM. However, the

specification of within-person changes in the CL-LCS model is equivalent to taking first-order differences of factor scores. For this reason, linear trends in data are eliminated and should not pose a threat to the accuracy of estimates in the CL-LCS model. I generated the data by specifying the model equations in R and I fit the models using Mplus Version 7.11 (Muthén & Muthén, 1998-2010). All code used for the proof of concept simulation can be found in the Appendix.

In sum, the purpose of the proof of concept was to elucidate the extent to which the CL-LCS model is appropriate for data with and without a linear trend, and to gain preliminary insight into the identifiability, parameter estimate recovery, and model fit under ideal (albeit unrealistic) conditions of no model misspecification and no missing data. Importantly, parameter estimate recovery can only be discussed on a descriptive level given the design of the proof of concept, as there is no way of understanding sampling variability in the case of a single replication.

### **Participants and Procedures**

To illustrate a systematic model-building strategy for the CL-LCS approach, I used the publically available RAND version of the Health and Retirement Study (HRS) data (RAND HRS, 2013). The HRS is a nationally representative study of aging Americans conducted through the University of Michigan. The study began in 1992 and the full dataset currently comprises over 37,000 Americans aged 50 or older who are contacted for interviewing every 2 years. The HRS also enrolls new participants every 6 years to replenish the dataset (Karp, 2007). Following previous guidelines for excluding participants (McArdle, Fisher, & Kadlec, 2007), individuals in the current analyses were omitted from analyses if they (1) were not the primary respondent, (2) had no data on gender or age, or (3) had a sampling weight of zero, indicating

they lived in a nursing home as opposed to the community. To account for within-wave age heterogeneity, I further limited my analyses to one cohort—the early baby boomers, who were born between 1948 and 1953. This group was first interviewed for the HRS in 2004; the current analyses utilized data from four of the most recent available waves (2004, 2006, 2008, and 2010; from here on, I refer to these as waves 1 through 4, respectively). The sample size used in the current analyses consisted of 2,687 elders. Demographically, the sample was 46.60% female and had a mean age of 65.3 years; moreover, 70.86% identified as White/Caucasian, 17.60% as Black/African American, and 11.54% identified as Other.

Sampling weights are available in the HRS dataset to account for oversampling of Blacks, Hispanics, households in the state of Florida, unequal selection probabilities in geographical areas, and response rate group differences by race and geographical location (see Karp, 2007 for more details). There are both household sampling weights and individual sampling weights. In the current analyses, individual sampling weights from 2004 were included in all analyses; as such, parameters were estimated by maximizing a weighted log-likelihood function and standard errors were computed using a sandwich estimator (Muthén & Muthén, 1998-2010). This was achieved in *Mplus* using the maximum likelihood robust (MLR) estimator.

## **Measures**

*Depression.* The HRS includes 8 items from the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977) to assess participants' depression; each item is scored as a binary (yes = 1, no = 0) variable. Six negatively-worded items prompt the participant to answer whether he or she experienced the following sentiments all or most of the time: felt depressed, everything is an effort, sleep is restless, felt alone, felt sad, and could not get going. There were also two positively worded items (felt happy, enjoyed life), but to maintain a

unidimensional measure of depression, these items were not included in the current analyses. To arrive at a single indicator with a continuous measurement scale, I fit two parameter item response theory (2PL-IRT) models to the six negatively worded items at each of the four time-points and estimated using the first wave as the calibration sample. Higher values in the resulting factor scores indicate more depression.

*Cognitive Ability.* I used an immediate recall scale and a delayed recall scale to operationalize cognitive ability (Chien et al., 2013). The immediate recall task consisted of a list of 10 nouns. Respondents could recall the words in any order and were given 1 point for each correctly recalled word. The delayed recall task consisted of the same 10 nouns; respondents were asked to recall these words after a delay of about five minutes spent answering other survey questions. Across waves, respondents were randomly assigned to different lists of nouns, such that they were not recalling the same set of words at every visit.

To arrive at a single indicator of cognitive ability, I fit confirmatory factor analyses to two indicators—immediate recall and delayed recall—at each time-point. To identify the models, the factor loadings were fixed to equality. As was the case with depression, I used the first wave as the calibration sample for estimating factor scores that I used in subsequent analyses, with higher factor scores indicating better cognitive ability.

*Disability.* I also estimated factor scores for disability by fitting IRT models to two scale scores, the Activities of Daily Living Scale (ADLS) and the Instrumental Activities of Daily Living Scale (IADLS). The ADLS includes five tasks (bathing, eating, dressing, walking across a room, and getting in or out of bed). The IADLS includes three tasks (using a telephone, taking medication, and handling money). Each item is scored as a binary variable, with 1 indicating that the participant endorsed having some difficulty with the given task and 0 indicating no difficulty.

Again, to identify the models, equality constraints were placed on the factor loadings (akin to a Rasch model). Higher values in the resulting factor scores indicate more disability.

### **Data Analysis**

In addition to providing the theoretical rationale for the development of the CL-LCS model, a major task of this project was to put forth a systematic model-building strategy for selecting the optimal specification, as longitudinal mediation can unfold in a variety of ways. To simplify the discussion of model building, assume that a researcher hypothesizes that the relation between changes in  $x$  and changes in  $y$  is mediated through changes in  $m$ —from here on, I refer to this causal mechanism as the “downstream” paths (note that in Figure 6 these paths are represented by downward pointing one-headed arrows). The downstream paths also include the direct effect from changes in  $x$  to changes in  $y$ . The competing hypothesis is that changes in  $y$  predict changes in  $m$ , which predict changes in  $x$ —I refer to these as the “upstream paths.” The upstream paths also include the direct effect from changes in  $y$  to changes in  $x$ . One recommended model-building strategy is as follows (note that alternative strategies are possible).

First, a series of models that include both upstream and downstream effects should be fit to the data; I refer to all such models as Models 1a through 1e. The first step is to fit Model 1a to the data, which is a fully-freed CL-LCS model. That is, all autoregressive and cross-lagged effects among the changes are freely estimated, as are the residual variances of the change scores and the within-time covariances among the change scores. The only equality constraints over time are on the residual variances of the manifest variables, which must be included to identify the model when single indicators are used.

In Model 1b, I argue for placing equality constraints on the autoregressive effects within each construct across time. A likelihood ratio test (LRT) should then be performed between

Model 1a and 1b. Because Model 1b is nested within Model 1a, if the LRT is significant, it suggests that Model 1b significantly decrements the fit to the data. In that event, the researcher would need to free the autoregressive effects for each construct, one at a time, and perform additional LRTs to determine which, if any, processes can be described by unchanging autoregressive effects across time.

In Model 1c, equality constraints should be placed on all cross-lagged effects that link the same two constructs. For example, the cross-lagged effect from changes in  $x1$  to changes in  $m2$  would be set to equality with the cross-lagged effect from changes in  $x2$  to changes in  $m3$ . Again, an LRT should be conducted to compare Model 1c to Model 1b; if nonsignificant, the researcher can move on to Model 1d. If the LRT is significant, the researcher must free the cross-lagged effects—one process at a time—to determine which, if any, can remain equal over time without significantly reducing the fit of the model.

In Model 1d, equality constraints are placed on the residuals of the endogenous latent changes within-construct. Again, an LRT should then be computed to determine which, if any, of these residuals can be fixed to equality without sacrificing model fit. Finally, in Model 1e, equality constraints should be placed on the within-time covariances among the latent change residuals (e.g., the residual covariance of changes in  $x1$  and changes in  $m1$  would be fixed to equal the residual covariance of changes in  $x2$  and changes in  $m2$ ). If the LRT is not significant, the equality constraints do not worsen the fit of the model and should be retained; otherwise, they must be freed one process at a time to determine which, if any, can be fixed to equality over time.

After the Model 1 series, which determines whether the downstream and upstream paths should be equal over time, is fit to the data, the researcher can move on to test the need for paths

that are not supportive of the motivating hypothesis. Namely, in Model 2, the upstream paths should be removed. If the LRT is not significant, it suggests that the upstream paths are not needed. Then, to test the alternate pattern of causality, in Model 3 the downstream paths should be removed; like Model 2, this model is nested within Model 1. If this LRT is significant, it suggests the downstream paths should not be removed. If Model 2 is selected as the optimal model, the motivating hypothesis is supported. If Model 3 is selected, the researcher must reconsider the original hypothesis and perhaps conduct additional studies to better understand the direction of effects. In some cases, it may be that neither Model 2 or Model 3 are selected; if one of the sub-models of Model 1 fits the data significantly better, then the researcher should conclude that there are bidirectional effects at play.

In the next section, I present results from the proof of concept and I demonstrate this model-building strategy using the HRS dataset.

## CHAPTER 3: RESULTS

### Proof of Concept

Spaghetti plots of the simulated no-trend and trend processes can be found in Figures 7 and 8, respectively. In both figures, there is considerable variability in individuals' trajectories. However, only in Figure 8 are there mean trends over time. Results from fitting the CL-LCS model to data without and with trends are located in Tables 1 and 2, respectively. In each of these tables, the first numerical column contains the data-generating parameters and the second column contains the estimated parameters. As each condition had just one replication, these results cannot be generalized to inform raw or relative bias. However, several conclusions can be drawn from this replication. Importantly, under ideal conditions including no misspecification and no missing data, the CL-LCS model seemed to adequately reproduce population parameters with decent accuracy in the case of both no-trend and trend data (see Tables 1 and 2). Note that these estimates are subject to sampling error, and only through a more comprehensive simulation study—which was not the goal of the current project—could accuracy of parameter recovery be discussed at length. In addition to fairly accurate parameter estimates, model fit indices reflect excellent fit to the data, as was expected. For the no-trend dataset, the model fit the data well,  $\chi^2(46) = 38.054$ ,  $p = .79$ , RMSEA = 0.000, SRMR = .002, BIC = 3,674,596.64. Likewise, for the trend data, the model also fit well,  $\chi^2(46) = 39.015$ ,  $p = .76$ , RMSEA = 0.0, SRMR = .002, BIC = 3,676,128.09.

In sum, the proof of concept supports the usefulness of the CL-LCS model for evaluating

change-to-change hypotheses, regardless of whether the constructs of interest show trends over time.

### **Illustration: Descriptive Statistics**

For simplicity, changes in depression, changes in cognitive ability, and changes in disability are notated as  $\Delta$ depression,  $\Delta$ cognitive ability, and  $\Delta$ disability, respectively. Table 3 contains the MLR estimated means, standard deviations, and correlations among the factor scores that were used as indicators in all subsequent models. The MLR estimator accounts for missing data and sampling weights, and was thus preferred over traditionally estimated descriptive statistics. As can be seen in Table 3, average depression levels fluctuated over the course of the study—means ranged from 0.138 at the second wave to .070 at the fourth wave, whereas standard deviations were fairly consistent across time. Cognitive ability also fluctuated over time, and the mean increased between waves 1 and 2, before consistently declining to  $M = -.065$  at wave 4. Standard deviations of cognitive ability increased slightly over time, suggesting greater between-person heterogeneity as individuals aged. Finally, disability increased consistently over time, from  $M = 0$  to  $M = 0.063$  over the course of the four waves. Standard deviations were relatively stable over time. Within-construct correlations over time were positive. Correlations among depression over time ranged from 0.50 to 0.56. Cognitive ability correlations over time were between 0.47 and 0.49. Finally, disability over time correlated between 0.46 and 0.61. Interestingly, although decreasing correlations are often observed as time lags increase in longitudinal data, in a few instances, this pattern was not observed in these data.

Cross-construct correlations were in the expected directions, but fairly small. Depression was negatively correlated with cognitive ability, with  $r$  in the  $-.20$  range. Depression was positively correlated with disability, with  $r$  ranging from .29 to .43. Finally, cognitive ability and

disability were negatively correlated, with  $r$  ranging from -.16 to -.23. Importantly, these descriptive results showcase between-person associations. For example, individuals with higher depression tended to have higher disability. These means and correlations do not inform the substantive hypothesis regarding within-person changes, however. Spaghetti plots of factor scores over time for a random subset of 100 individuals are plotted in Figures 9 through 11 for depression, cognitive ability, and disability. For each process, there is considerable heterogeneity in trajectories, yet the mean trajectories are fairly stable over time.

### **Illustration: CL-LCS Model Results**

Because the MLR estimator was used to estimate CL-LCS model parameters, likelihood ratio tests could not be performed using a simple chi-square difference test (as the difference between two scaled chi-squares for nested models is not distributed as chi-square). Thus, for comparison of all nested models, I computed the Satorra-Bentler scaled chi-square difference test (Satorra & Bentler, 2001).

Although residual variances of single indicators can be fixed to equality across time and estimated (e.g., Joreskog, 1978; also see results from proof of concept), models fit to this particular set of data either failed to converge after 60,000 iterations when specified this way, or resulted in non-positive definite psi matrices. To overcome these convergence issues, I fixed the residual variances of the manifest variables (which were factor scores) to zero in all subsequent analyses. Importantly, this solution imposes the unrealistic assumption that the factor scores are perfect reliable. I consider the implications of this solution further in the Discussion section.

First, I fit the fully-freed CL-LCS model (Model 1a). This model provided less than adequate fit to the data,  $\chi^2(29) = 1291.86, p < .01$ , RMSEA = 0.12, CFI = .79, SRMR = .16. Although each subsequent model in the model-building strategy is more parsimonious than

Model 1a—and thus cannot drastically improve model fit—I pursued the search for the optimal model for illustrative purposes.

The next step was to fit Model 1b, in which the autoregressive effects among the changes are fixed to equality across time. This model did not significantly decrement the fit of the model. The chi-square difference test revealed  $\Delta\chi^2(3) = 1.18, p > .05$ , suggesting that equating these effects over time did not significantly decrease the fit of the model. In Model 1c, I fixed the cross-lagged effects to equality over time, and again, the difference test suggested that these constraints did not decrement the fit of the model significantly,  $\Delta\chi^2(4) = 2.66, p > .05$

Next, I fit Model 1d to the data, which imposed equality constraints on the residual variances of the endogenous latent change scores. When compared to Model 1c, again, this model did not significantly reduce model fit,  $\Delta\chi^2(3) = 1.76, p > .05$ . The last model incorporating both upstream and downstream effects was Model 1e, in which I imposed equality constraints on the within-time residual covariances among the latent change scores. The chi-square difference test favored Model 1e over Model 1d,  $\Delta\chi^2(3) = 1.11, p > .05$ .

To test the need for the upstream paths, which were not theory-motivated, in Model 2 I omitted the effects from  $\Delta\text{disability}$  to  $\Delta\text{cognitive ability}$ , from  $\Delta\text{cognitive ability}$  to  $\Delta\text{depression}$ , and from  $\Delta\text{disability}$  to  $\Delta\text{depression}$ . Model 2 did not fit significantly worse than Model 1e,  $\Delta\chi^2(3) = 1.02, p > .05$ . Although Model 2 thus appeared to be an optimal model, it was important to test alternate patterns of causality. Thus, in Model 3 I omitted the downstream effects and included the upstream effects. Model 3 also did not fit significantly worse than Model 1e,  $\Delta\chi^2(3) = 1.72, p > .05$ .

Taken together, these results suggested the possibility that neither the upstream nor the downstream paths were required. To investigate this possibility, I fit a final model, Model 4, that

included no cross-lagged effects among any of the changes. This model did not fit significantly worse than Model 1e,  $\Delta\chi^2(6) = 2.62, p > .05$ , and was thus retained as the optimal model. However, the overall fit of the model was still poor,  $\chi^2(48) = 1249.28, p < .01$ , RMSEA = 0.10, CFI = .80, SRMR = .16, which precludes serious interpretation of model results. For illustrative purposes only, parameter estimates from Model 1d are located in Table 4. All autoregressive effects among the changes were negative and significant, suggesting that changes in a given construct are significantly related to subsequent changes. Moreover, the within-time residual covariances among  $\Delta$ depression and  $\Delta$ cognitive ability, as well as those among  $\Delta$ depression and  $\Delta$ disability, were significant, suggesting omitting causes in the model that, if included, would explain covariation among these change processes.

## CHAPTER 4: DISCUSSION

The purpose of this project was two-fold: to introduce a new approach for testing mediation when the motivating hypothesis predicts associations among within-person changes, and to demonstrate how to fit the model using a systematic model-building strategy. As was discussed at length, the CLPM is not suited for mediation hypotheses involving change as a predictor or mediator. The LCS-MM can be used to test change-to-change hypotheses, but its specification does not promote testing alternate patterns of causality. The newly developed CL-LCS model combines the best features of both these approaches.

Aside from providing both the theoretical and analytic details motivating the development of the CL-LCS model, there were two additional objectives of this project. First, I conducted a small single-replication simulation to serve as a proof of concept. This simulation was important to support the identifiability of the model and provide preliminary evidence of its suitability for data that exhibit and do not exhibit trends over time. Indeed, in both conditions, the CL-LCS approach converged to what appeared to be reasonable parameter estimates, and traditional fit indices indicated excellent fit. The second objective of this project was to fit the CL-LCS model to a real empirical dataset to showcase the utility of the model for testing theoretically motivated change-to-change hypotheses. I hypothesized that among this population, the relation between within-person changes in depression and changes in disability is mediated by changes in cognitive ability.

Fitting the CL-LCS model to real data from the HRS was an invaluable exercise, as a

number of issues were brought to attention. The final model that I selected failed to find any significant effects among changes in depression, changes in cognitive ability, and changes in disability—although each of these constructs had autoregressive effects operating among the within-person changes. Thus, the extent to which an individual changed on a given construct in one interval affected how much change the person experienced in the next interval on that same construct. Although the cross-lagged effects were not significant, the within-time correlations among the residual variances of the change scores were significant. Substantively, these correlations imply existence of omitted causes from the model; in other words, there is some unknown factor that contributes to changes in depression, cognitive ability, and disability. Future work in the aging literature should be conducted to better understand what these omitted causes may be.

Notably, I failed to replicate prior findings that changes in cognitive health mediate the relation between changes in depression and changes in disability (e.g., Henk & Castro-Schilo, 2015). However, there are several reasons why the cross-lagged effects may not have been significant—even if these processes truly are linked in the population. First, Henk and Castro-Schilo (2015) specified a latent change score mediation model with multiple indicator common factors. Specifying latent change scores with residual variances fixed to zero is equivalent to taking raw-score differences of the scale scores. Thus, the measurement error in the change factors of the current illustration is likely high and could potentially obscure true associations across processes. Although I made efforts to improve the measures I used (by using advanced psychometric techniques for obtaining factor scores), the measurement of the constructs could still have been improved if (1) I opted to use multiple indicator latent variables, or (2) had looked for further empirical data that have better measurement properties. I explored both of these

options in several preliminary analyses not reported in this paper. With regard to the first option, a myriad of issues surfaced when I tried to specify such complex models. For example, the indicators for depression and disability had to be modeled as binary, plus each of these two constructs had numerous indicators (8 items each), which increased the complexity of the models in a way that I believe distracted from the purpose of this project. Moreover, specifying common factors required factorial invariance tests for each of the three processes and, again, discussion of those results was tangential to the scope of this project. With regard to the second option, other scales in the HRS dataset did not exhibit better measurement properties. Despite the difficulty involved in the two options I outlined, in future work, I believe both of these must be explored further so applied researchers can use the CL-LCS model in an optimal way.

Another key issue to consider when modeling within-person change is how best to capture change at the intraindividual level. Many psychological instruments were developed to capture between-person differences and may not be sensitive enough to capture small but meaningful changes within an individual (e.g., Cranford et al., 2006). To the extent that the measures used in the empirical example are better suited to capture stable interindividual differences, power to detect relations among within-person changes may be diminished.

Indeed, much work remains to be done to understand the complexities of the CL-LCS model. For example, when autoregressive and cross-lagged effects are equated across time, more precise estimates can be obtained—as more information tends to yield smaller standard errors. A comprehensive simulation study is required to gain insight into the effects of fewer versus more waves of data and at what point accurate autoregressive and cross-lagged effects may be obtained through fewer equality constraints. From a theoretical perspective, for example, a researcher may hypothesize linkages among changes that begin after a critical point in the

study—thus, it might be ideal to have equality constraints on parameters before that point but not after. Other unknown aspects of the CL-LCS model include its power to find effects among changes as a function of the reliability of the measures, as well as its ability to discriminate between processes linked in terms of level/status versus processes truly linked through within-person changes. It would be interesting to run a simulation with two population-generating models (the CLPM and the CL-LCS model), to see how parameter estimates, significance tests, and fit indices come out when the wrong model is fit to the data. It would also be interesting in future research to compare the CL-LCS model results to those from a latent curve model with structured residuals (LCM-SR; Curran, Howard, Bainter, Lane & McGinley, 2014). In the current context, the latter model would capture how an individual's deviation from his/her average level on, say, depression, predicts subsequent deviations from his/her average cognitive ability level. There are subtle yet distinct differences in the types of hypotheses these models map onto; only the CL-LCS model explicitly models change-to-change relations. Moreover, both approaches separate the between-person effects from what is of primary interest—the within-person effects. The CL-LCS model accomplishes this through differencing, whereas the LCM-SR does so through estimating mean and variance parameters of the trajectory (similar to de-trending).

Finally, as it was presented here, the CL-LCS model is suited for observational data, yet mediation hypotheses are fundamentally about causation. The most fool-proof methods for understanding cause and effect rely on experimental manipulation (Shadish, Cook, & Campbell, 2002). A full discussion of the relative strengths and weaknesses of experimental versus observational designs in psychology (e.g., external versus internal validity) is beyond the scope of the current work, but there is—without question—great value in using experimental designs

for assessing mediation (Spencer, Zanna, & Fong, 2005; MacKinnon & Fairchild, 2009; Preacher, 2015). Longitudinal, experimental designs for studying mediation may be the best course of action for ruling out equivalent models (e.g., MacCallum et al., 1993). Another future direction for the CL-LCS model involves outlining how this model may be applied to data collected in the context of an experiment.

In summary, the newly developed CL-LCS model allows for tests of change-to-change hypotheses, which cannot be accomplished by the causal steps approach or the CLPM. Moreover, the CL-LCS model overcomes important limitations of both the CLPM and the LCS-MM. This project certainly does not address all aspects of the CL-LCS model. However, it does represent a solid first step to better match psychological theories of within-person change with an appropriate modeling technique, particularly when those theories are focused on uncovering mechanisms by which relations among changes operate.

Table 1.

*Unstandardized Results: CL-LCS Model, No Linear Trend Simulation*

Parameter	Population	Estimate	SE
<b>Autoregressions</b>			
$\Delta\text{DEP}_t$ ON $\Delta\text{DEP}_{t-1}$	0.30	0.3	0.006
$\Delta\text{COG}_t$ ON $\Delta\text{COG}_{t-1}$	-0.40	-0.395	0.004
$\Delta\text{DIS}_t$ ON $\Delta\text{DIS}_{t-1}$	0.20	0.197	0.006
<b>Cross Construct Regressions</b>			
$\Delta\text{COG}_t$ ON $\Delta\text{DEP}_{t-1}$	-0.20	-0.201	0.003
$\Delta\text{DIS}_t$ ON $\Delta\text{COG}_{t-1}$	0.30	0.302	0.004
<b>Exogenous Covariances</b>			
LDEP1, $\Delta\text{DEP1}$	-0.40	-0.405	0.007
LDEP1, LCOG1	-0.30	-0.296	0.005
LDEP1, LDIS1	0.25	0.251	0.005
LDEP1, $\Delta\text{COG1}$	-0.08	-0.082	0.005
LDEP1, $\Delta\text{DIS1}$	0.10	0.091	0.005
$\Delta\text{DEP1}$ , LCOG1	-0.10	-0.098	0.005
$\Delta\text{DEP1}$ , LDIS1	0.10	0.094	0.005
$\Delta\text{DEP1}$ , $\Delta\text{COG1}$	-0.30	-0.298	0.006
$\Delta\text{DEP1}$ , $\Delta\text{DIS1}$	0.25	0.268	0.005
LCOG1, LDIS1	0.40	0.401	0.004
LCOG1, $\Delta\text{COG1}$	-0.20	-0.191	0.006
LCOG1, $\Delta\text{DIS1}$	-0.15	-0.152	0.005
LDIS1, $\Delta\text{COG1}$	-0.20	-0.193	0.005
LDIS1, $\Delta\text{DIS1}$	-0.10	-0.105	0.006
$\Delta\text{COG1}$ , $\Delta\text{DIS1}$	-0.20	-0.207	0.005
<b>Intercepts</b>			
MV_DEP1	0.00	-0.004	0.004
MV_DEP2	0.00	0	0.004
MV_DEP3	0.00	0	0.005
MV_DEP4	0.00	0.008	0.006
MV_COG1	0.00	-0.006	0.004
MV_COG2	0.00	-0.004	0.004
MV_COG3	0.00	-0.007	0.005
MV_COG4	0.00	-0.002	0.005
MV_DIS1	0.00	-0.011	0.004
MV_DIS2	0.00	-0.007	0.005
MV_DIS3	0.00	-0.008	0.005
MV_DIS4	0.00	-0.003	0.006
<b>Variances</b>			

LDEP1	1.00	1.003	0.008
LCOG1	1.00	0.993	0.007
LDIS1	1.00	1.003	0.007
ΔDEP1	1.00	1.003	0.011
ΔCOG1	1.00	0.993	0.01
ΔDIS1	1.00	1.01	0.01
<b>Residual Variances</b>			
MV_DEP1	0.64	0.638	0.004
MV_COG1	0.36	0.362	0.004
MV_DIS1	0.36	0.358	0.003
ΔDEP_t	0.49	0.496	0.008
ΔCOG_t	0.42	0.419	0.006
ΔDIS_t	0.25	0.248	0.006

\*note: DEP = depression; COG = cognitive health; DIS = disability; trailing numbers indicate the time-point (e.g., 1-4); trailing “t” indicates parameter invariant across time; ΔDEP indicates change score for depression; LDEP indicates the status/level of depression; IDEP1 through IDEP4 indicates the indicator for depression at waves 1 through 4, respectively. ΔCOG indicates change score for cognitive ability; LCOG indicates the status/level of cognitive ability; ICOG1 through ICOG4 indicates the indicator for cognitive ability at waves 1 through 4, respectively. ΔDIS indicates change score for disability; LDIS indicates the status/level of disability; IDIS1 through IDIS4 indicates the indicator for disability at waves 1 through 4, respectively.  $N = 100,000$ .

Table 2.

*Unstandardized Results: CL-LCS Model, Linear Trend Simulation*

Parameter	Population	Estimate	SE
<b>Autoregressions</b>			
$\Delta\text{DEP}_t$ ON $\Delta\text{DEP}_{t-1}$	0.30	0.306	0.006
$\Delta\text{COG}_t$ ON $\Delta\text{OCG}_{t-1}$	-0.40	-0.399	0.004
$\Delta\text{DIS}_t$ ON $\Delta\text{DIS}_{t-1}$	0.20	-0.399	0.004
<b>Cross Construct Regressions</b>			
$\Delta\text{COG}_t$ ON $\Delta\text{DEP}_{t-1}$	-0.20	-0.2	0.003
$\Delta\text{DIS}_t$ ON $\Delta\text{COG}_{t-1}$	0.30	0.306	0.004
<b>Exogenous Covariances</b>			
LDEP1, $\Delta\text{DEP}_1$	-0.40	-0.395	0.007
LDEP1, LCOG1	-0.30	-0.301	0.005
LDEP1, LDIS1	0.25	0.249	0.005
LDEP1, $\Delta\text{COG}_1$	-0.08	-0.086	0.005
LDEP1, $\Delta\text{DIS}_1$	0.10	0.106	0.005
$\Delta\text{DEP}_1$ , LCOG1	-0.10	-0.103	0.005
$\Delta\text{DEP}_1$ , LDIS1	0.10	0.107	0.005
$\Delta\text{DEP}_1$ , $\Delta\text{COG}_1$	-0.30	-0.295	0.006
$\Delta\text{DEP}_1$ , $\Delta\text{DIS}_1$	0.25	0.247	0.005
LCOG1, LDIS1	0.40	0.399	0.004
LCOG1, $\Delta\text{COG}_1$	-0.20	-0.198	0.006
LCOG1, $\Delta\text{DIS}_1$	-0.15	-0.146	0.005
LDIS1, $\Delta\text{COG}_1$	-0.20	-0.203	0.005
LDIS1, $\Delta\text{DIS}_1$	-0.10	-0.097	0.006
$\Delta\text{COG}_1$ , $\Delta\text{DIS}_1$	-0.20	-0.207	0.005
<b>Intercepts</b>			
MV_DEP1	10.00	10.000	0.004
MV_DEP2	12.00	12.001	0.004
MV_DEP3	14.00	14.000	0.005
MV_DEP4	16.01	16.007	0.006
MV_COG1	10.00	9.997	0.004
MV_COG2	9.00	8.999	0.004
MV_COG3	8.00	7.997	0.005
MV_COG4	7.00	7.002	0.005
MV_DIS1	10.00	9.997	0.004
MV_DIS2	11.50	11.502	0.005
MV_DIS3	13.00	13.001	0.005
MV_DIS4	14.51	14.506	0.006
<b>Variances</b>			

LDEP1	1.00	1.006	0.008
LCOG1	1.00	1.003	0.007
LDIS1	1.00	1.007	0.007
ΔDEP1	1.00	0.991	0.011
ΔCOG1	1.00	0.996	0.01
ΔDIS1	1.00	0.995	0.01
<b>Residual Variances</b>			
MV_DEP1	0.64	0.644	0.004
MV_COG1	0.36	0.361	0.004
MV_DIS1	0.36	0.360	0.003
ΔDEP_t	0.49	0.482	0.008
ΔCOG_t	0.42	0.421	0.006
ΔDIS_t	0.25	0.243	0.006

\*note: DEP = depression; COG = cognitive health; DIS = disability; trailing numbers indicate the time-point (e.g., 1-4); trailing “t” indicates parameter invariant across time; ΔDEP indicates change score for depression; LDEP indicates the status/level of depression; IDEP1 through IDEP4 indicates the indicator for depression at waves 1 through 4, respectively. ΔCOG indicates change score for cognitive ability; LCOG indicates the status/level of cognitive ability; ICOG1 through ICOG4 indicates the indicator for cognitive ability at waves 1 through 4, respectively. ΔDIS indicates change score for disability; LDIS indicates the status/level of disability; IDIS1 through IDIS4 indicates the indicator for disability at waves 1 through 4, respectively.  $N = 100,000$ .

Table 3.

*Descriptive Statistics: Indicators of Depression, Cognitive Ability, and Disability*

	1	2	3	4	5	6	7	8	9	10	11	12
<b>M</b>	0.114	0.138	0.096	0.070	-0.002	0.056	0.004	-0.065	0.000	0.013	0.018	0.063
<b>SD</b>	0.713	0.728	0.713	0.710	0.940	0.998	1.022	1.023	0.515	0.539	0.536	0.603
<b>1 IDEP1</b>	1											
<b>2 IDEP2</b>	0.52	1										
<b>3 IDEP3</b>	0.50	0.53	1									
<b>4 IDEP4</b>	0.51	0.53	0.56	1								
<b>5 ICOG1</b>	-0.20	-0.19	-0.18	-0.19	1							
<b>6 ICOG2</b>	-0.21	-0.18	-0.19	-0.21	0.47	1						
<b>7 ICOG3</b>	-0.21	-0.17	-0.21	-0.18	0.47	0.47	1					
<b>8 ICOG4</b>	-0.20	-0.18	-0.22	-0.22	0.47	0.49	0.48	1				
<b>9 IDIS1</b>	0.39	0.34	0.34	0.34	-0.20	-0.19	-0.17	-0.19	1			
<b>10 IDIS2</b>	0.34	0.40	0.35	0.34	-0.19	-0.20	-0.19	-0.23	0.57	1		
<b>11 IDIS3</b>	0.29	0.33	0.41	0.35	-0.18	-0.18	-0.18	-0.21	0.48	0.55	1	
<b>12 IDIS4</b>	0.33	0.35	0.36	0.43	-0.16	-0.20	-0.18	-0.23	0.46	0.54	0.61	1

\*note: Means, standard deviations, and correlations among indicators used in the empirical example. IDEP1 through IDEP4 are the factor scores for depression at waves 1 through 4, respectively. COG1 through ICOG4 are the factor scores for cognitive ability at waves 1 through 4, respectively. IDIS1 through IDIS4 are the factor scores for disability at waves 1 through 4, respectively.  $N = 2,687$ .

Table 4.  
Results: Final Selected Model

	Estimate	SE	Estimate / SE	p
<b>Autoregressive Effects</b>				
$\Delta$ DEP	-0.48	0.015	-32.506	< 0.001
$\Delta$ COG	-0.526	0.014	-37.554	< 0.001
$\Delta$ DIS	-0.429	0.024	-18.059	< 0.001
<b>Exogenous Covariances</b>				
LDEP1, DDEP1	-0.242	0.013	-18.971	< 0.001
LDEP1, LCOG1	-0.136	0.015	-9.222	< 0.001
LDEP1, LDIS1	0.144	0.01	13.708	< 0.001
LDEP1, DCOG1	-0.01	0.016	-0.631	0.528
LDEP1, DDIS1	-0.013	0.01	-1.213	0.225
DDEP1, LCOG1	0.009	0.015	0.612	0.54
DDEP1, LDIS1	-0.015	0.008	-1.844	0.065
DDEP1, DCOG1	0.004	0.016	0.233	0.816
DDEP1, DDIS1	0.04	0.009	4.409	< 0.001
LCOG1, LDIS1	-0.094	0.012	-7.892	< 0.001
LCOG1, DCOG1	-0.448	0.027	-16.847	< 0.001
LCOG1, DDIS1	-0.001	0.013	-0.076	0.939
LDIS1, DCOG1	-0.003	0.012	-0.245	0.807
LDIS1, DDIS1	-0.108	0.01	-10.791	< 0.001
DCOG1, DDIS1	-0.009	0.013	-0.702	0.482
<b>Within-Time Residual Covariances</b>				
$\Delta$ DEP, $\Delta$ COG	-0.018	0.009	-1.991	0.046
$\Delta$ DEP, $\Delta$ DIS	0.045	0.006	7.993	< 0.001
$\Delta$ COG, $\Delta$ DIS	-0.009	0.007	-1.173	0.241
<b>Intercepts</b>				
IDEP1	0.113	0.015	7.409	< 0.001
IDEP2	0.138	0.016	8.588	< 0.001
IDEP3	0.099	0.016	6.127	< 0.001
IDEP4	0.074	0.016	4.529	< 0.001
ICOG1	-0.001	0.02	-0.038	0.97
ICOG2	0.056	0.022	2.51	0.012
ICOG3	-0.005	0.023	-0.233	0.815
ICOG4	-0.082	0.024	-3.427	0.001

IDIS1	0	0.01	0	< 0.001
IDIS2	0.014	0.011	1.178	0.239
IDIS3	0.017	0.012	1.458	0.145
IDIS4	0.062	0.013	4.651	< 0.001

#### **Variances**

LDEP1	0.506	0.014	37.13	< 0.001
LCOG1	0.882	0.029	30.497	< 0.001
LDIS1	0.265	0.015	17.74	< 0.001
DDEP1	0.504	0.02	25.704	< 0.001
DCOG1	1	0.038	26.37	< 0.001
DDIS1	0.24	0.015	15.54	< 0.001

#### **Residual Variances**

ΔDEP	0.354	0.011	32.233	< 0.001
ΔCOG	0.792	0.023	34.783	< 0.001
ΔDIS	0.212	0.011	19.552	< 0.001

\*note: ΔDEP indicates change score for depression; LDEP indicates the status/level of depression; IDEP1 through IDEP4 indicates the indicator for depression at waves 1 through 4, respectively. ΔCOG indicates change score for cognitive ability; LCOG indicates the status/level of cognitive ability; ICOG1 through ICOG4 indicates the indicator for cognitive ability at waves 1 through 4, respectively. ΔDIS indicates change score for disability; LDIS indicates the status/level of disability; IDIS1 through IDIS4 indicates the indicator for disability at waves 1 through 4, respectively.

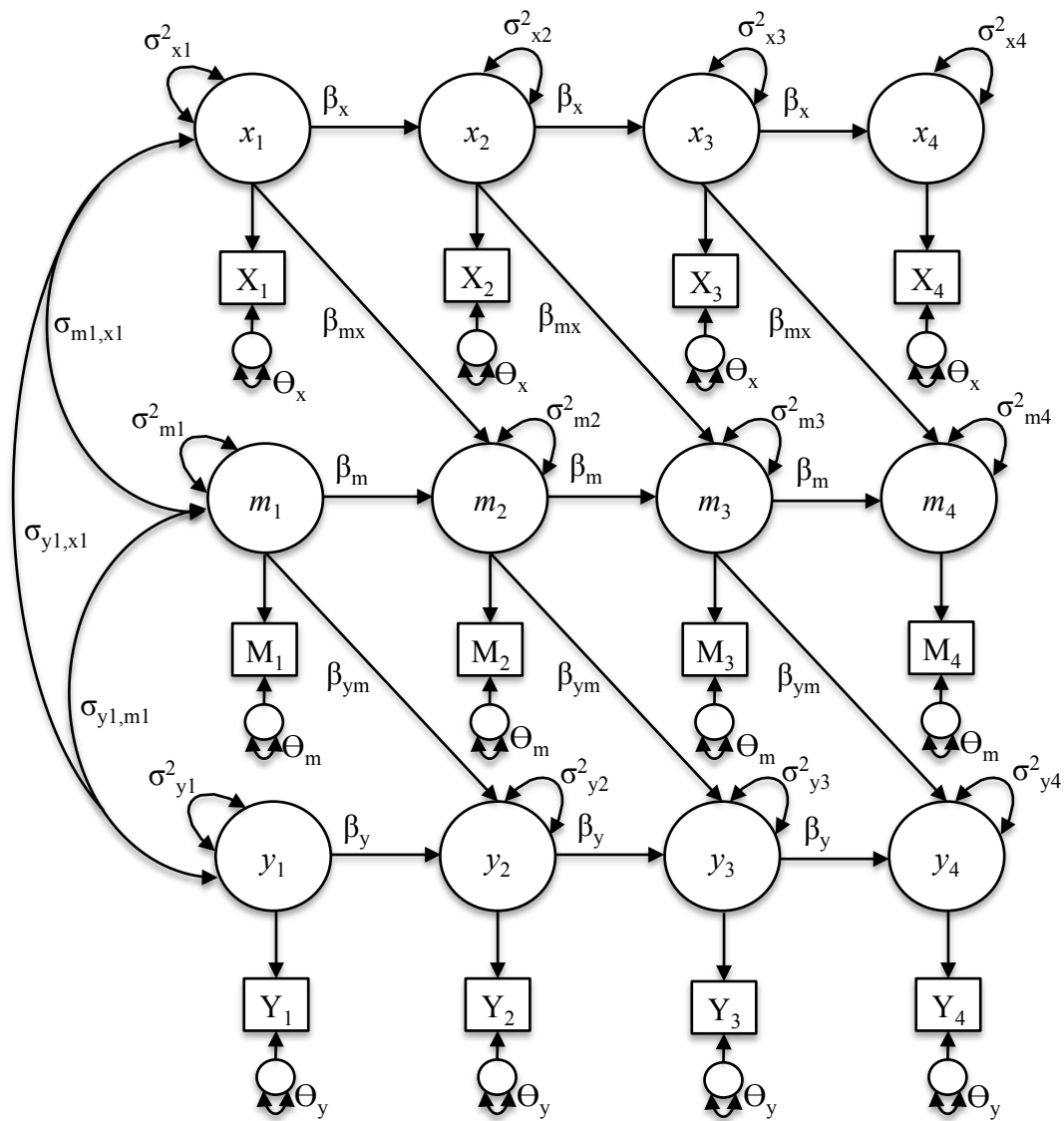


Figure 1. Path diagram of one specification of Cole and Maxwell's (2003) CLPM with latent variables. Direct effects from  $x$  to  $y$  omitted for simplicity.

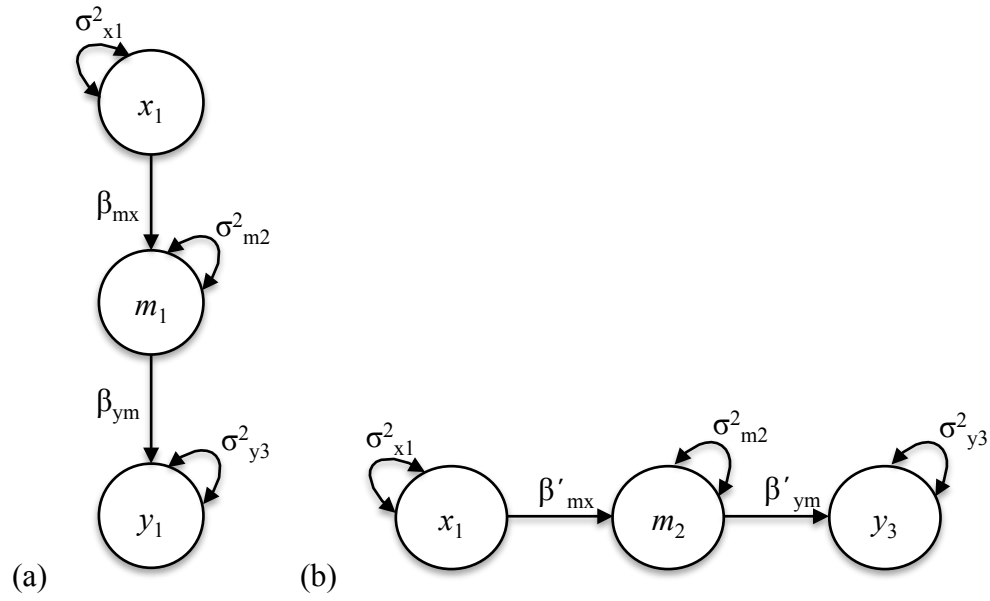


Figure 2. (a) Causal steps cross-sectional model. (b) Causal steps model with temporal separation. For simplicity, measurement model and direct effect paths (i.e.,  $x \rightarrow y$ ) are omitted.

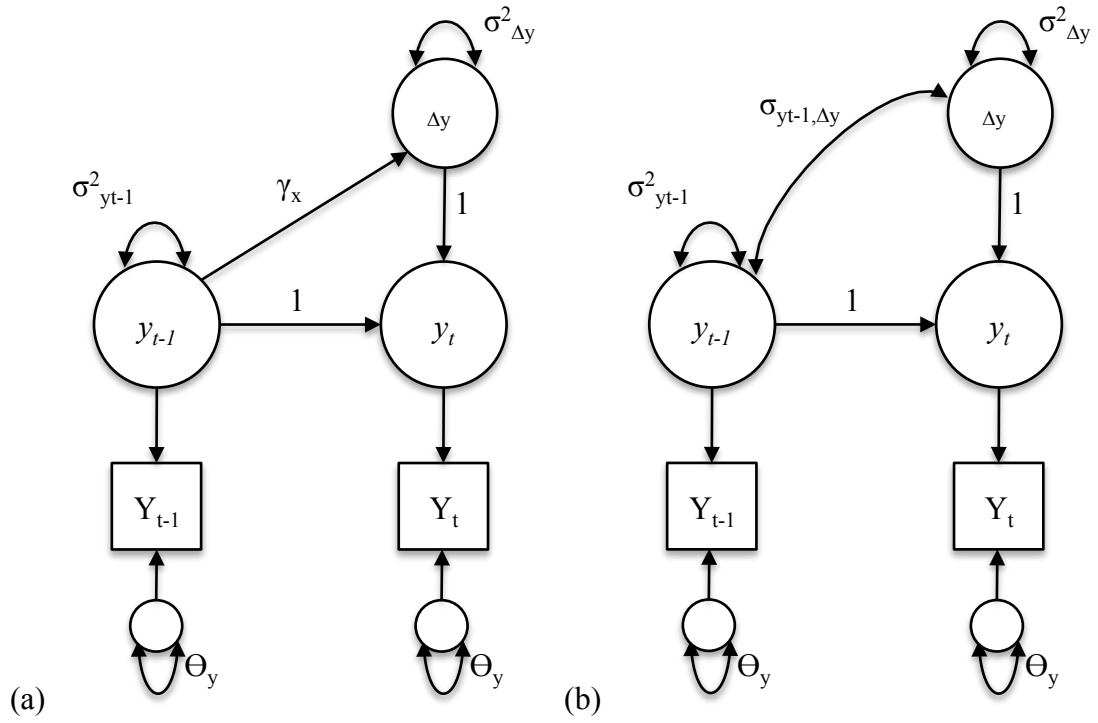


Figure 3. (a) Basic LCS building block with the initial level-to-change association specified as a regression. (b) Basic LCS building block with the initial level-to-change association specified as a correlation.

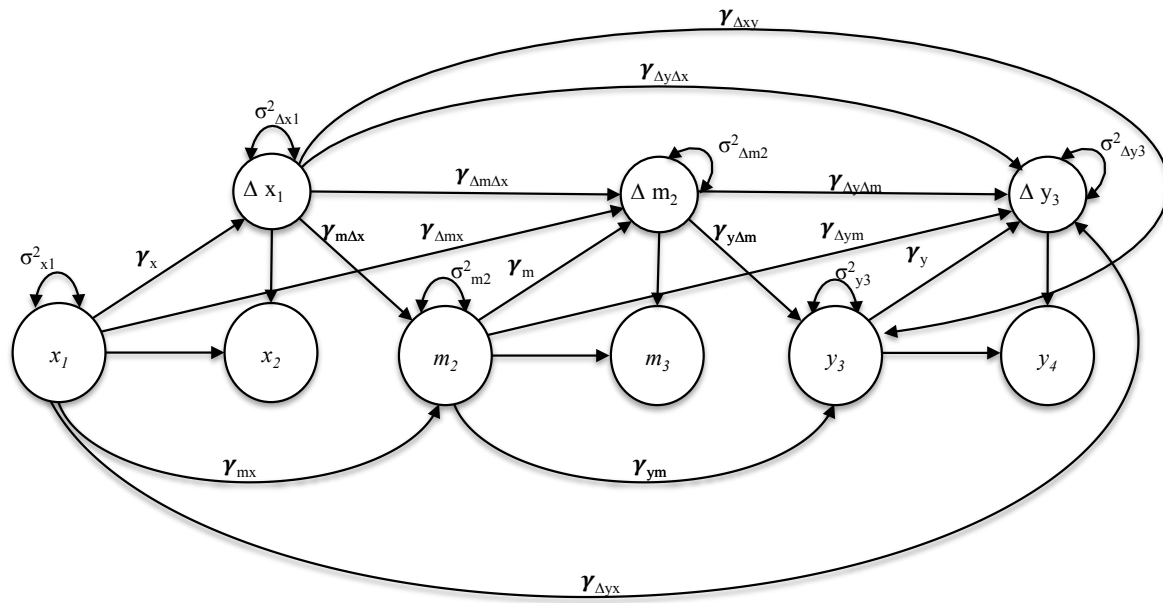


Figure 4. Path diagram of Selig and Preacher's (2009) LCS-MM. Observed variables omitted for simplicity.

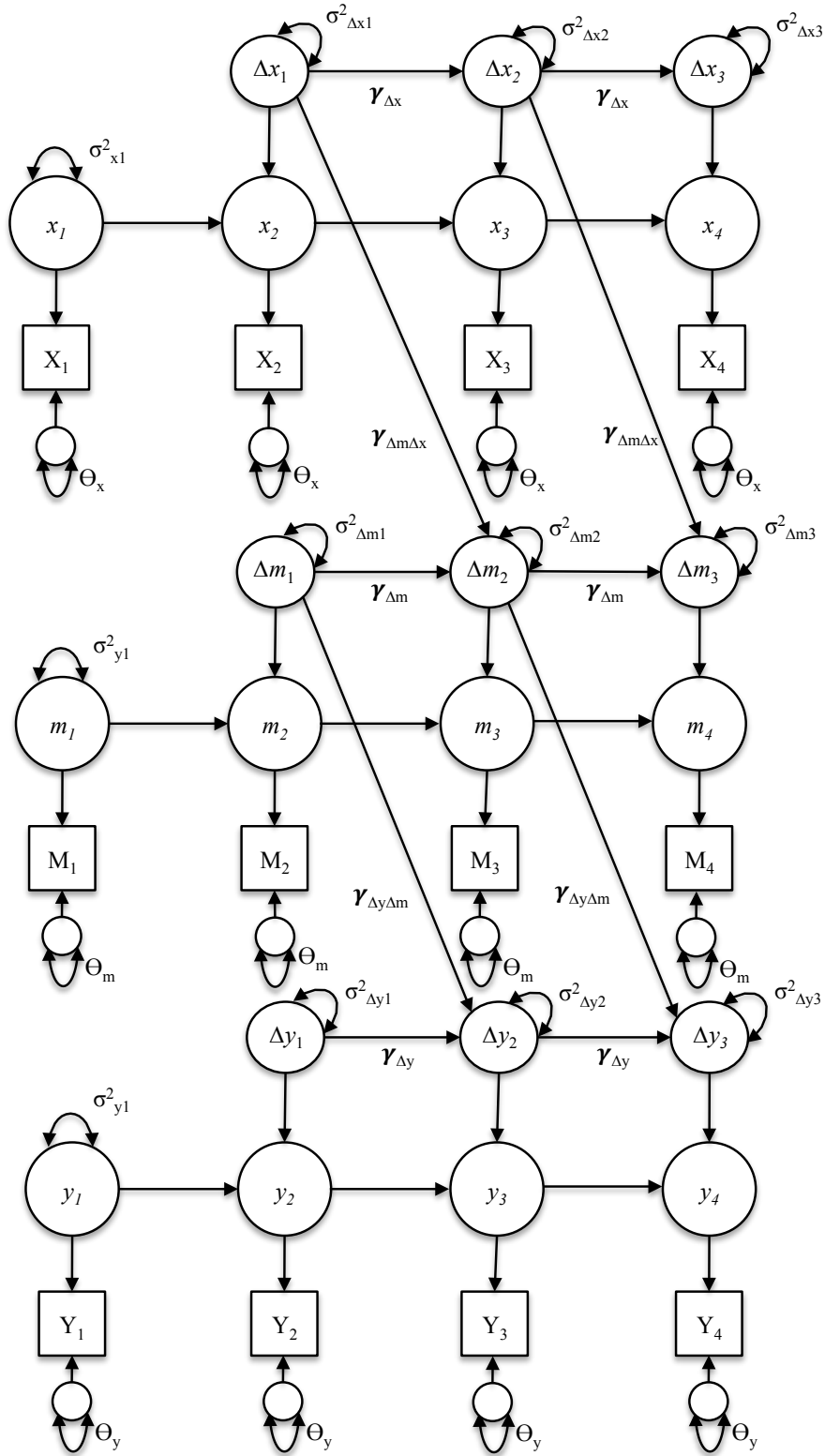


Figure 5. Path diagram of a fully mediated process using the CL-LCS. All unlabeled paths are fixed to 1. All exogenous variables are allowed to covary freely.

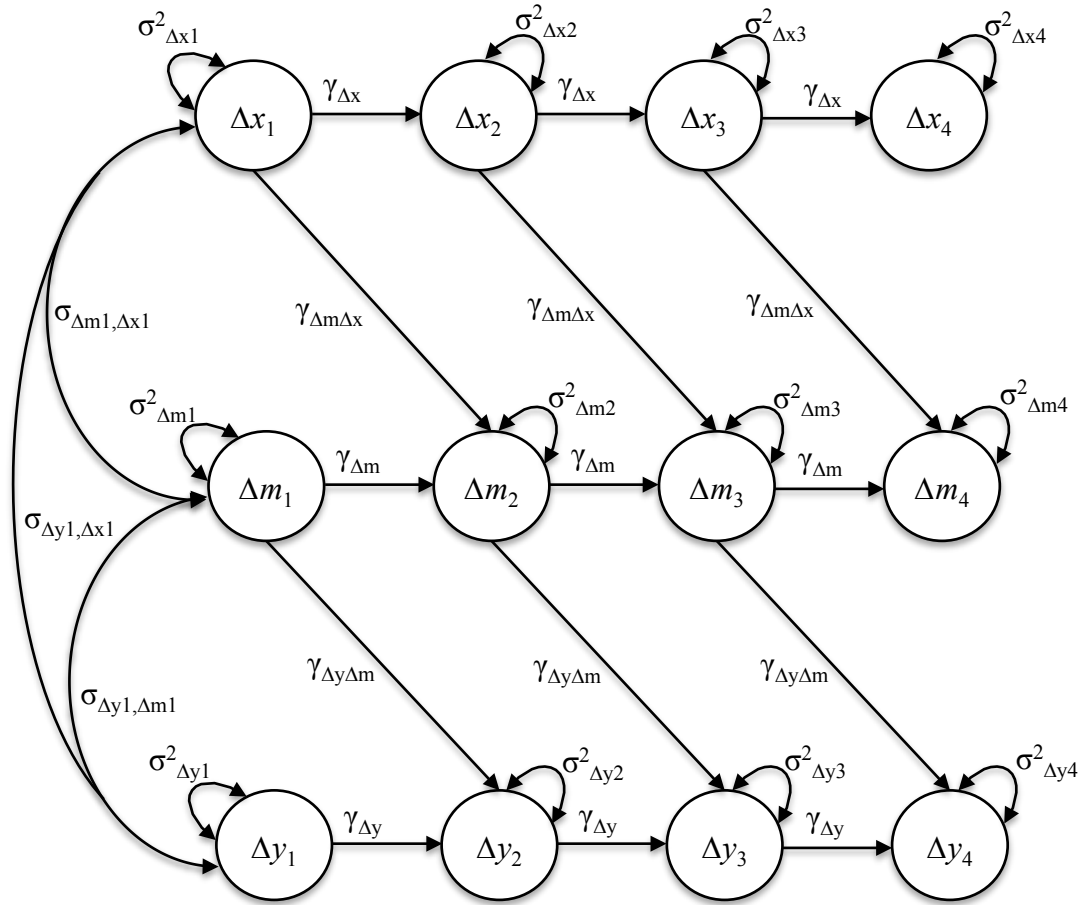
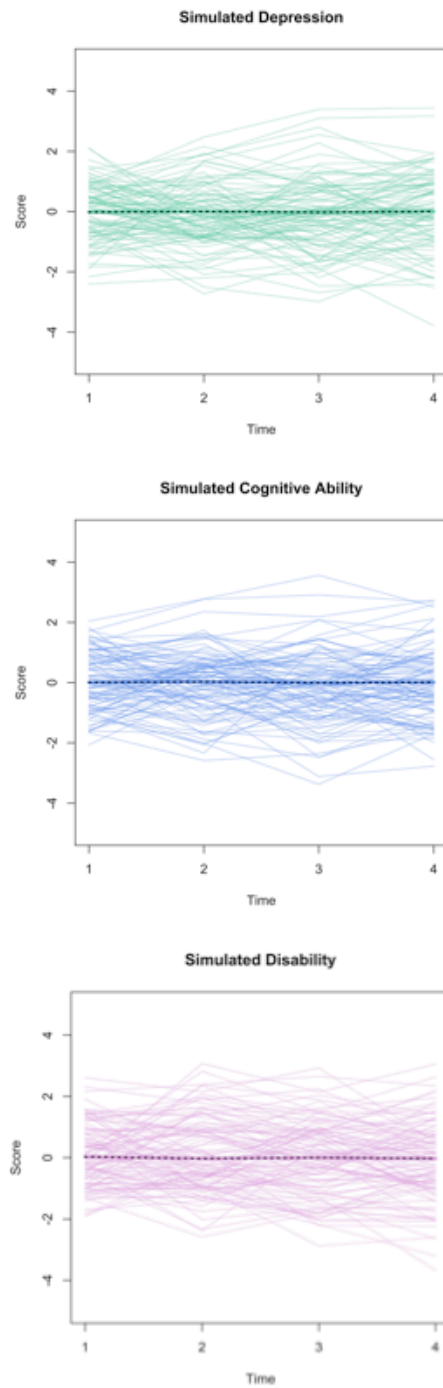
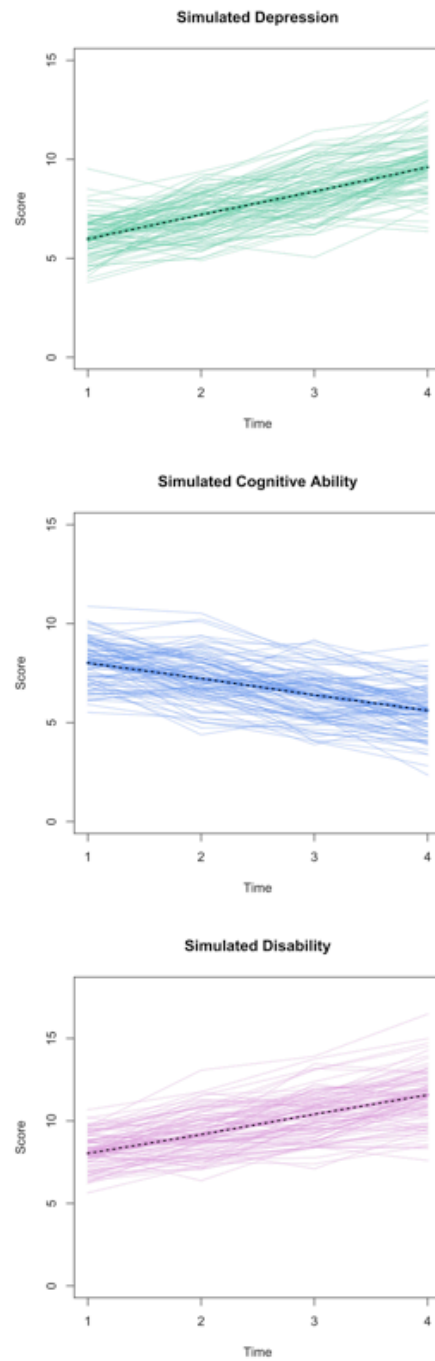


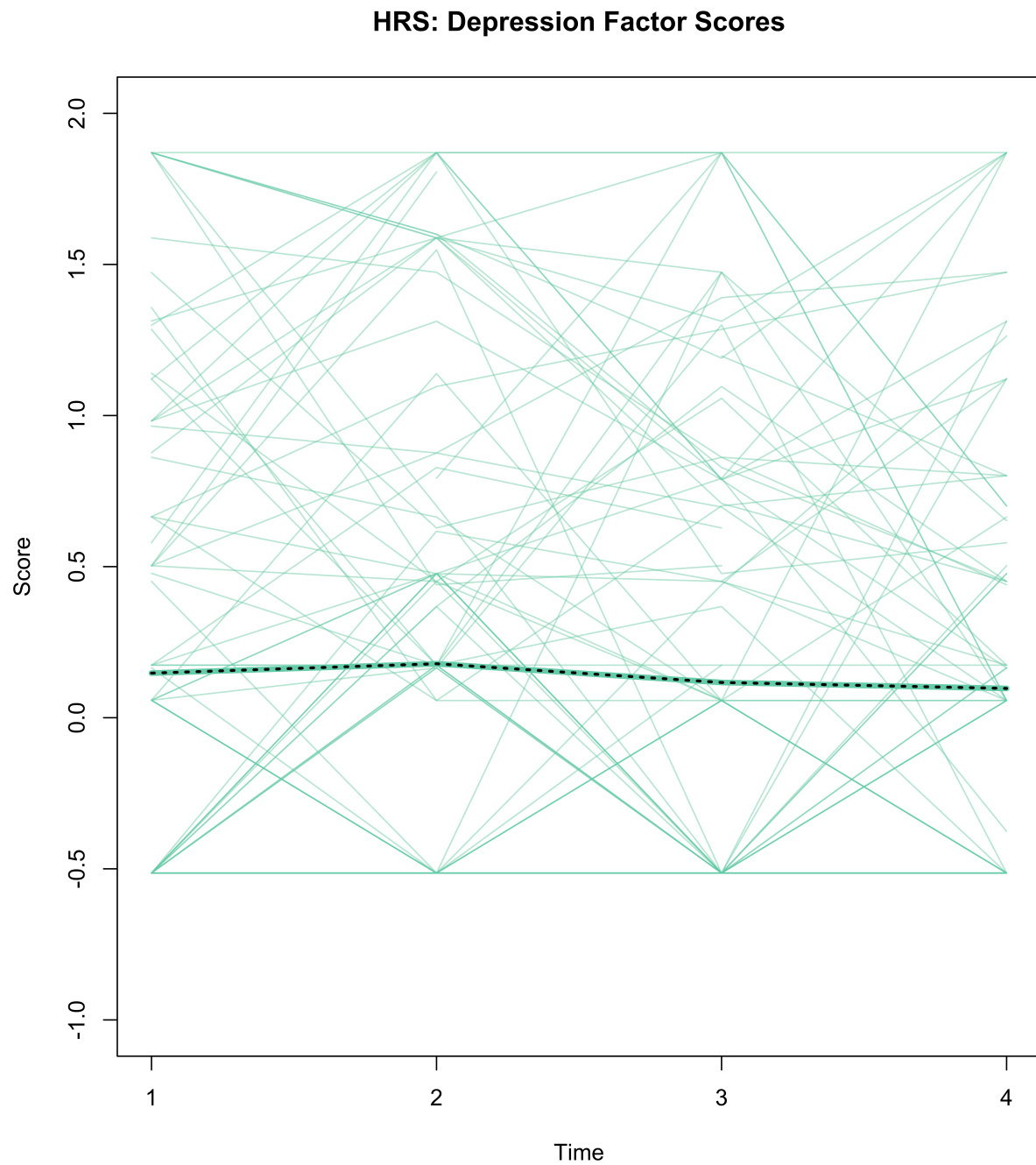
Figure 6. Conceptual diagram of the CL-LCS in the case of full mediation.



*Figure 7.* Random selection of 100 simulated trajectories governed by no trend across 4 waves.

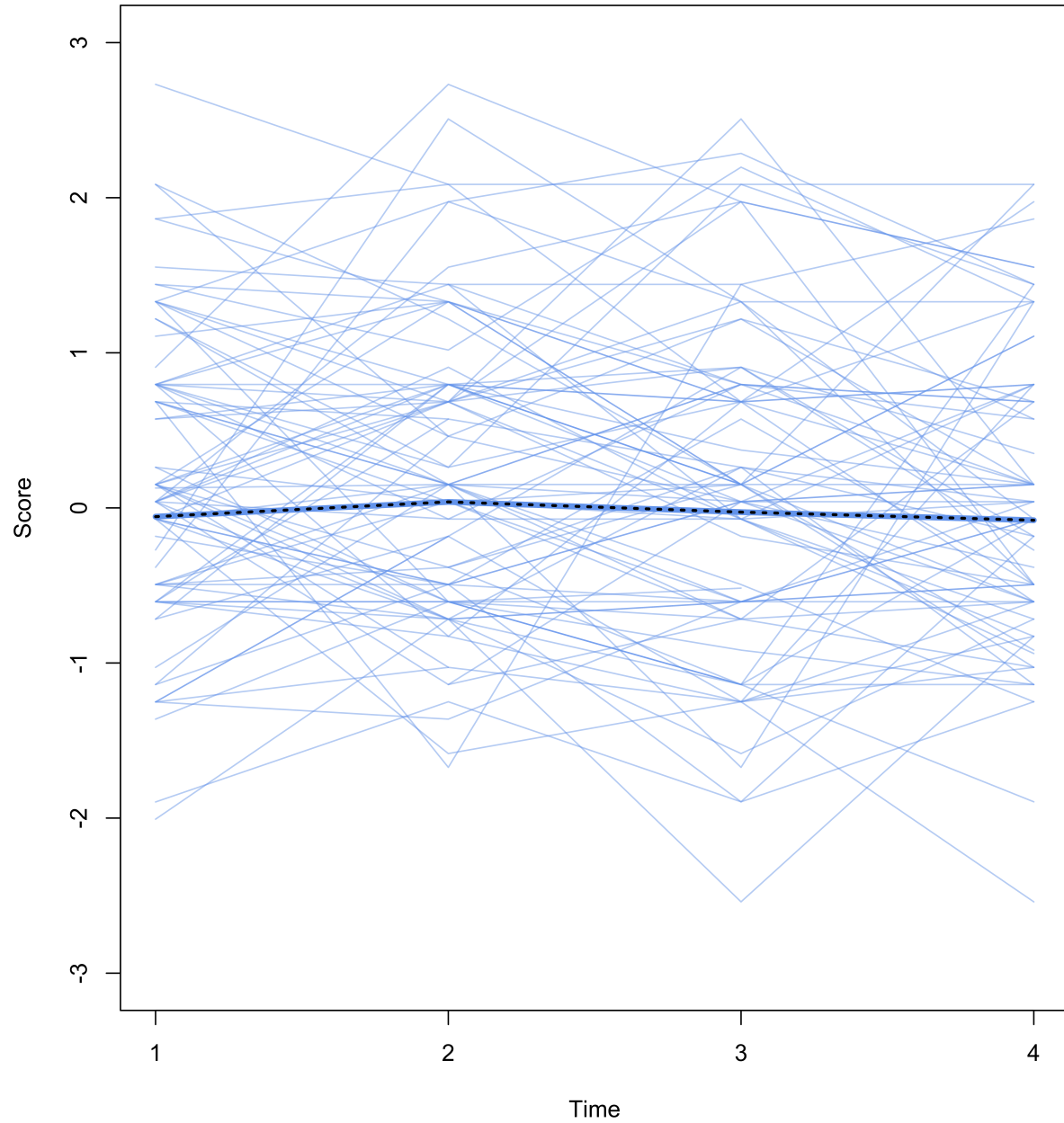


*Figure 8.* Random selection of 100 simulated trajectories governed by a linear trend across 4 waves.

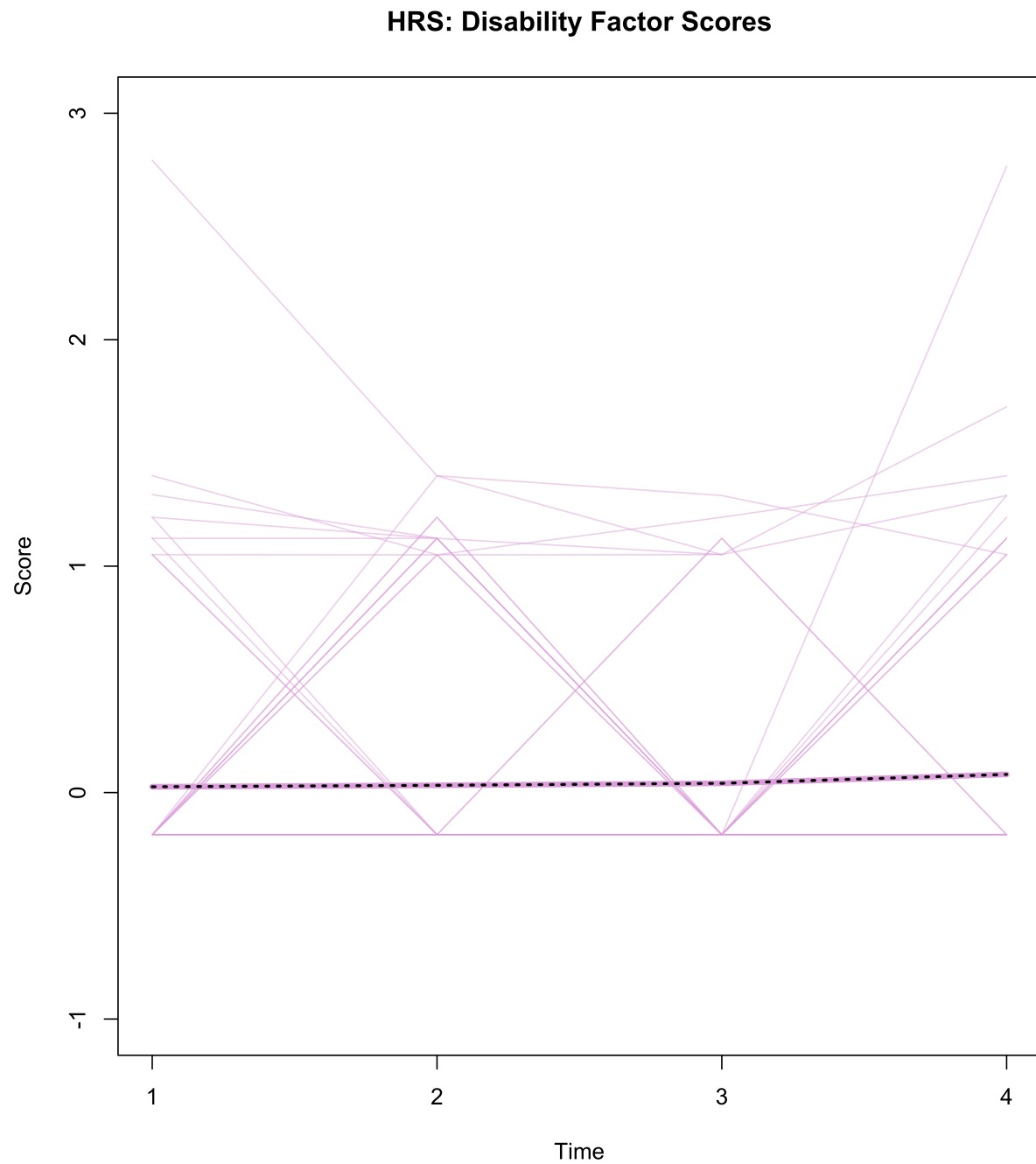


*Figure 9.* Individual trajectories of a random sample of 100 HRS participants' factor scores on depression over time. Dotted line represents the mean trajectory in the full sample. Darker lines represent overlap in trajectories.

### HRS: Cognitive Ability Factor Scores



*Figure 10.* Individual trajectories of a random sample of 100 HRS participants' factor scores on cognitive ability over time. Dotted line represents the mean trajectory in the full sample. Darker lines represent overlap in trajectories.



*Figure 11.* Individual trajectories of a random sample of 100 HRS participants' factor scores on disability over time. Dotted line represents the mean trajectory in the full sample. Darker lines represent overlap in trajectories.

## APPENDIX A: R CODE FOR SIMULATING CL-LCS DATA

```
#####  
# STATIONARY (all means are zero) #  
#####  
  
set.seed(6424235)  
sig = matrix(c(1, -.3, .25, -.4, -.08, .1,  
              -.3, 1, .4, -.1, -.2, -.15,  
              .25, .4, 1, .1, -.2, -.1,  
              -.4, -.1, .1, 1, -.3, .25,  
              -.08, -.2, -.2, -.3, 1, -.2,  
              .1, -.15, -.1, .25, -.2, 1), byrow=TRUE, ncol=6)  
  
my.names <- c("ldep1", "lcog1", "ldis1", "ddep1", "dcog1", "ddis1")  
colnames(sig) = my.names  
  
library(MASS)  
myexogenous = mvrnorm(n = 100000, mu = rep(0,6), Sigma = sig, empirical =  
FALSE)  
mydf = data.frame(myexogenous)  
names(mydf) = my.names  
head(mydf)  
cor(mydf)  
set.seed(123)  
  
#Depression LVs  
mydf$ldep2 = mydf$ldep1 + mydf$ddep1  
mydf$ddep2 = .3*mydf$ddep1 + rnorm(n = 100000, sd = .7)  
mydf$ldep3 = mydf$ldep2 + mydf$ddep2  
mydf$ddep3 = .3*mydf$ddep2 + rnorm(n = 100000, sd = .7)  
mydf$ldep4 = mydf$ldep3 + mydf$ddep3  
  
#Cognitive ability LVs  
mydf$lcog2 = mydf$lcog1 + mydf$dcog1  
mydf$dcog2 = -.4*mydf$dcog1 + ((-.2)*mydf$ddep1) + rnorm(n = 100000, sd =  
.65)  
mydf$lcog3 = mydf$lcog2 + mydf$dcog2  
mydf$dcog3 = -.4*mydf$dcog2 + ((-.2)*mydf$ddep2) + rnorm(n = 100000, sd =  
.65)  
mydf$lcog4 = mydf$lcog3 + mydf$dcog3  
  
#Disability LVs  
mydf$ldis2 = mydf$ldis1 + mydf$ddis1  
mydf$ddis2 = .2*mydf$ddis1 + (.3*mydf$dcog1) + rnorm(n = 100000, sd = .5)  
mydf$ldis3 = mydf$ldis2 + mydf$ddis2  
mydf$ddis3 = .2*mydf$ddis2 + (.3*mydf$dcog2) + rnorm(n = 100000, sd = .5)  
mydf$ldis4 = mydf$ldis3 + mydf$ddis3  
  
#Function to figure out theta-delta  
thde = function(loading){  
  loadsq = loading^2  
  itemvar = 1-loadsq  
  thesd = sqrt(itemvar)  
  return(thesd)  
}
```

```

#Depression MVs (idep_wave#_item#)
mydf$idep11 = mydf$ldep1 + rnorm(n = 100000, sd = .8)
mydf$idep21 = mydf$ldep2 + rnorm(n = 100000, sd = .8)
mydf$idep31 = mydf$ldep3 + rnorm(n = 100000, sd = .8)
mydf$idep41 = mydf$ldep4 + rnorm(n = 100000, sd = .8)

#Cognitive ability MVs (icog_wave#_item#)
mydf$icog11 = mydf$lcog1 + rnorm(n = 100000, sd = .6)
mydf$icog21 = mydf$lcog2 + rnorm(n = 100000, sd = .6)
mydf$icog31 = mydf$lcog3 + rnorm(n = 100000, sd = .6)
mydf$icog41 = mydf$lcog4 + rnorm(n = 100000, sd = .6)

#Disability MVs (idis_wave#_item#)
mydf$idis11 = mydf$ldis1 + rnorm(n = 100000, sd = .6)
mydf$idis21 = mydf$ldis2 + rnorm(n = 100000, sd = .6)
mydf$idis31 = mydf$ldis3 + rnorm(n = 100000, sd = .6)
mydf$idis41 = mydf$ldis4 + rnorm(n = 100000, sd = .6)

head(mydf)

#Export file for Mplus
which(names(mydf)=="idep11")
mymv = mydf[,22:33]

write.table(mymv, file = "apr24_stationary_largeN.csv", col.names = F,
row.names = F, sep=",")
names(mymv)

#####
#      NON-STATIONARY (LINEAR TREND)      #
#####

set.seed(2674663) #different seed from above

sig = matrix(c(1, -.3, .25, -.4,-.08, .1,
               -.3, 1, .4, -.1, -.2,-.15,
               .25, .4, 1, .1, -.2, -.1,
               -.4, -.1, .1, 1,-.3, .25,
               -.08, -.2, -.2, -.3, 1,-.2,
               .1,-.15, -.1, .25, -.2, 1), byrow=TRUE, ncol=6)

my.names <- c("ldep1", "lcog1", "ldis1", "ddep1", "dcog1", "ddis1")
colnames(sig) = my.names

#library(Matrix)
#sig2 = nearPD(sig, corr = TRUE, keepDiag = T, eig.tol = .001)
#sig2 = sig2$mat
library(MASS)
myexogenous = mvrnorm(n = 100000, mu = rep(0,6), Sigma = sig, empirical =
FALSE)
mydf = data.frame(myexogenous)
names(mydf) = my.names
head(mydf)

set.seed(123)

```

```

#Add means to relevant variables outside of mvnorm
mydf$ldep1 = mydf$ldep1 + 10
mydf$l cog1 = mydf$l cog1 + 10
mydf$l dis1 = mydf$l dis1 + 10

#Depression LVs (depression is increasing by 2 units every wave)
mydf$ldep2 = 2 + mydf$ldep1 + mydf$ddep1
mydf$ddep2 = .3*mydf$ddep1 + rnorm(n = 100000, sd = .7)
mydf$ldep3 = 2 + mydf$ldep2 + mydf$ddep2
mydf$ddep3 = .3*mydf$ddep2 + rnorm(n = 100000, sd = .7)
mydf$ldep4 = 2 + mydf$ldep3 + mydf$ddep3

#Cognitive ability LVs (cogab is declining by 1 unit every wave --
irrespective of depression)
mydf$l cog2 = -1 + mydf$l cog1 + mydf$d cog1
mydf$d cog2 = -.4*mydf$d cog1 + ((-.2)*mydf$ddep1) + rnorm(n = 100000, sd =
.65)
mydf$l cog3 = -1 + mydf$l cog2 + mydf$d cog2
mydf$d cog3 = -.4*mydf$d cog2 + ((-.2)*mydf$ddep2) + rnorm(n = 100000, sd =
.65)
mydf$l cog4 = -1 + mydf$l cog3 + mydf$d cog3

#Disability LVs (disability is increasing by 1.5 units every wave --
irrespective of cogab and dep)
mydf$l dis2 = 1.5 + mydf$l dis1 + mydf$d dis1
mydf$d dis2 = .2*mydf$d dis1 + (.3*mydf$d cog1) + rnorm(n = 100000, sd = .5)
mydf$l dis3 = 1.5 + mydf$l dis2 + mydf$d dis2
mydf$d dis3 = .2*mydf$d dis2 + (.3*mydf$d cog2) + rnorm(n = 100000, sd = .5)
mydf$l dis4 = 1.5 + mydf$l dis3 + mydf$d dis3

#Function to figure out theta-delta
thde = function(loading){
  loadsq = loading^2
  itemvar = 1-loadsq
  thesd = sqrt(itemvar)
  return(thesd)
}

#Depression MVs (idep_wave#_item#)
mydf$idep11 = mydf$ldep1 + rnorm(n = 100000, sd = .8)
mydf$idep21 = mydf$ldep2 + rnorm(n = 100000, sd = .8)
mydf$idep31 = mydf$ldep3 + rnorm(n = 100000, sd = .8)
mydf$idep41 = mydf$ldep4 + rnorm(n = 100000, sd = .8)

#Cognitive ability MVs (icog_wave#_item#)
mydf$icog11 = mydf$l cog1 + rnorm(n = 100000, sd = .6)
mydf$icog21 = mydf$l cog2 + rnorm(n = 100000, sd = .6)
mydf$icog31 = mydf$l cog3 + rnorm(n = 100000, sd = .6)
mydf$icog41 = mydf$l cog4 + rnorm(n = 100000, sd = .6)

#Disability MVs (idis_wave#_item#)
mydf$idis11 = mydf$l dis1 + rnorm(n = 100000, sd = .6)
mydf$idis21 = mydf$l dis2 + rnorm(n = 100000, sd = .6)
mydf$idis31 = mydf$l dis3 + rnorm(n = 100000, sd = .6)
mydf$idis41 = mydf$l dis4 + rnorm(n = 100000, sd = .6)

```

```
which(names(mydf)=="idep11")
mymv.nonst = mydf[,22:33]

#Export file for Mplus
write.table(mymv.nonst, file = "apr24_nonstationary_largeN.csv", col.names =
F, row.names = F, sep=",")
```

## APPENDIX B: EXAMPLE MPLUS CODE FOR FITTING CL-LCS MODEL

```
TITLE: One-Indicator Model - Masters - Large N;
DATA: FILE IS "Z:\apr24_nonstationary_largeN.csv";
VARIABLE: NAMES =
    idep11 idep21 idep31 idep41
    icog11 icog21 icog31 icog41
    idis11 idis21 idis31 idis41;

MISSING = .;

USEVARIABLES =
    idep11 idep21 idep31 idep41
    icog11 icog21 icog31 icog41
    idis11 idis21 idis31 idis41;

ANALYSIS:
    ESTIMATOR = ML;
    ITERATIONS = 15000;
    MODEL = NOCOVARIANCES;

MODEL:
    !Make all MVs LVs
    !Depression
        ldep1 BY idep11@1;
        ldep2 BY idep21@1;
        ldep3 BY idep31@1;
        ldep4 BY idep41@1;
    !Cog health
        lcog1 BY icog11@1;
        lcog2 BY icog21@1;
        lcog3 BY icog31@1;
        lcog4 BY icog41@1;
    !Disability
        ldis1 BY idis11@1;
        ldis2 BY idis21@1;
        ldis3 BY idis31@1;
        ldis4 BY idis41@1;

    !LVs' variances
        ldep1*; ldep2@0; ldep3@0; ldep4@0;
        lcog1*; lcog2@0; lcog3@0; lcog4@0;
        ldis1*; ldis2@0; ldis3@0; ldis4@0;

    !Estimate res vars and set to equality to identify
        idep11 idep21 idep31 idep41 (RV1);
        icog11 icog21 icog31 icog41 (RV2);
        idis11 idis21 idis31 idis41 (RV3);

    !Deltas
        ddep1 BY ldep2@1;
        ddep2 BY ldep3@1;
        ddep3 BY ldep4@1;

        dcog1 BY lcog2@1;
        dcog2 BY lcog3@1;
```

```

    dcog3 BY lcog4@1;

    ddis1 BY ldis2@1;
    ddis2 BY ldis3@1;
    ddis3 BY ldis4@1;

!Deltas' variances
    ddep1*; dcog1*; ddis1*;

    ddep2*(V1);
    ddep3*(V1);

    dcog2*(V2);
    dcog3*(V2);

    ddis2*(V3);
    ddis3*(V3);

!LV Covariances
    ldep1 WITH ddep1 lcog1 ldis1 dcog1 ddis1;
    ddep1 WITH lcog1 ldis1 dcog1 ddis1;
    lcog1 WITH ldis1 dcog1 ddis1;
    ldis1 WITH dcog1 ddis1;
    dcog1 WITH ddis1;

!Perfect (1) Autoregressive effects
    ldep4 ON ldep3@1;
    ldep3 ON ldep2@1;
    ldep2 ON ldep1@1;

    lcog4 ON lcog3@1;
    lcog3 ON lcog2@1;
    lcog2 ON lcog1@1;

    ldis4 ON ldis3@1;
    ldis3 ON ldis2@1;
    ldis2 ON ldis1@1;

!Change predicting change (Autoregressions)
    ddep2 ON ddep1 (B1);
    ddep3 ON ddep2 (B1);

    dcog2 ON dcog1 (B2);
    dcog3 ON dcog2 (B2);

    ddis2 ON ddis1 (B3);
    ddis3 ON ddis2 (B3);

!Change predicting change (Crosslags)
    dcog2 ON ddep1 (B4);
    dcog3 ON ddep2 (B4);

    ddis2 ON dcog1 (B5);
    ddis3 ON dcog2 (B5);

!Intercepts
    [idep11 idep21 idep31 idep41];

```

```
[icog11 icog21 icog31 icog41];  
[idis11 idis21 idis31 idis41];
```

## ENDNOTES

<sup>1</sup> More recently, some researchers have argued that it is not necessary to test the  $y$  on  $x$  regression because the effect size may be small or suppression effects may be operating (Shrout & Bolger, 2002).

<sup>2</sup> The reason these are indistinguishable explanations can be illustrated most simply in the context of bivariate regression. Consider fitting two models: Model A treats  $X$  as the predictor and  $Y$  as the outcome, and Model B treats  $Y$  as the predictor and  $X$  as the outcome. Model A and Model B will yield the exact same  $R^2$  and provide the same fit to the data.

<sup>3</sup> With three waves of data, there is exactly one time-specific indirect effect, which is equivalent to the overall indirect effect; in this case, there is little need to distinguish between the two.

<sup>4</sup> To verify that the parameters are in fact unbiased, one would need to include multiple replications. However, this was not the purpose of the current project.

## REFERENCES

- Baltes, P. B. & Nesselroade, J. R. (1979). History and rationale of longitudinal research. *Longitudinal Research in the Study of Behavior and Development*, 1-59.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, 51, 1173-1182.
- Castro-Schilo, L. & Grimm, K. J. (in press). Using residualized change versus difference scores for longitudinal research. *Journal of Social and Personal Relationships*.
- Castro-Schilo, L., Ferrer, E., Hernández, M. M., & Conger, R. D. (2015). Developmental Outcomes of School Attachment Among Students of Mexican Origin. *Journal of Research on Adolescence*. doi: 10.1111/jora.12223
- Chien, S., Campbell, N., Hayden, O. Hurd, M., Main, R., Mallett, J., Martin, C., Meijer, E., Miu, A., Moldoff, M., Rohwedder, S., & St. Clair, P. (2013). RAND HRS Data Documentation.
- Cheong, J., MacKinnon, D. P., & Khoo, S. T. (2003). Investigation of mediational processes using parallel process latent growth curve modeling. *Structural Equation Modeling*, 10, 238-262.
- Cole, D. A. & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, 112, 558–577. doi:10.1037/0021-843X.112.4.558
- Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: can mood measures in diary studies detect change reliably?. *Personality and Social Psychology Bulletin*, 32, 917-929.
- Cronbach, L. J., & Furby, L. (1970). How we should measure" change": Or should we?. *Psychological bulletin*, 74, 68-80.
- Ferrer, E., & McArdle, J.J. (2010). Longitudinal modeling of developmental changes in psychological research. *Current Directions in Psychological Science*, 19, 149-154.
- Grimm, K. J., Castro-Schilo, L., Davoudzadeh, P. (2013). Modeling intraindividual change in nonlinear growth models with latent change scores. *The Journal of Gerontopsychology and Geriatric Psychiatry*, 26, 153-162.
- Grimm, K. J., An, Y., McArdle, J. J., Zonderman, A. B., & Resnick, S. M. (2012). Recent changes leading to subsequent changes: Extensions of multivariate latent difference score models. *Structural Equation Modeling: A Multidisciplinary Journal*, 19, 268–292.

- Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods, 1*, 102-116.
- Henk, C. M. & Castro-Schilo, L. (2015) Preliminary Detection of Relations Among Dynamic Processes With Two-Occasion Data, *Structural Equation Modeling: A Multidisciplinary Journal*. doi:10.1080/10705511.2015.1030022
- Hertzog, C., Dixon, R. A., Hultsch, D. F., & MacDonald, S. W. S. (2003). Latent change models of adult cognition: Are changes in processing speed and working memory associated with changes in episodic memory? *Psychology and Aging, 18*, 755-769.
- Jöreskog, K. G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika, 43*, 443-477.
- Karp, F. (2007). *Growing older in America: The health and retirement study*. National Institute on Aging, National Institutes of Health, US Department of Health and Human Services.
- Curran, P. J., Howard, A. L., Bainter, S. A., Lane, S. T., & McGinley, J. S. (2014). The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology, 82*, 879.
- Lord, F. M. (1978). A paradox in the interpretation of group comparisons. *Psychological Bulletin, 68*, 304-305.
- MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin, 1*, 185-199.
- MacKinnon, D. P., & Fairchild, A. J. (2009). Current directions in mediation analysis. *Current Directions in Psychological Science, 18*, 16-20.
- McArdle, J. J. (2001). A latent difference score approach to longitudinal dynamic structural analyses: a festschrift in honor of Karl Jöreskog. In R. C. K. G. Jöreskog, S. H. C. D. Toit & D. Sörbom (Eds.), *Structural equation modeling: Present and future* (pp. 342-380).
- McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577-605.
- McArdle, J. J., Fisher, G. G., & Kadlec, K. M. (2007). Latent variable analyses of age trends of cognition in the Health and Retirement Study, 1992-2004. *Psychology and Aging, 22*, 525-545.
- McArdle, J. J., & Hamagami, F. (2001). Latent difference score structural models for linear dynamic analyses with incomplete longitudinal data. In L. Collins & A. Sayer (Eds.), *New methods for the analysis of change* (pp. 137-176). Washington, DC: APA Press.

- McArdle, J. J., & Nesselroade, J. R. (1994). Using multivariate data to structure developmental change. In S. H. Cohen & H. W. Reese (Eds.), *Life-span developmental psychology: Methodological innovations* (pp. 223–267). Hillsdale, NJ: Erlbaum.
- Muthén, L.K. and Muthén, B.O. (1998-2010). *Mplus User's Guide*. Sixth Edition. Los Angeles, CA: Muthén & Muthén.
- Preacher, K. J. (2015). Advances in mediation analysis: A survey and synthesis of new developments. *Annual Review of Psychology*, 66, 825-852.
- Radloff, L. S. (1977). The CES–D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385–401.
- RAND HRS. (2013). Data, Version M. Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA: RAND.
- Satorra, A. & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507-514.
- Shumway, R. H., & Stoffer, D. S. (2010). *Time series analysis and its applications: with R examples*. Springer Science & Business Media.
- Selig, J. P. & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development*, 6, 144–164.
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: new procedures and recommendations. *Psychological Methods*, 7, 422-445.
- Smith, L. G. E., Amiot, C. E., Smith, J. R., Callan, V. J., & Terry, D. J. (2013). The social validation and coping model of organizational identity development: A longitudinal test. *Journal of Management*, 39, 1952–1978. doi:[10.1177/0149206313488212](https://doi.org/10.1177/0149206313488212)
- Sobel, M. E. (1990). Effect analysis and causation in linear structural equation models. *Psychometrika*, 55, 495–515.
- Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology*, 89, 845-851.
- van Breukelen, G. J. (2013). ANCOVA versus CHANGE from baseline in nonrandomized studies: The difference. *Multivariate Behavioral Research*, 48, 895-922.