

# A Structural Equation Modeling Approach for the Analysis of Cortisol Data Collected Using Pre–Post–Post Designs

Michael Willoughby  
*FPG Child Development Institute,  
University of North Carolina at Chapel Hill*

Nathan Vandergrift  
*FPG Child Development Institute,  
University of North Carolina at Chapel Hill  
L. L. Thurstone Psychometric Laboratory,  
University of North Carolina at Chapel Hill*

Clancy Blair  
*Department of Human Development and Family Studies,  
Pennsylvania State University*

Douglas A. Granger  
*Department of Biobehavioral Health,  
Pennsylvania State University*

This study introduces a novel application of structural equation modeling (SEM) for the analysis of cortisol data that are collected using a pre–post–post design. By way of an extended example, an SEM model is developed that permits an examination of both the overall level of cortisol, as well as changes in cortisol (reactivity and regulation), as predictors of cognitive (executive) and behavioral functioning in 3- to 5-year-old children ( $N = 171$ ) attending Head Start. The SEM model makes use of the

parameterization of latent curve models. Throughout the extended example, the strengths of using an SEM approach for the analysis of cortisol data that are collected using pre–post–post designs is highlighted.

Stress can be defined as occurring when the internal or external demands made on an organism exceed its available resources. When confronted with stress, animals and humans alike initiate a complex set of neurophysiological changes that are designed to accommodate the stressor and return biological systems back to a point of homeostasis (McEwen, 2000). The limbic hypothalamic pituitary adrenal (LHPA) axis is centrally implicated in all major models of stress response, and corticosteroid measures (cortisol in primates, corticosterone in rodents)—the major output of the LHPA axis—are frequently used as an objective index of an organism's ability to successfully mount a stress response. For purposes here, it is useful to establish that cortisol is centrally implicated in adaptive affective, behavioral, and cognitive functioning, including the neural and neurotransmitter systems that subserve this functioning (Erickson, Drevets, & Schulkin, 2003).

An inverted U-shaped function has been used to describe two different aspects of the relation between the stress response system (as defined by cortisol activity) and adaptive functioning. First, an inverted U-shaped function has been used to describe the short term time course of cortisol following exposure to a stressor or novel event. Thus, a “healthy” cortisol response has been described as the up-regulation of cortisol in the face of a stressor that is followed by the down-regulation of cortisol once the (perception of the) stressor has diminished (Dickerson & Kemeny, 2004; McEwen, 2000). Second, an inverted U-shaped function has also been used to describe the relation between overall levels of cortisol and corresponding affective and cognitive functioning (Erickson et al., 2003; Lupien & McEwen, 1997). Extremely low and high levels of cortisol have been implicated in cognitive dysfunction and atypical or pathological states, whereas moderate levels are implicated in adaptive functioning (Heim, Ehlert, & Hellhammer, 2000). The statistical model introduced here considers baseline levels of cortisol, as well as time-dependent changes in cortisol that may occur following exposure to a novel or stressful event.

Although studies have yielded consistent mean-level differences in patterns of cortisol response following exposure to a stressful or novel event (Dickerson & Kemeny, 2004), interindividual differences in patterns of cortisol response are routinely found (Granger, Stansbury, & Henker, 1994). Even among those individuals who exhibit the expected pattern of effects (increasing cortisol values that subsequently decline), there is variation in the magnitude of the response. Moreover, there are also individuals who exhibit unexpected patterns of change, including no change, as well as continuously increasing or decreasing change at least during the time period in which cortisol was sampled. Researchers are increasingly interested

in the use of research designs and data analytic methods that can characterize interindividual differences in time-dependent changes in cortisol activity following exposure to a stressful or novel event, as well as exploring factors that help to explain this variation (Granger & Kivlighan, 2003). The statistical model introduced here conveys information about average, group-level changes in cortisol over time, as well as information about the extent and magnitude of individual differences around these average effects.

Given an interest in delineating interindividual differences in time-dependent patterns of change in cortisol following exposure to a stressful or novel event, applied researchers routinely collect three (and only three) samples of cortisol. The first of the three samples is typically taken prior to the event and represents a baseline measure. The second and third samples are typically obtained approximately 20 and 40 min following the experience of the event, respectively (Gunnar & White, 2001). Data collected in this matter are referred to here as originating from pre–post–post designs. Differences between the first and second samples are referred to as cortisol reactivity, whereas differences between the first (or second) and third samples are referred to as cortisol regulation. Of key interest is whether individual differences in cortisol reactivity and regulation are associated with individual differences in affective, behavioral, or cognitive functioning. Although it is well known that obtaining more than three measures of cortisol would permit a more detailed understanding of individual differences in cortisol reactivity and regulation, real-world constraints related to the collection of saliva samples (from which cortisol is assayed), especially in studies involving young children or that collect a variety of other measures, often impose practical limitations on the number of saliva samples that can be collected. Given these real-world constraints, the statistical model introduced here makes optimal use of three time points in characterizing time-dependent changes in cortisol activity.

The primary goal of this study is to introduce a structural equation modeling (SEM) approach to the analysis of cortisol data that are collected in pre–post–post designs. To accomplish this task, data that two of the authors (Clancy Blair and Douglas A. Granger) recently published relating patterns of cortisol reactivity and regulation to measures of cognitive and behavioral functioning in preschoolers from low-income homes (Blair, Granger, & Razza, 2005) are analyzed. The models investigated here are directly related to a submodel of SEM, namely the latent curve model (LCM). The parameterization of the LCM is described, which provides a framework for understanding the proposed models.

## THE GENERAL LCM

Early forms of the LCM came from both the factor analysis tradition in psychology (Tucker, 1958) and from the biostatistics literature (Rao, 1958). These traditions

were motivated by the idea of determining an underlying true curve representing the development of a construct. This was the weighted sum of common basis curves, which can be thought of as the mean curve or function. Researchers also recognized that individual differences were an important aspect that needed to be explicitly accounted for and they are represented as weights on the basis curves that freely vary by individual. The contemporary LCM was first proposed by Meredith and Tisak (1984, 1990) and expanded by McArdle (1989), Browne and Du Toit (1991), B. Muthén and Shedden (1999), and many others. The contemporary formulation of the LCM as described by Bollen and Curran (2006) was used as the reference here.

The LCM, like all structural equation models, contains a measurement model relating the observed variables to the latent variables, and a structural model that relates the latent variables to one another as well as to other observed variables. The measurement model for the LCM is

$$Y_i = \nu_i + \Lambda_y \eta_i + \varepsilon_i,$$

where  $Y_i$  is a vector of measurements from an individual over time,  $\nu$  is a vector of intercepts,  $\Lambda_y$  is a matrix of factor loadings,  $\eta_i$  is a vector of latent variable means, and  $\varepsilon_i$  is a vector of residual variances. To capture the mean structure of  $Y_i$ , the intercepts are set to zero, thereby passing the mean structure to the structural model (described later). The factor loadings in  $\Lambda_y$  take on fixed values structured to reflect the time course under study. The residual variances in  $\varepsilon_i$  refer to the combination of measurement error and the time-specific residual between the observed value and a model-implied trajectory. The structural model for the LCM is

$$\eta_i = \mu_\eta + \zeta_i,$$

where  $\eta_i$  is a vector of latent variable true scores;  $\mu_\eta$  is a vector of latent variable means; and  $\zeta_i$  is a vector of disturbances. The important parameters for these models are the means of the LCM trajectory parameters,  $\mu_\eta$ , and the variances,  $\text{VAR}(\zeta_i)$ , contained in the parameter covariance matrix  $\Psi_{\eta\eta}$ .

The LCM can be easily expanded to include exogenous predictors to explain interindividual differences in intraindividual change. Conditional models are accommodated by extending the structural model described earlier

$$\eta_i = \mu_\eta + \Gamma Z_i + \zeta_i,$$

where  $\Gamma$  is a vector of regression coefficients for vector  $Z_i$  of exogenous variables. Furthermore, the predictors that are exogenous to the LCM can be endogenous to another model, such that they are latent variables themselves.

## MODIFICATIONS TO THE LCM FOR PRE-POST-POST DESIGNS

We propose adapting the LCM to accommodate the analysis of cortisol data that are collected using pre-post-post designs. The unconditional and conditional models considered in detail are depicted in Figures 1 and 2, respectively. The measurement model for Figure 1 is

$$\begin{bmatrix} cort_1 \\ cort_2 \\ cort_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} baseline \\ react \\ reg \end{bmatrix},$$

and the structural representation is

$$\begin{bmatrix} baseline \\ react \\ cort \end{bmatrix} = \begin{bmatrix} \mu_{base} \\ \mu_{react} \\ \mu_{reg} \end{bmatrix} + \begin{bmatrix} \Psi_{base}^2 & & \\ \Psi_{base,react} & \Psi_{react}^2 & \\ \Psi_{base,reg} & \Psi_{reg,react} & \Psi_{reg}^2 \end{bmatrix}$$

This parameterization is equivalent to a piecewise linear model approach where each piece consists of only two time points. This parameterization is also equivalent to a simple difference score approach where reactivity and regulation refer to the simple differences between cortisol values obtained at Times 2 and 1, and Times 3 and 1, respectively. The key aspects of the model are parameters that directly correspond to (a) average levels of baseline cortisol level, as well as cortisol reactivity and regulation (i.e., means of latent variables,  $\mu_\eta$ ); (b) individual differences in baseline cortisol, as well as cortisol reactivity and regulation (i.e., vari-

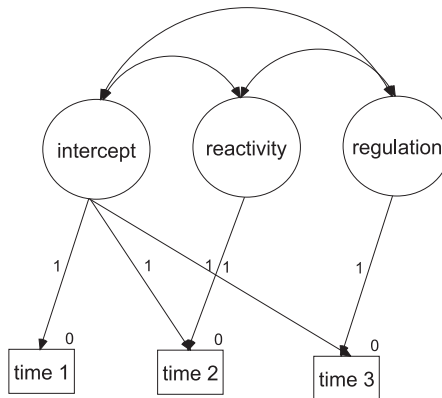


FIGURE 1 Unconditional model.

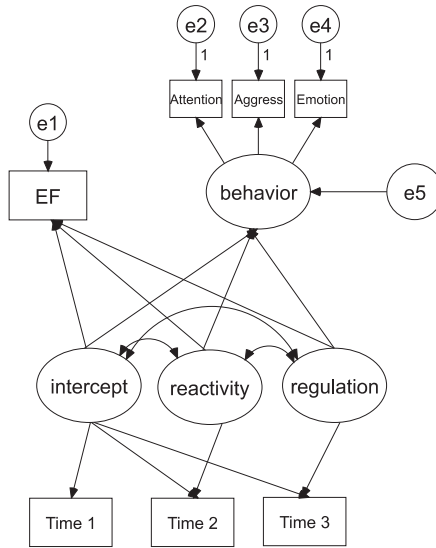


FIGURE 2 Conditional model.

ances of latent variables,  $\zeta_{\eta}$ ); and (c) indexes of the concordance among cortisol level, reactivity, and regulation across persons (i.e., covariances of latent variables,  $\Psi_{\eta\eta}$ ).

The conditional model considered in this study is depicted in Figure 2. This model tests whether interindividual differences in baseline levels of cortisol, as well as cortisol reactivity and regulation, are predictive of executive functioning (EF) and behavioral functioning. The measurement and structural models for the pre–post–post model remain the same. The matrix expression of the measurement model of behavior is

$$\begin{bmatrix} att \\ agg \\ emo \end{bmatrix} = \begin{bmatrix} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \end{bmatrix} [beh] + \begin{bmatrix} e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

and the structural model is

$$[beh] = [\mu_{beh}] + \begin{bmatrix} \gamma_{base} \\ \gamma_{react} \\ \gamma_{reg} \end{bmatrix} \begin{bmatrix} \mu_{base} & \mu_{react} & \mu_{reg} \end{bmatrix} + [e_5],$$

$$[EF] = [\mu_{EF}] + \begin{bmatrix} \gamma_{base} \\ \gamma_{react} \\ \gamma_{reg} \end{bmatrix} [\mu_{base} \quad \mu_{react} \quad \mu_{reg}] + [e_1].$$

Finally, the LCM can be extended to consider multiple groups. That is, separate measurement and structural models can be estimated for two or more groups simultaneously. The corresponding measurement and structural models are denoted as

$$Y_i = \nu^g + \Lambda_y^g \eta_i + \varepsilon_i$$

and

$$\eta_i = \mu_{\eta}^g + \zeta_i.$$

Superscripting parameter vectors with  $g$  signifies that separate parameters are being estimated for each group ( $g_1, g_2 \dots G$ ). The ability to impose equality constraints on estimated parameters across groups provides a means of testing for cross-group similarities and differences in substantively interesting parameters. The multiple-group approach is used to test whether relations between cortisol activity and EF and behavioral functioning are moderated by child gender.

This study demonstrates how framing questions regarding cortisol level, reactivity, and regulation from an SEM perspective can expand the scope and improve the rigor of questions that applied researchers routinely ask in at least five ways. First, unlike the unconditional model in Figure 1, the conditional model in Figure 2 is overidentified, which provides the ability to test overall model fit prior to interpretation of parameter estimates. We establish good fitting models prior to interpreting the significance and magnitude of cortisol reactivity and regulation parameters. Second, through the use of (robust) full information maximum likelihood (FIML) estimation methods, this SEM analysis provides a flexible method for handling missing data, as well as the inclusion of nonnormal outcome data. We demonstrate how up to 25% more observations are used than would be the case if listwise deletion methods were used. Third, the provision of latent versus manifest variables provides a strategy for taking measurement error into account. We are interested in relations between cortisol reactivity and regulation and the construct of behavioral competence, not specific indicators (attention, aggression, emotional competence) of the construct, per se. Fourth, the use of indirect effects facilitates formal tests of mediation. We test whether the association between cortisol reactivity and regulation and behavioral competence is mediated through EF. Fifth, the use of multiple group models provides a formal test of whether the relations between cortisol reactivity or regulation and EF and behavioral competence are moderated by child gender.

## METHOD

### Participants

The sample included 171 children attending Head Start programs predominantly serving White families living in rural and nonurban locations. Mean age of the children at the time of testing was 5 years, 1 month (range = 3 years, 9 months–5 years, 8 months). Eighty of the children were girls and 91 were boys. All children were either from households in which family income was below the poverty line as defined by federal poverty thresholds, or were eligible for Head Start due to the fact that in the absence of child care, family income would be below the poverty line. Fifty-nine percent of the sample reported highest level of education as a high school diploma or equivalent, 23% reported some education beyond high school, and 14% reported less than a high school education (4% did not list their educational level). The majority of children, 71%, lived in a home with two adults and 88% of the sample had one or more sibling in the household.

### Procedure

Children were seen individually in two 45-min sessions either in the morning or in the afternoon. Due to the constraints of the Head Start day, no morning session began prior to 9:00 a.m. and no afternoon session began later than 2:30 p.m. Children were seen in a quiet testing area at the Head Start center. During one of the sessions, the first session for approximately two thirds of the children participating in the study, three saliva samples were collected to assess cortisol reactivity and regulation over the course of the testing session.

During the session in which saliva samples were collected, children were also administered a measure of receptive vocabulary, an attention-shifting measure of EF, and a measure of letter knowledge. In the second session, a peg-tapping measure of EF was collected along with measures of false-belief understanding and emotion knowledge. All measures were administered in a standard order.

### Measures

*Executive function: Flexible Item Selection Task.* Children were also administered the Flexible Item Selection Task (FIST; Jacques & Zelazo, 2001). In this task, children are presented with pictures of three items that vary along some combination of two of three dimensions, including size, shape, and color. Following a pretest, children are presented with 16 trials in which they are instructed to point two objects that go together in one way. Children are then instructed to point to two objects that go together in another way. The task requires children to identify two of the three objects that are similar along one dimension (i.e., shape) but

then to shift set and identify two of the three objects that are similar along a second dimension (i.e., size). Jacques and Zelazo (2001) extensively investigated the task in a cross-sectional sample of children between the ages of 2 and 5 years and found that results converge well with those of similar dimensional shift measures. The number of correct responses divided by the total number of trials was used as a measure of performance on the task.

*Saliva collection and cortisol assay.* Children were asked to mouth with assistance from the investigator, a 6-in. sterile cotton rope, the saturated end of which was cut and placed into a needleless 10 cc plastic syringe, expressed into a plastic vial, and stored in a portable cooler. Samples were frozen for temporary storage prior to assay. All samples were assayed in duplicate for cortisol using a 510k cleared high-sensitive enzyme immunoassay specifically designed for use with saliva to measure adrenal function (Salimetrics, LLC, State College, PA). The test uses 25  $\mu$ l of saliva (for singlet determinations), has a lower limit of sensitivity of .007  $\mu$ g/dl, range of sensitivity from .007 to 1.8  $\mu$ g/dl, and average intra- and interassay coefficients of variation less than 15% and 10%, respectively. Following recommendations of Schwartz, Granger, Susman, Gunnar, and Laird (1998), sample pH was screened prior to assay. The criterion for agreement between duplicates was no more than 7% error, and the average of the duplicate tests was used in all analyses. All cortisol values were rescaled (multiplied by 100) to avoid working with extremely small numbers.

*Child classroom behavior.* Teachers reported on child attention and behavior using a preschool version of the Teacher Observation of Classroom Adaptation—Revised (TOCA-R; Werthamer-Larrson, Kellam, & Wheeler, 1991). Teachers rate on a scale from 1 to 6 the extent to which they think a particular item is true for a given child. The attention scale is composed of seven items including items relating to working hard, concentrating, staying on-task, paying attention, working at an age-appropriate level, not being distractible, and staying focused. The aggressive behavior scale is composed of five items including items relating to taking others' property, yelling at others, fighting, teasing classmates, and harming others. The social and emotional competence scale is composed of seven items relating to showing empathy, turn taking, listening, sharing, initiating interactions in a positive manner, recognizing and labeling feelings, and ease in talking to the teacher. These scales were identified through confirmatory factor analysis of data from a sample of 378 children. The internal consistency reliability of these scales is .95 for the attention items, .93 for the aggressive behavior items, and .92 for the social and emotional competence items (Kam & Greenberg, 2001).

## Missing Data and Outliers

In total data were collected on 171 children. Univariate analyses identified 3 cases with abnormally high cortisol values, which is indicative of a fever or illness, who were excluded from all analyses. Of the remaining 168 cases, 126 (75%) had complete cortisol data (three of three observations), and an additional 28 (17%) had one or two observations of cortisol. Fourteen children (8%) had outcome data but no cortisol data. Thus, whereas listwise deletion methods would result in a 17% to 25% reduction in sample size (depending on whether models were based exclusively on cortisol vs. cortisol plus outcome data), the FIML estimation methods utilized later make full use of all available data. The ability to utilize all available data is a major contribution of the SEM approach described here. General discussions of FIML estimation in the context of SEM models were provided elsewhere (Arbuckle, 1996; Schafer & Graham, 2002). All SEM analyses were completed using *Mplus* (version 3.13; Muthén & Muthén, 2004).

## RESULTS

Descriptive statistics for cortisol, covariate, and outcome data are provided in Tables 1 and 2. The variables labeled reactivity and regulation in these tables are simple difference scores (reactivity = Time 2 – Time 1; regulation = Time 3 – Time 1) that characterize changes in cortisol across three assessments. On average, there was a negligible reduction in cortisol between the first and second time assess-

TABLE 1  
Descriptive Statistics

<i>Cortisol</i>	<i>All Available Data</i>					<i>Complete Cortisol Data</i>				
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Skew</i>	<i>Kurtosis</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Skew</i>	<i>Kurtosis</i>
Time 1	146	11.29	5.91	1.95	6.89	126	11.08	5.62	2.09	9.23
Time 2	144	11.21	5.88	1.18	1.24	126	10.90	5.59	1.15	1.27
Time 3	140	10.14	5.26	1.33	2.21	126	9.70	4.56	1.11	1.17
Reactivity	137	-.08	4.93	-0.62	7.41	126	-.18	4.65	-1.09	9.36
Regulation	133	-1.31	3.81	-0.58	5.49	126	-1.38	5.50	-1.71	11.43
FIST	146	0.69	.25	-1.55	2.15	—	—	—	—	—
Aggress	166	6.44	5.29	0.93	0.48	—	—	—	—	—
Attention	161	22.86	7.43	-0.03	-1.01	—	—	—	—	—
Emotion	168	22.49	6.84	0.01	-11.09	—	—	—	—	—
Income	152	1.66	1.08	1.89	5.17	—	—	—	—	—
Age	168	5.08	0.41	-0.98	1.77	—	—	—	—	—
Male	168	0.53	0.50	-0.12	-2.01	—	—	—	—	—

*Note.* FIST = Flexible Item Selection Task.

TABLE 2  
Pairwise Correlations

	1	2	3	4	5	6	7	8	9	10
1. Cortisol—Time 1	1.0									
2. Cortisol—Time 2	.62438	1.0								
3. Cortisol—Time 3	.47145	.76145	1.0							
4. FIST	.05062	.06787	-.05205	1.0						
5. Aggression	.14170	.01715	-.02772	-.20320	1.0					
6. Attention	-.07824	-.06658	-.04384	.26853	-.56878	1.0				
7. Emotion	-.02779	.00478	-.00597	.29952	-.59525	.74160	1.0			
8. Income	-.46229	-.09983	-.14617	.06677	-.11601	.11783	.06428	1.0		
9. Age	.00494	-.09396	.02166	.07890	-.03512	.14983	.17378	.04426	1.0	
10. Male	.00706	.03861	.10293	-.29854	.21490	-.29460	-.31996	.02146	-.04360	1.0

*Note.* *N*s = 131–168; FIST = Flexible Item Selection Task.

ments and a small additional decrease between the second and third samples. As shown in Table 1, the exact amount of decrease in cortisol between successive samples depends on whether difference scores are computed using all available versus complete (i.e., listwise deletion) methods (e.g., reactivity range from  $-.08$  to  $-.18$ ). Although on average there were decreasing levels of cortisol over time, there are substantial individual differences in patterns of change between successive cortisol samples (note that the standard deviations are far greater than the means for cortisol reactivity and regulation). Of theoretical interest was whether children who show increased levels of cortisol between first and second samples (i.e., reactivity scores  $> 0$ ) and decreased levels of cortisol between the second and third samples (i.e., regulation scores  $< 0$ ) exhibit optimal EF and improved behavioral competence.

### Unconditional Models

The model depicted in Figure 1 was fit to the observed data using both listwise deletion and FIML estimation methods. As noted earlier, all of the unconditional models are just identified and hence have no formal index of overall model fit. Nonetheless, as seen in Table 3, the model implied cortisol values (cortisol level, reactivity, regulation) from the unconditional SEM model are virtually identical to the corresponding raw data, when both are based on complete data ( $N = 126$ ). The model implied values are very close, although not identical, to observed data when the FIML estimator is used (e.g., note that the point estimate for reactivity is now greater than 0 in the FIML model but less than 0 in the listwise deletion model). This is due to the fact that the FIML estimator uses all available information ( $N = 154$ ). That is, the FIML analysis includes participants who have one or more cortisol values whereas the listwise deletion analysis includes participants who have all three cortisol values.

The primary goals of the unconditional model are to establish (a) the average values for cortisol (baseline) level, reactivity, and regulation (i.e., latent means); (b) to determine whether there is significant variation around these average values (i.e., latent variances); and (c) to determine the interrelations between cortisol level, reactivity, and regulation scores (i.e., latent covariances). The means of the latent variables that corresponded to cortisol level, reactivity, and regulation were 11.452 ( $z = 22.7, p < .001$ ), 0.087 ( $z = 0.2, p = .84$ ), and  $-1.253$  ( $z = -2.52, p < .05$ ), respectively. Thus, the mean cortisol value at Time 1 was 11.452 and was significantly different than 0. The mean difference between cortisol values at Times 1 and 2, defined here as reactivity, was .087. Although positive, this difference was not significantly different than 0. The mean difference between cortisol values at Times 1 and 3, defined here as regulation, was  $-1.253$ . This value was negative and significantly different than 0.

TABLE 3  
Recovery of Mean, Variances, and Correlations Between Cortisol Measures Using Unconditional SEM Model

	<i>Raw Data (N = 126)</i>			<i>SEM (Listwise, N = 126)</i>			<i>SEM (FIML, N = 154)</i>		
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>
1. Cortisol (initial)	1.0			1.0			1.0		
2. Cortisol reactivity	-.420	1.0		-.420	1.0		-.356	1.0	
3. Cortisol regulation	-.664	0.742	1.0	-.664	0.742	1.0	-.613	0.763	1.0
<i>M</i>	11.082	-0.183	-1.381	11.082	-0.183	-1.381	11.452	0.087	-1.253
Variance	31.555	21.601	30.250	31.305	21.430	30.010	36.797	25.321	33.526

*Note.* FIML = full information maximum likelihood.

A principal advantage of this approach is the simultaneous consideration of latent means and variances. The variances of the latent variables that corresponded to cortisol level, reactivity, and regulation were 36.80 ( $z = 4.2, p < .001$ ), 25.32 ( $z = 4.1, p < .001$ ), and 33.53 ( $z = 3.67, p < .001$ ), respectively. Thus, there was significant interindividual variability in the initial level of cortisol, as well as in the magnitude of change in cortisol between Times 2 and 1 (i.e., reactivity) and Times 3 and 1 (i.e., regulation).

Finally, the latent correlations between cortisol level with reactivity and regulation were  $-.36$  ( $z = -1.63, p = .10$ ) and  $-.61$  ( $z = -2.58, p < .01$ ), respectively. Thus, baseline levels of cortisol were significantly related to the magnitude of regulation but not reactivity (e.g., individuals with higher baseline cortisol scores exhibited less regulation). The latent correlation between reactivity and regulation was  $.76$  ( $z = 3.17, p < .01$ ). Thus, the magnitude of reactivity scores were significantly and positively related with the magnitude of regulation scores (e.g., individuals who showed greater amounts of change between Times 1 and 2 also showed greater amounts of change between Times 1 and 3).

### Conditional Models

The results of the unconditional models demonstrated that there was significant variability both in baseline levels of cortisol and in cortisol reactivity and regulation. The primary goals of the conditional models were to relate variability in cortisol level, reactivity, and regulation to outcome variables, including a single manifest indicator of child EF and a latent variable of behavior problems (see Figure 2). Consistent with Blair et al. (2005), cortisol (level, reactivity, and regulation), EF, and child behavior were simultaneously regressed on covariates including child gender, age, and family income. These covariate paths are omitted from Figure 2 to simplify the visual presentation. A total of three variants of this model were fit to the observed data. A synopsis of model description and fit is provided in Table 4.

The primary goal of the first conditional model was to test whether interindividual differences in cortisol level, reactivity, and regulation were predictive of children's EF scores, teacher-rated behavioral competence, or both. The first conditional model fit the observed data well,  $\chi^2(15, N = 168) = 16.6, p = .35$ . The emotional subscale of the TOCA was used as an indicator of the latent variable representing behavioral competence. The factor loadings for aggression and attention were significant and in the expected direction, and the residual and latent variances were significantly different than zero. With respect to covariates, child age and gender were significantly related to the behavioral competence latent variable (younger children and boys exhibited less competent behavior), and child gender was also significantly related to EF (boys performed poorer on the FIST task relative to girls). The covariates were unrelated to all of the cortisol latent variables

TABLE 4  
General SEM Model Fit and Description

<i>Model Number</i>	<i>Model Description/Change From Baseline</i>	$\chi^2$	<i>df</i>	<i>p</i>	<i>CFI</i>	<i>RMSEA</i>	<i>90CI</i>	
1	Direct effects (cortisol → EF, behavior)	16.565	15	.3455	.996	.025	.000	.079
2	Indirect effects (cortisol → EF → behavior)	10.037	14	.7594	1.000	.000	.000	.053
3	Multiple Group (gender; no constraints)	31.789	28	.2832	.990	.040	.000	.097
4	Multiple Group (gender; constraints)	35.534	35	.4431	.999	.013	.000	.080

*Note.* *N* = 168 for all models; CFI = comparative fit index; RMSEA = root mean square error of approximation; 90CI = 90% confidence intervals; EF = executive functioning.

(level, reactivity, and regulation). With respect to substantive paths of interest, there was a trend for elevated cortisol reactivity scores to be associated with more socially competent behavior ( $z = 1.84, p = .06, \beta = .26$ ). Both cortisol reactivity and regulation scores were significantly related to EF. Specifically, greater cortisol reactivity ( $z = 2.00, p < .05, \beta = .30$ ) and cortisol regulation ( $z = -2.03, p < .05, \beta = -.34$ ) were both associated with improved EF performance. Although the initial level of cortisol was unrelated to EF and behavioral outcomes, consistent with the unconditional model, it was significantly and negatively intercorrelated with cortisol reactivity ( $\phi = -.34$ ) and regulation ( $\phi = -.61$ ).

The first model did not allow a relation between the manifest indicator of EF and the latent variable of behavioral competence. The primary goal of the second conditional model was to test whether interindividual differences in cortisol were predictive of child behavioral competence indirectly, through EF. To test this, the second model introduced a direct path from EF to the latent variable of behavioral competence. In addition to considering the significance of this main effect, the indirect effects from cortisol reactivity and regulation to behavioral competence via EF were also considered. Evidence of significant indirect effects (literally a test of the product of the coefficients linking reactivity and regulation to EF and EF to behavioral competence) provided an explicit test of mediation. The second model fit the data well,  $\chi^2(14, N = 168) = 1.04, p = .76$ . Four results are noteworthy. First, the direct effects of cortisol reactivity ( $z = 2.2, p < .05, \beta = .32$ ) and regulation ( $z = -2.1, p < .05, \beta = -.36$ ) on EF continued to be statistically significant. Second, EF was a significant predictor of behavioral competence ( $z = 2.6, p < .05, \beta = .24$ ). The addition of EF as a predictor resulted in an increased  $R^2$  for the behavioral competence latent variable from .17 in the previous model to .23 in the current model. The inclusion of this path was also the reason for improved overall model fit,  $\chi^2$  difference (1) = 6.528,  $p = .01$ . Third, the previously significant direct effect from cortisol reactivity to behavioral competence was no longer statistically significant ( $z = 1.2, p = .12, \beta = .17$ ), which is consistent with, although not a direct test of, mediation. Fourth, the test of the indirect effects from cortisol reactivity ( $z = 1.6, p = .11, \beta = .08$ ) and regulation ( $z = -1.6, p = -.11, \beta = -.08$ ) through EF in the prediction of behavioral competence were not significant. Hence, the formal test of mediation was not significant.

The first and second conditional models used child age and gender and family income as covariates. The primary goal of the last set of conditional models was to test whether child gender moderated the relation among cortisol level, reactivity, and regulation in the prediction of EF and behavioral competence, as well as the prediction of behavioral competence from EF. This was accomplished using a multiple groups approach. Specifically, a third conditional model was estimated. The third model was similar to the second model except that child gender was no longer used as a covariate. Rather, the second model was estimated simultaneously for boys and girls, without imposing any equality constraints on parameters across

groups. The third conditional model fit the observed data well,  $\chi^2(28, N = 168) = 31.79, p = .28$ . This model was then reestimated simultaneously across gender groups while imposing equality constraints on the regression paths from cortisol level, reactivity, and regulation in the prediction of EF and behavioral competence, as well as the path relating behavioral competence to EF (a total of seven constraints). This fourth conditional model also fit the data well,  $\chi^2(35, N = 168) = 35.53, p = .44$ . Because the fourth model was nested within the third model, a chi-square difference test was to determine whether the imposition of equality constraints degraded model fit. It did not,  $\chi^2$  difference (7) = 3.745,  $p = .81$ , which indicated that child gender did not moderate the relation among cortisol, EF, and behavioral competence.

## DISCUSSION

There is a well-established literature demonstrating that optimal patterns of up-followed by down-regulation of cortisol (i.e., cortisol reactivity and regulation) in response to stressors are associated with adaptive functioning. Although, on average, this pattern of effects holds true, there is a growing interest in delineating individual differences in patterns of intraindividual change in cortisol (Granger & Kivlighan, 2003; Ramsay & Lewis, 2003). This study introduced and developed, by way of an extended example, a novel structural equation model that could accommodate this goal, specifically for cortisol data that are collected using pre–post–post designs. The results of this study demonstrated that, after controlling for individual differences in child gender, age, and family income, cortisol reactivity and regulation were independently predictive of executive and behavioral functioning, that the relation between cortisol reactivity and regulation and behavioral competence was not mediated through EF, and that these relations were not moderated by child gender. These conclusions are generally consistent with those drawn from a previous analysis of these data but resulted from a model that is more closely aligned with motivating questions (Blair et al., 2005).

Building on the overall strengths of SEM, the models described in this study permit tests of overall model fit prior to interpretation of parameter estimates, accommodate missing data, nonnormal data, or both; accommodate measurement error in both predictors and outcome variables; and can explicitly test questions involving mediation and moderation. There are a variety of ways that the models introduced here could be extended. First, the proposed models can easily be extended to include parallel processes. Parallel processes mean simultaneous changes in two or more sets of repeated measures data. For example, elsewhere one of us has emphasized the theoretical value of considering parallel changes in cortisol and alpha-amylase, which is a separate hormone that is also related to the stress response system (Granger et al., 2006). Parallel process models could also

be extended to dyadic data. For example, to the extent that caregivers and infants are exposed to some common stressor, it may be of interest to examine relations between caregiver and infant cortisol reactivity and regulation. Such analyses would inform questions regarding caregiver–child dyad coupling and synchrony (Woody & Sadler, 2005).

Second, recent developments in SEM involve the integration of categorical and continuous latent variables into a single modeling framework (Arminger & Stein, 1997; Arminger, Stein, & Wittenberg, 1999; Jedidi, Jagpal, & DeSarbo, 1997; Muthén, 2001; Yung, 1997). The inclusion of categorical latent variables into general SEMs is conceptually similar to multiple group models, except that group membership is unknown (Muthén & Shedden, 1999). Expanding the models demonstrated here to include categorical latent variables may help inform whether the hypothesized relations among cortisol reactivity and regulation, EF, and behavior competence are specific to only a subset of the sample. Creating profiles of children based both on their patterns of cortisol values over time, as well as the relation between patterns and various outcomes, has the potential to further enrich questions that have thus far been largely addressed using variable-oriented approaches (Muthén & Muthén, 2000).

Third, to the extent that cortisol reactivity and regulation data are collected in longitudinal designs, the single-indicator latent variables designating cortisol level, reactivity, and regulation, could be construed as indicators of a subsequent LCM. Such models have variously been referred to as second-order and multiple indicator latent growth models (Chan, 1998; Sayer & Cumsille, 2001).

There are limitations to the models described here, too. All of the models that were considered in this study were predicated on an assumption that researchers are unable to collect more than three samples of cortisol around the experience of some novel or stressful event. This is due to real-world limitations involving the collection of saliva samples in the context of larger studies for which cortisol data are only one part. Indeed, the goal of the models considered here was to make optimal use of limited information contained in three samples of cortisol. However, design changes involving the collection of more frequent samples of cortisol across the time period under consideration would provide a far better means for testing whether and how interindividual differences in cortisol reactivity and regulation are related to various outcomes (Willet, 1997).

Second, although the value of using a FIML estimation method was emphasized for dealing with missing data, this is only the case when it is reasonable to assume that data are missing at random. However, one can easily imagine situations where this assumption will not be valid. For example, studies of infant cortisol response following immunization shots will invariably have some missing data due to children falling asleep as a regulation strategy (Ramsay & Lewis, 1994). In these cases, the reasons for missing data are likely nonignorable, rendering this treatment of missing data inappropriate.

Third, it is generally understood that SEM models require large samples. Sample size requirements are related to the reliance on asymptotic theory (assuming that the sample mean vector and covariance matrix are consistent, unbiased, and efficient estimates of the population mean vector and covariance matrix), the type of estimation algorithm that is used, the measurement characteristics of the data (e.g., effect sizes, reliability), and the general complexity of the models being estimated (Bollen, 1989; Klein, 2005). There are no good rules of thumb for determining a minimally sufficient sample size (MacCallum, Widaman, Zhang, & Hong, 1999). Although as many as 20% of recently published papers using SEM methods have involved samples of fewer than 100 cases, with samples this small technical problems are more likely to arise (e.g., nonpositive definite covariance matrices, Heywood cases), the power to detect effects (including the ability to reject overall poor model fit) may be compromised, and the precision of parameter estimates is unknown (MacCallum & Austin, 2000). Researchers interested in applying the model described here to small samples may want to consider alternative estimation algorithms that may help guard against some of the potential problems noted earlier (e.g., two-stage least squares; Bollen, 2001) and should consider keeping models as simple as possible.

In sum, an SEM model that makes novel use of the parameterization of LCMs was introduced for the analysis of cortisol data that are collected using a pre–post–post design. The model includes parameters for means, variances, and covariances of cortisol level, reactivity, and regulation. This model facilitates a closer correspondence between the types of questions that substantive researchers ask and the methods that they use to empirically answer these questions.

## REFERENCES

- Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling* (pp. 243–277). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Arminger, G., & Stein, P. (1997). Finite mixtures of covariance structure models with regressors. *Sociological Methods and Research*, 26, 148–182.
- Arminger, G., Stein, P., & Wittenberg, J. (1999). Mixtures of conditional mean and covariance structure models. *Psychometrika*, 64, 475–494.
- Blair, C., Granger, D., & Razza, R. P. (2005). Cortisol reactivity is positively related to executive function in preschool children attending Head Start. *Child Development*, 76, 554–567.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bollen, K. A. (2001). Two-stage least squares and latent variable models: Simultaneous estimation and robustness to misspecifications. In R. Cudeck, S. Du Toit, & D. Sörbom (Eds.), *Structural equation modeling: Present and future* (pp. 119–138). Lincolnwood, IL: Scientific Software.
- Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley.

- Browne, M. W., & Du Toit, S. H. C. (1991). Models for learning data. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 47–68). Washington, DC: American Psychological Association.
- Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach for incorporating longitudinal mean and covariance structures analysis (LMACS) and multiple indicator latent growth modeling (MLGM). *Organizational Research Methods, 1*, 421–483.
- Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin, 130*, 355–391.
- Erickson, K., Drevets, W., & Schulkin, J. (2003). Glucocorticoid regulation of diverse cognitive functions in normal and pathological emotional states. *Neuroscience and Biobehavioral Reviews, 27*, 233–246.
- Granger, D., & Kivlighan, K. T. (2003). Integrating biological, behavioral, and social levels of analysis in early child development: Progress, problems, and prospects. *Child Development, 74*, 1058–1063.
- Granger, D. A., Kivlighan, K. T., Blair, C., El-Sheikh, M., Mize, J., Lisonbee, J. A., et al. (2006). Integrating the measurement of salivary alpha-amylase into studies of child health, development, and social relationships. *Journal of Social and Personal Relationships, 23*, 267–290.
- Granger, D., Stansbury, K., & Henker, B. (1994). Preschoolers' behavioral and neuroendocrine responses to social challenge. *Merrill Palmer Quarterly, 40*, 190–211.
- Gunnar, M., & White, B. P. (2001). Salivary cortisol measures in infant and child assessment. In L. T. Singer & P. S. Zeskind (Eds.), *Biobehavioral assessment of the infant* (pp. 167–189). New York: Guilford.
- Heim, C., Ehlert, U., & Hellhammer, D. H. (2000). The potential role of hypocortisolism in the pathophysiology of stress-related bodily disorders. *Psychoneuroendocrinology, 25*, 1–35.
- Jacques, S., & Zelazo, P. D. (2001). The Flexible Item Selection Task (FIST): A measure of executive function in preschoolers. *Developmental Neuropsychology, 20*, 573–591.
- Jedidi, K., Jagpal, H., & DeSarbo, W. S. (1997). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science, 16*, 39–59.
- Kam, C., & Greenberg, M. (2001). *A version of the teacher observation of classroom adaptation for use with preschoolers*. State College: Pennsylvania State University.
- Klein, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford.
- Lupien, S. J., & McEwen, B. S. (1997). The acute effects of corticosteroids on cognition: Integration of animal and human model studies. *Brain Research Reviews, 24*, 1–27.
- MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51*, 201–226.
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 84–99.
- McArdle, J. J. (1989). A structural modeling experiment with multiple growth functions. In R. Kanfer, E. Phillip, L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology: The Minnesota Symposium on Learning and Individual Differences* (pp. 71–117). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- McEwen, B. S. (2000). The neurobiology of stress: From serendipity to clinical relevance. *Brain Research, 886*, 172–189.
- Meredith, W., & Tisak, J. (1984, July). *Tuckerizing curves*. Paper presented at the annual meeting of the Psychometric Society, Santa Barbara, CA.
- Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*, 107–122.
- Muthén, B. O. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In L. M. Collins & A. Sayer (Eds.), *New methods for the analysis of change* (pp. 291–322). Washington, DC: American Psychological Association.

- Muthén, B. O., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, *24*, 882–891.
- Muthén, B. O., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, *55*, 463–469.
- Muthén, L. K., & Muthén, B. O. (2004). *Mplus users guide* (3rd ed.). Los Angeles: Author.
- Ramsay, D. S., & Lewis, M. (1994). Developmental change in infant cortisol and behavioral response to inoculation. *Child Development*, *65*, 1491–1502.
- Ramsay, D., & Lewis, M. (2003). Reactivity and regulation in cortisol and behavioral responses to stress. *Child Development*, *74*, 456–464.
- Rao, C. R. (1958). Some statistical methods for comparison of growth curves. *Biometrics*, *14*, 1–17.
- Sayer, A. G., & Cumsille, P. E. (2001). Second-order latent growth models. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 179–201). Washington, DC: American Psychological Association.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*, 147–177.
- Schwartz, E., Granger, D., Susman, E. J., Gunnar, M., & Laird, B. (1998). Assessing salivary cortisol in studies of child development. *Child Development*, *69*, 1503–1513.
- Tucker, L. R. (1958). Determination of parameters of a functional relation by factor analysis. *Psychometrika*, *23*, 19–23.
- Werthamer-Larrson, L., Kellam, S. G., & Wheeler, L. (1991). Effect of first grade classroom environment on shy behavior, aggressive behavior, and concentration problems. *American Journal of Community Psychology*, *19*, 890–897.
- Willet, J. B. (1997). Measuring change: What individual growth modeling buys you. In E. Amsel & K. A. Renninger (Eds.), *Change and development: Issues of theory, method, and application* (pp. 213–243). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Woody, E., & Sadler, P. (2005). Structural equation models for interchangeable dyads: Being the same makes a difference. *Psychological Methods*, *10*, 139–158.
- Yung, Y. (1997). Finite mixtures in confirmatory factor analysis models. *Psychometrika*, *62*, 297–330.