# A Correlated Traits Correlated Methods Model for Thin-Slice Child Personality Assessment

Jennifer L. Tackett[1], Jonas W. B. Lang[2], Kristian E. Markon[3], and Kathrin Herzhoff[1]

[1]Northwestern University

[2]Ghent University

[3]University of Iowa

**Manuscript in press, *Psychological Assessment***

Author Note

Jennifer L. Tackett, Department of Psychology, Northwestern University; Jonas W. B. Lang, Department of Personnel Management, Work and Organizational Psychology, Ghent University; Kristian E. Markon, Department of Psychology, University of Iowa; Kathrin Herzhoff, Department of Psychology, Northwestern University.

Support for this research was provided by funding from the Social Sciences and Humanities Research Council (SSHRC) of Canada awarded to Dr. Jennifer L. Tackett.

Correspondence concerning this manuscript should be directed to: Jennifer L. Tackett, Department of Psychology, Northwestern University, Evanston, IL, 60208. E-mail: jennifer.tackett@northwestern.edu

**version of the article. Please do not copy or cite without author's permission. The final article will be available, upon publication.**

Abstract

Recent research has illustrated the utility and accuracy of a thin-slice (TS) approach to child personality assessment, whereby unacquainted observers provide personality ratings of children after exposure to brief behavioral episodes. The current study sought to expand on this approach by exploring formal multitrait-multimethod (MTMM) models for child TS data comprising ratings from a comprehensive set of TS situations. Results using data from a sample of 326 9-to-10-year-old community children indicate that a correlated traits, correlated methods (CTCM) model can be used to represent individual differences in children's behavior as manifest across different situations. Indicator variables derived from a CTCM differentially correlate with traditional parental ratings of behavior, moreover, and provide predictive and incremental validity with regard to child competencies and behavior. Results illustrate the utility of a TS approach in the assessment of childhood personality, and inform understanding of issues encountered in applying different MTMM models to these types of empirical data.

Keywords: Multitrait multimethod matrix, thin slice, child personality, multiple informants

Public Significance Statement: Child personalities can be accurately observed by strangers exposed to only brief observations of the child’s behavior. The current study examined a statistical model aimed at integrating this type of data across multiple raters and multiple situations in which the children were observed.

.

# A Correlated Traits Correlated Methods Model for Thin-Slice Child Personality Assessment

Although under-represented in the broader field of personality research, research on child personality is alive and well (Soto & Tackett, 2015; Tackett & Durbin, 2017). A primary reason for this under-representation is the difficulty in measuring personality in childhood, a developmental period where self-reports are less common, more challenging to implement, and more psychometrically compromised. The field has largely progressed by emphasizing questionnaire methods, commonly completed by parents or teachers, supplemented by intensive laboratory-based protocols that are micro-coded for personality or temperament information (Tackett, Herzhoff, Kushner, & Rule, 2016). Each method has its advantages and drawbacks, which have been extensively discussed elsewhere (Durbin, 2010; Rothbart & Bates, 2006; Tackett et al., 2016). This eclectic measurement history has led to a common emphasis on collecting data from multiple informants and multiple methods, whenever possible. However, a multi-informant multi-method design typically yields more disagreement than agreement (De Los Reyes & Kazdin, 2005), with even highly similar and knowledgeable informants showing only moderate convergence for child personality traits (Clark, Durbin, Donnellan, & Neppl, 2017; Göllner et al., 2017; Luan, Hutteman, Denissen, Asendorpf, & van Aken, 2017; Tackett, 2011). The current paper sought to examine a variety of combinatorial modeling approaches to handling multi-informant, multimethod studies of child personality, with a focus on cross-situational thin-slice personality ratings (Tackett et al., 2016).

Given the limitations of relying solely on parent-report questionnaires to assess child personality, along with the resource-intensive nature of extensively micro-coded lab-based protocols (Durbin, 2010; Rothbart & Bates, 2006), a recent study sought to examine a complementary alternative to child personality assessment which we refer to as the “thin-slice” (TS) approach (Tackett et al., 2016). Unacquainted observers can provide accurate personality ratings of (adult) strangers after exposure to brief behavioral episodes, or “thin slices”, and the accuracy of these ratings improves upon aggregating 5-7 “thin-slice” situations (e.g., Allik, de Vries, & Realo, 2016; Borkenau, Mauer, Riemann, Spinath, & Angleitner, 2004). We recently extended this work to a large sample of 326 children, ages 8-12 years old, and established that unacquainted raters can provide valid and reliable information about child personality after only brief behavioral exposures, as well (Tackett et al., 2016). Children engaged in 15 different situational tasks in the lab, with each task rated by 3 unacquainted observers. The initial investigation of these data sought to examine psychometric properties of these ratings, correlations with parental report, and prediction of psychopathology and competence outcomes. In addition, these ratings were aggregated across all 3 raters per slice AND all 15 situations in the previous investigation, yielding a single TS personality score for each trait. The current study sought to build on this initial work by applying more sophisticated models of multi-trait multi-method data to: 1) examine the best way to model these types of multi-trait, multi-method, multi-informant child personality data and 2) provide useful recommendations (and the necessary stimulus materials) for researchers, clinicians, and educators who may be interested in applying thin-slice assessment of child personality in their own work.

**Combining multiple informants and methods**

In multi-trait, multi-method (MTMM) data, all methods are combined with all traits and vice versa to gain insight into the degree to which traits generalize across methods and methods bias is correlated between traits (Campbell & Fiske, 1959). Psychometricians have discussed alternative models for combining MTMM data for decades (e.g., Campbell & Fiske, 1959; Eid, Geiser, & Koch, 2016; Lance & Fan, 2016; Widaman, 1985). This literature focuses on the use of confirmatory factor analysis (CFA), which is particularly well suited for modeling MTMM data because of its flexibility in allowing researchers to specify, test, and compare loading patterns based on theory. Figures 1 and 2 illustrate some common CFA models for MTMM data.

The most common, and arguably most important, of these approaches is the correlated traits, correlated methods model (CTCM) shown in Figure 1. The CTCM model allows latent variables for both the traits and the methods. Each measurement indicator (observed variable) consequently loads both on its method and on its trait.

The CTCM structure is relatively restrictive because it assumes that variance in the indicators is only explained by the general variance for the trait and the method. That is, a trait like extraversion and a method (or situation) like sharing a sad memory only exist in the model as general factors. There is no factor that specifically accounts for extraversion in the restricted context of the method/situation; this specific type of variance would be considered error variance in the context of the CTCM. This feature of the CTCM is in line with the original concept behind the MTMM approach (see Campbell and Fiske, 1959) that suggests that traits and methods should generalize. Routinely, the CTCM is specified with uncorrelated traits with methods—in other words, trait factors are correlated with other trait factors and methods factors with other methods factors, but trait and methods factors are uncorrelated. This approach is often necessary to achieve model convergence (Lance, Noble, & Scullen, 2002; Widaman, 1985, 1992) and is also in line with the original ideas behind the MTMM approach (Campbell and Fiske, 1959) and is similar to the popular bi-factor approach in personality and ability assessment (Reise, Moore, & Haviland, 2010).

Research on MTMM data suggests that the CTCM model is the preferred model because it generally avoids biases in loading assumptions that might otherwise be necessary with MTMM data (Eid et al., 2016; Lance et al., 2002; Lance, Lambert, Gewin, Lievens, & Conway, 2004, also see the appendix for detail). However, a common problem in the literature is that

CTCM models for MTMM data often do not converge properly (Eid et al., 2016; Lance et al., 2002). A common recommendation in the literature is to fit the CTCM and use it when it converges and to only consider other models when the CTCM does not converge (Lance et al., 2002). Problems with non-fitting CTCM, however, are frequently tied to conceptual issues with a small number of methods or traits—a challenge which may be overcome by a comprehensive thin-slice battery (such as the one used here), in which data are available from many situations.

**Present Study**

Our primary goal in the present study was to examine a situational MTMM measurement approach to a comprehensive battery of child personality. More specifically, our study had three objectives. Our first objective was to study whether a situational MTMM approach has sufficient construct validity and reliability to be useful. We investigated this research question by fitting a series of MTMM CFA models to our data. Furthermore, we also estimated the reliability of measures based on our CFA factors using contemporary methods for estimating reliability in latent variable models. In interpreting our MTMM data, we were also interested in the degree to which various situations load on multi-method Big 5 factors, and whether certain situations may be more suited to observe certain personality traits than others. These questions are of conceptual interest for personality researchers and also relevant for informing future situational assessment approaches. Our second objective was to assess the degree to which latent traits estimated via a comprehensive MTMM model using TS child personality data are differentially associated with parent ratings, as parent ratings are often the default measurement approach in child personality research. Our approach is also an extension of earlier work on the relationship between observer ratings and thin slice assessments outside of a comprehensive modeling approach (Tackett et al., 2016). Finally, our third objective was to examine evidence for predictive and incremental validity of child competencies and psychopathology, extending findings from Tackett and colleagues (2016) by examining these questions in a comprehensive MTMM modeling framework.

# Methods

## Participants

Participants were 326 primarily 9-to-10-year-old children (52% female, $M_{age}$ = 9.96, SD = 0.81) and their parents. They were recruited through flyers posted in the community and a community participant pool maintained by the university. Exclusion criteria were intellectual disability, neurodevelopmental, or psychotic disorders in the child. Inclusion criteria were fluency in English for both child and parent. Informed consent was ascertained from the parents and verbal assent from children. The following ethnic breakdown for children was reported by their parents: 72% Caucasian, 17% Other/Multiracial, 8% Asian-Canadian, 3% African-Canadian, 1% Hispanic, 0.3% Pacific Islander, and 0.3% not reporting ethnicity. Using the Hollingshead (1975) Occupational Scale where a score of 9 indicates the highest occupational attainment, parental occupation was coded as *M* = 7.04, *SD* = 1.48; *range* = 3 – 9.

**Measures**

**Inventory of Children's Individual Differences – Short Form (ICID-S; Deal, Halverson, Martin, Victor, & Baker, 2007).** Children's Big Five personality was measured using the ICID-S, which thin slice coders and parents completed. The ICID-S consists of 50 items that were rated on a 7-point scale (1 = *much less than the average child or not at all* to 7 = *much more than the average child*; thin slice coders had the additional option of endorsing not observable, which was discouraged as a response choice). Parents' items were extracted from the 144-item long form of the ICID (Halverson et al., 2003) and all items were scored according to Halverson's item loadings with the exception of the lower-order openness facet items which were scored on the higher-order Openness factor. Cronbach's alpha ranged from .82 (mother-reported Agreeableness and father-reported Neuroticism) to .93 (thin-slice-coded Conscientiousness).

**Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001).** Children's competence in different areas of living as well as their problem behavior (psychopathology) was measured using the CBCL, which parents completed. The CBCL Competence scale consists of 16 items that were rated on varying scales (e.g., 0 = *less than average* to 3 = *more than average* or 0 = *less active* to 3 = *more active* etc.). The CBCL Internalizing and Externalizing scales consist of 34 and 35 items, respectively, that were rated on a 3-point scale (0 = *Not True (as far as you know)* to 2 = *Very True or Often True*). In accordance with Achenbach and Rescorla (2001), all items were summed into Total Competence, Internalizing, and Externalizing scale scores.

**Procedure**

Full procedures are described in Tackett et al. (2016). In brief, parents rated their child's personality at home and judges, who did not interact with the child, rated the target child's personality using the ICID-S based on 15 video-recorded tasks (see Table 1 for a brief description of each task). These video recordings were full-length (~1 min to 10 min long) clips of each task with sound that included a standard set of instructions given by the research assistant interacting with the child in a separate room from their parents. Most thin slice ratings were completed by research assistants in the lab (92%) whereas some were completed by undergraduates participating for course credit (9%; due to rounding error, the percentages do not add up to 100%). Thin slice coders were oriented to the protocol, which initially allowed for coders to watch up to three clips per child with at least one month in between coding (25% of the clips) but later allowed coders to watch only one clip per child. Ethics approval was obtained from the institutional review board and assent and consent were obtained from child and adult participants, respectively. Parents received $40 CAD and children two small gifts.

## Results

Analyses were conducted in R (R Core Team, 2016) using the lavaan package (Rosseel, 2012). Table 1 reports the standardized loading for each trait on each situation, and overall shows substantial factor loadings for the higher-order traits across all situations. Table 2 provides a complete correlation matrix between trait ratings in all 15 situations. Table 3 reports average trait correlation across distinct situations. These estimates suggest that, on average, traits assessed in one situation are moderately-highly associated with the same trait assessed in another situation, supporting validity of the ratings based on expectations of cross-situational trait generalizability.

**Objective 1: Structure of Situational MTMM Data**

Table 4 provides fit statistics for (a) a model with only traits (Model CT), (b) a model with only methods (Model CM), and (c) the CTCM. All three models converged, although the CT did so with a warning that the covariance matrix of the latent traits was not positive definite. The fit statistics in Table 4 indicate that both the trait- and the methods-only models provided unacceptable fits. In contrast, the CTCM model almost reached commonly desired minimum levels of model fit (CFI > .90; RMSEA < .09; SRMR < .07).

In the next step, we conducted a closer examination of the CTCM model. The loadings for the CTCM are shown in Table 1 and the factor intercorrelations in Table 5. Given that our CTCM model had an acceptable fit to the data, we proceeded by examining the reliability of the measures.

Table 6 provides reliability estimates for the CTCM model. We specifically estimated a CFA-based Cronbach's alpha and additionally three different versions of coefficient Omega based on the procedures described by Raykov (2001). As indicated by Table 6, all traits provided an acceptable level of reliability that was similar to commonly observed levels of reliability for self- and observer ratings of personality. These findings suggest that the number of situations we sampled were more than adequate to measure the respective trait using thin slices. Full syntax and output of all models is provided at https://osf.io/8fupx/ and https://osf.io/9k8da/, respectively.

**Objective 2: Correspondence Between Situational MTMM factors and Parent Ratings**

Given that our analyses in the preceding section provided evidence that a situational MTMM approach can be successfully fit to child personality data, our next research objective was to assess the degree to which these measurements correspond to parent ratings. We therefore estimated factor scores for the CTCM models using the lavPredict function in the lavaan package (Rosseel, 2012). Correlations between these factor scores and the parent ratings are provided in Table 7. As indicated by Table 7, the situational Big 5 factors showed considerable correspondence to the parents' ratings. Table 7 also presents the corresponding thin-slice parent correlations based on standard scoring procedures (not using latent traits), for comparison (Tackett et al., 2016).

**Objective 3: Predictive and Incremental Validity**

Table 8 presents results from regression analyses that predicted competencies and psychopathology from parent ratings, situational ratings, and both sources combined. Specifically, model 1 included all five parent-rated traits as predictors of parent-rated CBCL Total Competence, CBCL Internalizing Problems, and CBCL Externalizing Problems, model 2 included all five thin-slice coded traits as predictors of the same outcomes, whereas model 3 included both the five parent-rated and the five thin-slice coded traits as predictors. In terms of CBCL Competence, as indicated by Table 8, the situational measures and the parent ratings both demonstrated substantial predictive validity for CBCL Competence ($R^2$ = .20 and .35, $p < .001$, respectively). Additionally, and interestingly, the situational ratings also showed incremental validity over the parent ratings for CBCL Competence, $\Delta R^2 = .07$, $p < .001$. In addition, the primary predictive traits differed for parent ratings and thin-slice ratings. Specifically, both parent and thin-slice-rated high conscientiousness uniquely predicted CBCL Total Competence. However, only parent-rated low neuroticism and high extraversion uniquely predicted CBCL Total Competence whereas thin-slice-rated low agreeableness and low openness predicted CBCL Total Competence. Further, we can compare these results to those using standard scoring (in Tackett et al., 2016), which illustrates some divergence. Specifically, at least at the bivariate level, standard scoring showed that all thin-slice-rated traits were associated with all CBCL competence scales (except for the thin-slice rated agreeableness – CBCL Activities correlation, which did not reach significance). In contrast, in terms of the CBCL psychopathology scales, as indicated by Table 8, the situational measures did not demonstrate the level of predictive validity that the parent ratings did (Internalizing Problems: $R^2$ = .01 and .37, respectively; Externalizing Problems: $R^2$ = .04 and .53, respectively). This difference is also reflected in the bivariate correlations presented in Table 9, which showed significant correlations across all five traits and CBCL Total Competence, but significant correlations only between low thin-slice coded conscientiousness and agreeableness with CBCL Externalizing Problems.

## Discussion

The present study found several variations of latent variable MTMM models to adequately capture variance in a multi-method multi-informant battery of child personality data. Specifically, we found that the CTCM model could be fitted to TS child personality ratings across 15 situations, mother and father ratings of their child’s personality are associated with latent personality traits derived from CTCM model of situational TS ratings, and latent trait scores derived from these models illustrated incremental variance in competence outcomes predicted by thin-slice personality ratings, after accounting for parental personality ratings. Next, we summarize the findings of these three primary aims in turn, followed by an examination of the implications for implementation of thin-slice child personality ratings in future research.

Regarding the first primary objective, the present investigation aimed to examine the fit of comprehensive latent variable models following from an MTMM approach to these complex thin-slice child personality ratings across 15 situations. The findings support the application of these sophisticated models to comprehensively and concisely summarize patterns of covariation in multi-situational TS data. Simple models characterizing only latent personality traits or only situational (methods) factors did not characterize the data well, unsurprisingly. However, the desirable CTCM model (e.g., Lance & Fan, 2016; Widaman, 1985), which researchers often report difficulty fitting to MTMM data, actually fit these data reasonably well, perhaps facilitated by the availability of multiple (15) situations. In the CTCM model, latent traits representing all five higher-order personality traits showed excellent reliabilities estimated by alpha (ranging from 0.90 for Openness to 0.94 for Conscientiousness and Extraversion) and omega (ranging from 0.94 for Openness to 0.98 for Extraversion). The model fit could be improved via modification indices, altering trait or situational definitions, but here we interpret results from the overarching CTCM model, given cautions about relying heavily on modification indices to improve model fit (details of these additional analyses are available on request). Overall, the fit of this model supports evidence for convergent validity, indexed by high trait-factor loadings, as well as situational specificity, as indexed by method-factor loadings, in the TS personality ratings (Eid et al., 2016).

Regarding the second primary objective, we then demonstrated that different sources of child personality information (in this case, questionnaire reports from mothers and fathers) are associated with latent personality traits derived from the CTCM model of TS personality ratings. As illustrated in Table 7, the factor scores derived from the CTCM model showed reasonable convergent correlations with analogous parental questionnaire ratings, with associations between CTCM TS traits and mother report ranging from $r = .16$ (A) to $r = .33$ (C and O), and with father report ranging from $r = .02$ (N) to $r = .31$ (C). These associations are modest but largely significant, and best interpreted in the broader context of child personality research, wherein the potential “ceiling” of convergent associations is itself likely in the .50-.60 range (Tackett, 2011). Comparing the CTCM TS associations with mothers versus fathers offers an interesting perspective on parental informant differences, as well. For example, mothers’ and fathers’ ratings of C and A (the “self-regulatory” traits) were highly comparable in their overlap with CTCM TS ratings of these same traits. On the other hand, mother report was more overlapping with CTCM TS ratings for E and O, and particularly for N, than was father report. It is unclear why fathers may have a more biased perception of their children’s N than do mothers (as an example of one such interpretation of these differences), but these results point to important future directions in understanding informant differences for child personality. In this vein, we can further compare these correlations with correlations between thin-slice trait ratings derived from typical summed scores and the parental questionnaire ratings (also in Table 7). Overall, the magnitude of findings does not look substantially different for the CTCM factor scores versus the traditional summed scores, as they are associated with parental personality ratings.

Finally, we examined the ability of the latent personality traits derived from the CTCM model of TS personality ratings to predict relevant outcomes of interest, namely, child competencies and psychopathology symptoms. These results indicated that the CTCM personality scores did significantly predict childhood competencies, particularly CTCM predictors of high Conscientiousness and low Openness to Experience and Agreeableness. Furthermore, when added to the same model as parental ratings of child personality (which share method variance with the DV, given that parent questionnaires were used for both), the CTCM personality scores added significant variance to the prediction of child competencies. Although CTCM personality scores largely predicted psychopathology outcomes in expected directions (both in the regression models and at the bivariate correlation level), they did not add significant variance above and beyond parental report, when entered simultaneously into the model. This difference in predictive validity of CTCM personality scores for predicting competence versus psychopathology may be because the current thin-slice battery taps more into competence than vulnerability for psychopathology, but of course, it also reflects shared method variance in associating parent-reported personality with parent-reported psychopathology. It is also worth noting that performance in many of these situational tasks also reflects cognitive ability, or IQ (e.g., Hoffman, Kennedy, LoPilato, Monahan, & Lance, 2015). Future investigations should make efforts to measure cognitive ability as well, to disentangle these influences from personality traits and examine their unique contributions side-by-side.

**Implications for Assessment: Constructing a Thin-Slice Battery**

Given the results of the present study, what implications do they have for researchers, clinicians, and educators who might be interested in incorporating thin-slice ratings into their own work? We hope to offer some practical recommendations, as well as access to our stimulus materials, which have been posted on the Open Science Framework (https://osf.io/tgmcv/ and https://osf.io/4a3uy/). First, it is important to note that the CTCM TS personality scores show reasonably expectable means and standard deviations when compared to other common approaches to child personality assessment, such as parent-reported questionnaires (Table 2). The pattern of intercorrelations does not evidence the level of discriminant validity one might hope for, particularly for Neuroticism-Extraversion and Neuroticism-Openness (Table 5). Nonetheless, discriminant validity is not actually worse for thin-slice ratings relative to parent-reported questionnaires (Tackett et al., 2016), and thus may rather reflect less-than-ideal properties of existing child personality questionnaires and taxonomies rather than problems of the thin-slice method and/or CTCM approach to aggregating thin-slice ratings. It is also worth noting that none of the confidence intervals around these parameters exceed an upper bound of |.82|, and most are substantially lower. Future studies of child personality might use different child personality measures to conduct the TS ratings (i.e., here we use the ICID but there are certainly other options available). In manipulating the personality measure used, researchers can begin to disentangle the source of discriminant validity problems. Ultimately, the TS approach may be used to inform better construction of child personality trait measures by pointing toward better sources of distinct trait information within and across measures.

It is also interesting to examine the extent to which CTCM TS personality scores were differentially correlated with different situations. On average, within-trait cross-situation correlations were fairly high (mostly $r = .40$-$.50$; see Table 3), as would be expected if the traits manifest stability across different lab-based tasks. However, a more nuanced look at these results suggests substantial variability at the specific level of trait-situation “fit” (see full syntax and results posted on OSF; https://osf.io/8fupx/ and https://osf.io/9k8da/). The complexity and quantity of output was too cumbersome to present in the manuscript itself; thus, aggregate results are presented here but comprehensive specific results posted on OSF. It is this level of analysis that may be most informative when making decisions about constructing a thin-slice battery to measure child personality traits. For example, one situation evidenced some of the highest loadings on all five latent CTCM traits: asking children to introduce themselves as though they were meeting a child their own age for the first time (Situation E). Thin-slice ratings of children in this task were strong indicators of the latent CTCM TS traits: C ($r = .72$), N ($r = .75$), E ($r = .72$), O ($r = .62$), and A ($r = .73$). An advantage of the CTCM model is that it further allows us to examine how situationally specific variance in the "introduce yourself" task was associated with CTCM TS trait scores. Unsurprisingly, the situation was most associated with high extraversion ($r = .78$) and low neuroticism ($r = -.76$), above and beyond actual trait ratings.

We can further evaluate these specific trait-situation associations (based on the comprehensive output as posted on OSF) to seek guidance in constructing a TS battery for child personality. Beyond "introduce yourself", two situations were strong indicators of four of the five latent CTCM TS traits: making up definitions to nonsense words (situation J; traits C, N, E, and O) and coming up with creative uses for a sheet of paper (situation M; traits C, N, O, and A). Three tasks were similarly strong indicators of two latent CTCM TS traits: making up stories to TAT cards (situation F; traits O and A), verbally responding to a challenging spelling task (situation H; traits N and E), and responding to the receipt of a disappointing gift (situation N; traits C and A). Researchers might consider construction of an abbreviated TS battery consisting of these 6 tasks, which might take approximately 30 minutes. This estimate is also potentially high, as we used three TAT cards when potentially only 1 or 2 would be adequate. Notably, the three potentially most informative tasks (Situations E, J, and M) take only 9 minutes in aggregate to administer in a protocol. More research is needed to precisely approximate information loss, but these findings offer researchers some guidance in the creation of such a battery, selection of tasks, and options for evaluating potential costs and benefits of selecting different situations. Furthermore, by making these detailed results and all stimulus materials available on OSF, we hope that other researchers will continue to examine questions around situational specificity and ideal TS battery construction and incorporate these same tasks into their own data collection efforts.

**Limitations and Future Directions**

The present study reflects a methodological extension of previous work (Tackett et al., 2016), and thus shares limitations of the sample and study design. Although data were available from both mother and father, parents are likely highly overlapping as “judges”, given that they share similar roles in their child’s life, similar levels of acquaintanceship, and are likely exposed to the child in fairly similar settings (Tackett, 2011). It will be critically important for future work to extend the types of informants examined in this research area, for example, to self- and teacher-report, also common informants in child personality research (e.g., Göllner et al., 2017; Luan et al., 2017; Teglasi et al., 2015). It should be expected that different “types” of informants will introduce greater disagreement, and subsequently impact model fit. These informant-level extensions will enhance generalizability of the findings to the extent that they highlight other types of informants used in both research and practice.

Similarly, future work should also extend the types of situations examined. Although the 15 situations measured here were different in that they reflect different types of lab-based tasks (which would be expected to differentially invoke trait expression), all tasks were performed in the lab with fairly constrained situational parameters. An easy way to expand the types of situations examined is to capitalize on archival video data collected for other purposes – for example, video recording from clinical contexts could be thin-sliced to obtain similar situationally based child personality measurements (e.g., Tackett, Smack, et al., 2017). For example, these results suggested that the thin-slice tasks did a “better” job of capturing trait variance relevant for competencies than for psychopathology. Yet, thin-slice application of a clinical sample of children engaged in a conflict scenario better captured psychopathology variance and those traits most relevant for psychopathology (i.e., N; Tackett, Smack, et al., 2017), suggesting that pursuing new situations or tasks may be fruitful for predicting different types of outcomes. Future work could expand this to even more ecologically valid situations at home, school, and play, to better understand situational specificity and variability in observed child personality.

In application of these more complex CTCM models to multi-method/ multi-situation data, researchers must be thoughtful about their strengths and limitations. In much psychological research focusing on individual differences, such as that described here, the primary goal of statistical analysis is precise estimation of effect sizes (McShane, Gal, Gelman, Robert, & Tackett, 2017; Tackett, Lilienfeld, et al., 2017). Thus, researchers must prioritize those analytic models that maximize accuracy and precision, including more complex modeling approaches such as that demonstrated here. Increased accuracy and precision is a clear strength of this analytic approach, and comparison of these results with those presented in Tackett, Lilienfeld, et al. (2016) reveal many overarching similarities but a number of differences, as well, suggesting that estimates obtained from this model are not reproduced in simpler, more common, standard-scoring approaches. Limitations include the model complexity, challenges in getting these models to converge, urges to achieve greater model fit by examining modification indices – all of which may also impact the potential replicability of this type of research (Tackett, Lilienfeld, et al., 2017). These strengths and limitations must be struggled with as researchers move to improve accuracy and precision of effect sizes with more complex models while also contributing toward a replicable set of scientific findings.

A final potential limitation of our investigation is that it focused only on the CTCM. Our decision to focus on the CTCM was primarily motivated by recommendations in the literature that the CTCM is the preferred model when it does actually converge and our expectation that a comprehensive thin-slice battery would enable us to overcome the typical convergence issues with the CTCM. These arguments notwithstanding, fitting alternatives can provide additional insights and especially can provides insights into potential improvements in model fit by using less restrictive models. In supplementary analyses, we therefore also fitted the most popular alternative to the CTCM—the correlated traits, correlated uniquenesses model (CTCU) shown in Figure 2 (e.g., Marsh, 1989; Marsh & Bailey, 1991). The CTCU specifies correlated traits but does not specify latent variables for the methods. Instead, all indicators sharing the same method have correlated uniquenesses (errors). Conceptually, the model eliminates the assumption that method effects can be modeled through a latent variable, which may be a plausible assumption. However, researchers have criticized the CTCU and have suggested that it typically leads to biased factor loadings. One reason is that the method uniquenesses in the CTCU are not allowed to be correlated across methods. Marsh (1989; also see Bauer et al., 2013) has argued that the CTCU can successfully deal with a situation in which a second-order general trait factor (i.e., higher order factor) affects all the trait ratings. In a situation of this type, correlated method factors may absorb the general trait variance (for statistical detail, see the appendix). However, this specification of the CTCU is also a weakness when the methods are truly correlated because there is conceptual overlap between the methods. In forcing the methods to be uncorrelated, the CTCU assigns this shared method variance to the traits thereby inflating the trait factor loadings (Lance et al., 2002, 2004; Lievens

& Conway, 2001). In the present study, the fit of the CTCU was slightly better than the CTCM, $\chi^2$ (df = 2540) = 4075.90, $p < .001$, CFI = 0.939, RMSEA = 0.043, SRMR = 0.077. Given the fact that the CTCM converged and also fit well, and is the theoretically preferred model, we felt that using and interpreting the CTCU was not worth the loss in potential explanatory power and interpretational clarity.

Other alternatives to the CTCM are versions of this model that define the methods relative to a reference method (Eid et al., 2016; Lance et al., 2002). Although these approaches frequently improve model fit, it can be difficult to decide on the appropriate reference method—a difficulty which is exacerbated in the absence of self-report data, typical in child personality research. When a researcher is interested in an improved model fit, the researcher could also consider post-hoc modifications of the CTCM. In our data, leaving out one or two traits or methods that do not fit very well can relatively quickly improve fit to common standard recommended in the literature. We made the data of this study available only so that interested readers can also explore post-hoc modifications of this type.

**Conclusions**

In conclusion, the present study demonstrated success in fitting a CTCM model—the most theoretically desirable instantiation of the larger class of MTMM models—to child personality data obtained from a comprehensive battery of TS ratings across 15 situations. The CTCM model estimated latent child personality traits which demonstrated desirable reliability and validity, showed expected convergence with traditional parent-reported questionnaires, and predicted incremental variance in child competence outcomes, above and beyond parent-report. An advantage of the CTCM model is the nuanced information it offers about specific trait-situation patterns of "fit", which extend our theoretical understanding of how child personality traits differentially manifest across various situations, but also offer practical guidance in the creation of briefer TS batteries to assess child personality in practice. Finally, we aimed to make the study maximally useful by including all stimulus materials, code, detailed results, and de-identified data on OSF, in hopes that other researchers and practitioners will be encouraged to include or broaden measurement of child personality in their ongoing and future work.

## Appendix: Formal description of the MTMM models used in the Study and their differences

1. Correlated traits, correlated traits model (CTCM)

In a formalized way (Lance et al., 2002), the model can be written as

$$TMU_{ij} = \lambda_{Tij} T_i + \lambda_{Mij} M_j + \delta_{ij}$$

In this formula, TMU refers to the combination of trait $i$ with method $j$. $\lambda$ are the factor loadings that link the unit to the trait factor $T_i$ and the method factor $M_j$. Finally, $\delta$ includes the residual variance that is not explained by the trait and method factors.

The monotrait-heteromethod correlations in the CTCM are explained by the correlations between the method as follows

$MTHM = \lambda_{Tij}\ \lambda_{Tij'} + \lambda_{Mij}\ \lambda_{Mij'}\ \psi_{MjMj'}$

where $\lambda_{Mij}$ and $\lambda_{Mij'}$ refer to the standardized loadings of the TMUs on the method factors and $\psi_{MjMj'}$ captures the correlation between the methods (see Lance et al., 2004). In other words, the CTCM removes the variance that is explained by the correlation between the methods from the MTHM correlation.

2. Correlated traits, correlated uniquenesses model (CTCU)

In a formalized way, the CTCU can be written as follows:

$TMU_{ij} = \lambda_{Tij}T_i + \delta_{ij}$

As explained previously, TMU refers to the combination of trait $i$ with method $j$, $\lambda_{Tij}$ are the factor loadings that link the unit to the trait factor $T_i$. However, the CTCU differs from the CTCM with regard to the residual variance $\delta_{ij}$. In the CTCU, $\delta_{ij}$ not only includes specific ($s_{ij}$) and nonsystematic effects ($e_{ij}$) but also the methods effects $\lambda_{Mij}M_j$. More formally, this can be written as $\delta_{ij} = s_{ij} + e_{ij} + \lambda_{Mij}M_j$.

Formally, the CTCU defines the monotrait-heteromethod (MTHM) correlations as follows.

$MTHM = \lambda_{Tij}\ \lambda_{Tij'}$

In this equation, $\lambda_{Tij}$ is the standardized loading of the trait-method unit (TMU or indicator) and $\lambda_{Tij'}$ refers to the standardized loading of another TMU on the trait factor. In the CTCU, the correlation between two traits measured by the same method is accordingly only determined/explained by the trait loadings and not by their shared method like in the CTCM.

References

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles.* Department of Psychiatry, University of Vermont, Burlington, VT.

Allik, J., de Vries, R.E., Realo, A. (2016). Why are moderators of self-other agreement difficult to establish? Manuscript in press, *Journal of Research in Personality, 63,* 72-83. doi: http://dx.doi.org/10.1016/j.jrp.2016.05.013

Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology, 86,* 599-614. doi:10.1037/0022-3514.86.4.599

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81-105. http://dx.doi.org/10.1037/h0046016

Clark, D. A., Durbin, C. E., Donnellan, M. B., & Neppl, T. K. (2017). Internalizing symptoms and personality traits color parental reports of child temperament. Manuscript in press, *Journal of Personality.* doi: 10.1111/jopy.12293

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study.

Psychological Bulletin, 131, 483–509. doi:10.1037/0033-2909.131.4.483

Deal, J. E., Halverson, C. F., Jr., Martin, R. P., Victor, J., & Baker, S. (2007). The inventory of children's individual differences: Development and validation of a short version. Journal of Personality Assessment, 89, 162–166. http://dx.doi.org/10.1080/00223890701468550

Durbin, C. E. (2010). Modeling temperamental risk for depression using developmentally sensitive laboratory paradigms. Child Development Perspectives, 4, 168-173. doi:10.1111/j.1750-8606.2010.00137.x

Eid, M., Geiser, C., & Koch, T. (2016). Measuring method effects: From traditional to design-oriented approaches. Current Directions in Psychological Science, 25, 275-280. https://doi.org/10.1177/0963721416649624

Göllner, R., Roberts, B. W., Damian, R. I., Lüdtke, O., Jonkmann, K., & Trautwein, U. (2017). Whose "storm and stress" is it? Parent and child reports of personality development in the transition to early adolescence. Journal of Personality, 85, 376-387. doi: 10.1111/jopy.12246

Halverson, C. F., Havill, V. L., Deal, J., Baker, S. R., Victor, J. B., Pavlopoulos, V., . . . Wen, L. (2003). Personality structure as derived from parental ratings of free descriptions of children: The inventory of child individual differences. Journal of Personality, 71, 995–1026. http://dx.doi.org/10.1111/1467-6494.7106005

Hoffman, B. J., Kennedy, C. L., LoPilato, A. C., Monahan, E. L., & Lance, C. E. (2015). A review of the content, criterion-related, and construct-related validity of assessment center exercises. *Journal of Applied Psychology*, *100*, 1143.

Hollingshead, A. B. (1975). *Four factor index of social status*. Unpublished manuscript, Yale University, New Haven, CT.

Lance, C. E., & Fan, Y. (2016). Convergence, admissibility, and fit of alternative confirmatory factor analysis models for MTMM data. *Educational and Psychological Measurement*, *76*, 487-507. https://doi.org/10.1177/0013164415601884

Lance, C. E., Lambert, T. A., Gewin, A. G., Lievens, F., & Conway, J. M. (2004). Revised estimates of dimension and exercise variance components in assessment center postexercise dimension ratings. *Journal of Applied Psychology*, *89*, 377-385. doi: 10.1037/0021-9010.89.2.377

Lance, C. E., Noble, C. L., & Scullen, S. E. (2002). A critique of the correlated trait-correlated method and correlated uniqueness models for multitrait-multimethod data. *Psychological Methods*, *7*, 228-244. doi:10.1037//1082-989X.7.2.228

Lievens, F., & Conway, J. M. (2001). Dimension and exercise variance in assessment center scores: A large-scale evaluation of multitrait-multimethod studies. *Journal of Applied Psychology*, *86*, 1202-1222. https://doi.org/10.1037/0021-9010.86.6.1202

Luan, Z., Hutteman, R., Denissen, J. J., Asendorpf, J. B., & van Aken, M. A. (2017). Do you see my growth? Two longitudinal studies on personality development from childhood to young adulthood from multiple perspectives. *Journal of Research in Personality*, *67*, 44-60. https://doi.org/10.1016/j.jrp.2016.03.004

Marsh, H.W. (1989). Confirmatory factor analyses of multitrait multimethod data: Many problems and a few solutions. *Applied Psychological Measurement*, *13*, 335-361. https://doi.org/10.1177/014662168901300402

Marsh, H.W., & Bailey, M. (1991). Confirmatory factor analyses of multitrait multimethod data: A comparison of alternative models. *Applied Psychological Measurement*, *15*, 47-70. https://doi.org/10.1177/014662169101500106

McShane, B. B., Gal, D., Gelman, A., Robert, C., & Tackett, J. L. (2017). Abandon statistical significance. *arXiv preprint arXiv:1709.07588*.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Raykov, T. (2001). Bias of coefficient α for fixed congeneric measures with correlated errors. *Applied Psychological Measurement*, *25*, 69-76. https://doi.org/10.1177/01466216010251005

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1-36.

Rothbart, M. K., & Bates, J. E. (2006). Temperament. In N. Eisenberg (Vol. Ed.) and W. Damon and R. M. Lerner (Eds.-in-Chief), *Handbook of Child Psychology 6th Ed Vol. 3: Social emotional and personality development* (pp. 99-166). Hoboken, NJ: Wiley.

Soto, C. J., & Tackett, J. L. (2015). Personality traits in childhood and adolescence: Structure, development, and outcomes. *Current Directions in Psychological Science, 24*, 358-362. https://doi.org/10.1177/0963721415589345

Tackett, J. L. (2011). Parent informants for child personality: Agreement, discrepancies, and clinical utility. *Journal of Personality Assessment 93* 539-544. doi:10.1177/1073191113509686

Tackett, J. L., & Durbin, C. E. (2017). Advances in research on youth personality: Introduction to the special issue. *Journal of Research in Personality 67*, 1-2. doi.org/10.1016/j.jrp.2017.03.004

Tackett, J. L., Herzhoff, K., Kushner, S. C., & Rule, N. (2016) Thin slices of child personality: Perceptual, situational, and behavioral contributions. *Journal of Personality and Social Psychology 110* 150-166. http://dx.doi.org/10.1037/pspp0000044

Tackett, J. L., Lilienfeld, S. O., Patrick, C. J., Johnson, S. L., Krueger, R. F., Miller, J. D., ... & Shrout, P. E. (2017). It's time to broaden the replicability conversation: Thoughts for and from clinical psychological science. *Perspectives on Psychological Science, 12*, 742-756.

Tackett, J. L., Smack, A. J., Herzhoff, K., Reardon, K. W., Daoud, S., & Granic, I. (2017). Measuring child personality when child personality was not measured: Application of a thin-slice approach. *Personality and Mental Health* 11 4-13. doi:10.1002/pmh.1351

Teglasi, H., Schussler, L., Gifford, K., Annotti, L. A., Sanders, C., & Liu, H. (2015). Child behavior questionnaire-short form for teachers: Informant correspondences and divergences. *Assessment* 22 730-748. https://doi.org/10.1177/1073191114562828

Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement*, *9*, 1-26. https://doi.org/10.1177/014662168500900101

Table 1

Description of Thin-slice Situations and Factor Loadings from the Correlated Traits Correlated Methods Model

| | Description | Standardized Loadings | | | | | |
|---|---|---|---|---|---|---|---|
| | | C | N | E | O | A | Mean |
| Sit A | Answering demographic interview questions | 0.67 | 0.60 | 0.61 | 0.52 | 0.63 | 0.61 |
| Sit B | Completing a Go-No-Go-like task | 0.67 | 0.48 | 0.52 | 0.45 | 0.61 | 0.55 |
| Sit C | Playing Jenga | 0.70 | 0.58 | 0.56 | 0.46 | 0.63 | 0.59 |
| Sit D | Whispering cartoon character names | 0.65 | 0.56 | 0.58 | 0.49 | 0.62 | 0.58 |
| Sit E | Pretending to introduce oneself to a peer | 0.72 | 0.64 | 0.60 | 0.55 | 0.69 | 0.64 |
| Sit F | Telling a story based on ambiguous pictures | 0.68 | 0.62 | 0.61 | 0.49 | 0.71 | 0.62 |
| Sit G | Sharing one's best memory | 0.68 | 0.63 | 0.61 | 0.58 | 0.59 | 0.62 |
| Sit H | Spelling words | 0.56 | 0.69 | 0.76 | 0.46 | 0.67 | 0.63 |
| Sit I | Sharing a sad memory | 0.66 | 0.59 | 0.61 | 0.50 | 0.67 | 0.61 |
| Sit J | Defining neologisms | 0.69 | 0.65 | 0.68 | 0.54 | 0.67 | 0.65 |
| Sit K | Solving math problems | 0.60 | 0.71 | 0.78 | 0.48 | 0.68 | 0.65 |
| Sit L | Singing one's favorite song | 0.74 | 0.49 | 0.59 | 0.49 | 0.55 | 0.57 |
| Sit M | Generating uses for paper | 0.76 | 0.68 | 0.66 | 0.56 | 0.71 | 0.67 |
| Sit N | Receiving a disappointing gift after a desirable one | 0.72 | 0.47 | 0.52 | 0.40 | 0.66 | 0.55 |
| Sit O | Completing a digit span task | 0.67 | 0.61 | 0.62 | 0.49 | 0.73 | 0.62 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Mean | 0.68 | 0.60 | 0.62 | 0.50 | 0.65 | 0.61 |

*Note*. Sit = situation; C = Conscientiousness; N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness.

Table 2

Mean Correlations Between the Traits (Estimated in Each Situation and Then Averaged)

| | Correlations | | | | M | SD |
|---|---|---|---|---|---|---|
| | Conscientiousness | Neuroticism | Extraversion | Openness | | |
| Conscientiousness | 1.00 | | | | 4.39 | 0.84 |
| Neuroticism | -.26 | 1.00 | | | 3.37 | 0.77 |
| Extraversion | .34 | -.82 | 1.00 | | 4.32 | 0.76 |
| Openness | .52 | -.67 | .73 | 1.00 | 4.40 | 0.72 |
| Agreeableness | .63 | -.27 | .43 | .29 | 4.69 | 0.85 |

*Note*. Correlations higher than $r = |.11|$ are significant at $p < .05$.

Correlations presented are calculated using a two-step procedure: first, correlations of each trait pair were calculated based on each of the 15 thin-slice situation, then, these correlation estimates were averaged across the 15 thin-slice situation for each of the correlation pairs.

Table 3

Mean Correlations Between the Situations (Estimated for Each Trait and Then Averaged)

| Sit. | Sit. A | Sit. B | Sit. C | Sit. D | Sit. E | Sit. F | Sit. G | Sit. H | Sit. I | Sit. J | Sit. K | Sit. L | Sit. M | Sit. N | Sit. O |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | -- | | | | | | | | | | | | | | |
| B | .51 | -- | | | | | | | | | | | | | |
| C | .48 | .49 | -- | | | | | | | | | | | | |
| D | .46 | .46 | .45 | -- | | | | | | | | | | | |
| E | .55 | .45 | .43 | .43 | -- | | | | | | | | | | |
| F | .51 | .46 | .46 | .45 | .59 | -- | | | | | | | | | |
| G | .52 | .40 | .42 | .41 | .58 | .62 | -- | | | | | | | | |
| H | .45 | .34 | .34 | .40 | .47 | .44 | .47 | -- | | | | | | | |
| I | .47 | .36 | .39 | .37 | .61 | .57 | .62 | .45 | -- | | | | | | |
| J | .49 | .48 | .46 | .48 | .54 | .61 | .57 | .48 | .55 | -- | | | | | |
| K | .43 | .37 | .37 | .41 | .45 | .40 | .47 | .48 | .42 | .48 | -- | | | | |
| L | .45 | .42 | .40 | .40 | .56 | .58 | .51 | .38 | .50 | .50 | .39 | -- | | | |
| M | .47 | .47 | .45 | .45 | .61 | .63 | .58 | .49 | .57 | .60 | .45 | .57 | -- | | |
| N | .42 | .43 | .46 | .43 | .47 | .50 | .46 | .41 | .41 | .47 | .37 | .45 | .53 | -- | |
| O | .52 | .50 | .50 | .49 | .49 | .46 | .42 | .51 | .41 | .49 | .46 | .44 | .51 | .47 | |
| M | 4.32 | 4.43 | 4.42 | 4.25 | 4.25 | 4.25 | 4.19 | 4.09 | 4.18 | 4.20 | 4.15 | 4.15 | 4.24 | 4.13 | 4.29 |
| SD | 0.77 | 0.70 | 0.71 | 0.71 | 0.89 | 0.87 | 0.84 | 0.80 | 0.79 | 0.84 | 0.74 | 0.84 | 0.78 | 0.84 | 0.71 |

Note. Correlations higher than $r = |.11|$ are significant at $p < .05$.

Sit. = Situation.

Correlations presented are calculated using a two-step procedure: first, correlations of each situation pair were calculated for each of the Big Five traits, then, these correlation estimates were averaged across the Big Five traits for each of the correlation pairs.

Table 4

Fit Statistics for the Four Compared Models

| Model | $\chi^2$ | df | p | CFI | RMSEA | SRMR |
|---|---|---|---|---|---|---|
| Correlated Traits | 14166.90 | 2690 | <.001 | 0.543 | 0.114 | 0.088 |
| Correlated Methods | 13197.58 | 2595 | <.001 | 0.578 | 0.112 | 0.161 |
| Correlated Traits, Correlated Methods | 5546.64 | 2510 | <.001 | 0.879 | 0.061 | 0.070 |

*Note*. CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual.

Table 5

Factor Intercorrrelations from the Correlated Traits Correlated Methods Model

| | Neuroticism | Extraversion | Openness | Agreeableness |
|---|---|---|---|---|
| Conscientiousness | -0.09 | 0.24 | 0.40 | 0.81 |
| Neuroticism | | -0.92 | -0.81 | -0.06 |
| Extraversion | | | 0.81 | 0.34 |
| Openness | | | | 0.19 |

Table 6

Estimated Reliabilities for the Latent Multi-Method Traits in the CTCM Model

| | Cronbach's Alpha | Omega |
|---|---|---|
| Conscientiousness | 0.94 | 0.96 |
| Neuroticism | 0.93 | 0.96 |
| Extraversion | 0.94 | 0.98 |
| Openness | 0.90 | 0.94 |
| Agreeableness | 0.92 | 0.95 |

Note. CTCM = Correlated Traits, Correlated Methods

Table 7

Comparison of Correlations between Mother- and Father-Reported Traits and Thin-slice Factor Scores and Thin-slice Summed Scale Scores

| | Factor Scores from the CTCM model | | | | | Summed Scale Scores | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | C | N | E | O | A | C | N | E | O | A |
| Mother-reported | | | | | | | | | | |
| C | .33 | -.10 | .13 | .20 | .23 | .38 | -.16 | .17 | .23 | .26 |
| N | -.04 | .25 | -.21 | -.19 | .00 | -.12 | .32 | -.28 | -.26 | -.07 |
| E | .08 | -.23 | .24 | .18 | .08 | .14 | -.29 | .31 | .22 | .13 |
| O | .06 | -.18 | .15 | .33 | -.08 | .16 | -.27 | .26 | .43 | .02 |
| A | .19 | -.01 | .04 | .05 | .16 | .24 | -.10 | .13 | .14 | .21 |
| Father-reported | | | | | | | | | | |
| C | .31 | -.09 | .11 | .17 | .23 | .33 | -.10 | .11 | .15 | .23 |
| N | -.22 | .02 | -.02 | -.09 | -.14 | -.26 | .11 | -.11 | -.15 | -.18 |
| E | .17 | -.09 | .12 | .08 | .14 | .21 | -.15 | .19 | .13 | .17 |
| O | .14 | -.04 | .02 | .17 | .01 | .20 | -.12 | .12 | .25 | .08 |
| A | .14 | .04 | .01 | -.02 | .16 | .16 | -.03 | .06 | .06 | .20 |

Note. CTCM = Correlated Traits, Correlated Methods; C = Conscientiousness; N = Neuroticism; E = Extraversion: O = Openness; A = Agreeableness.

Table 8

Regression Models of Parent-rated and Thin-slice Latent Traits Predicting Child Behavior Checklist Total Competence, Internalizing Problems, and Externalizing Problems Scores

| | Total Competence | | | | | Internalizing Problems | | | | | Externalizing Problems | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | B | t-value | $R^2$ | $\Delta R^2$ vs. M1 | $\Delta R^2$ vs. M2 | B | t-value | $R^2$ | $\Delta R^2$ vs. M1 | $\Delta R^2$ vs. M2 | B | t-value | $R^2$ | $\Delta R^2$ vs. M1 | $\Delta R^2$ vs. M2 |
| Model 1 (M1) | | | 0.35*** | | | | | .37*** | | | | | .53*** | | |
| Parent C | 0.33 | 5.77*** | | | | 0.00 | 0.06 | | | | -0.23 | -4.66*** | | | |
| Parent N | -0.25 | -3.99*** | | | | 0.81 | 12.91*** | | | | 0.13 | 2.42* | | | |
| Parent E | 0.18 | 3.05*** | | | | 0.23 | 3.90*** | | | | 0.06 | 1.17 | | | |
| Parent O | -0.03 | -0.62 | | | | 0.19 | 3.39*** | | | | 0.10 | 2.16* | | | |
| Parent A | 0.00 | -0.02 | | | | 0.12 | 2.29** | | | | -0.5 | -12.82* | | | |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | 9 | ** | | | |
| Model 2 (M2) | | | 0.20*** | | | | | .01 | | | | | .04* | | |
| TS C | 0.97 | 3.75** | | | | 0.15 | 0.52 | | | | -0.15 | -0.52 | | | |
| TS N | -0.16 | -0.51 | | | | 0.19 | 0.53 | | | | -0.19 | -0.55 | | | |
| TS E | 0.44 | 0.98 | | | | 0.15 | 0.29 | | | | -0.18 | -0.36 | | | |
| TS O | -0.42 | -2.07* | | | | -0.03 | -0.15 | | | | 0.01 | 0.04 | | | |
| TS A | -0.70 | -2.21* | | | | -0.15 | -0.43 | | | | -0.02 | -0.05 | | | |
| Model 3 | | | 0.42*** | .07*** | .21*** | | | .39*** | .02 | .38*** | | | .54*** | .01 | .50*** |
| Parent C | 0.25 | 4.29*** | | | | -0.04 | -0.73 | | | | -0.25 | -4.78*** | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Parent N | -0.21 | -3.45*** | 0.84 | 13.20*** | 0.16 | 2.92*** |
| Parent E | 0.17 | 2.73** | 0.25 | 3.93*** | 0.06 | 1.01 |
| Parent O | -0.04 | -0.70 | 0.22 | 3.49*** | 0.14 | 2.55* |
| Parent A | 0.03 | 0.55 | 0.13 | 2.33** | -0.58 | -12.53*** |
| TS C | 0.54 | 2.28* | 0.12 | 0.51 | 0.08 | 0.37 |
| TS N | -0.16 | -0.55 | -0.46 | -1.60 | -0.45 | -1.79 |
| TS E | 0.13 | 0.32 | -0.44 | -1.06 | -0.30 | -0.84 |
| TS O | -0.16 | -0.83 | -0.07 | -0.36 | -0.15 | -0.89 |
| TS A | -0.39 | -1.37 | 0.08 | 0.27 | 0.03 | 0.13 |

Note. *** $p < .001$ ** $p < .01$ * $p < .05$

C = Conscientiousness; N = Neuroticism; E = Extraversion: O = Openness; A = Agreeableness; TS = Factor Score from the Correlated Traits, Correlated Methods Model. Model 1 included all five parent-rated traits as predictors of parent-rated CBCL Total Competence, CBCL Internalizing Problems, and CBCL Externalizing Problems, model 2 included all five thin-slice coded traits as predictors of the same outcomes, whereas model 3 included both the five parent-rated and the five thin-slice coded traits as predictors.

Table 9

Correlations between Thin-slice Latent Traits and Child Behavior Checklist Total Competence, Internalizing Problems, and Externalizing Problems Scores

| | Total Competence | Internalizing Problems | Externalizing Problems |
|---|---|---|---|
| TS Conscientiousness | 0.33 | 0.03 | -0.19 |
| TS Neuroticism | -0.25 | 0.07 | -0.03 |
| TS Extraversion | 0.23 | -0.07 | -0.03 |
| TS Openness | 0.31 | -0.04 | -0.03 |
| TS Agreeableness | 0.18 | 0.01 | -0.19 |

*Note*. Correlations higher than $r = |.11|$ are significant at $p < .05$.

TS = Factor Score from the Correlated Traits, Correlated Methods Model.

Figure 1

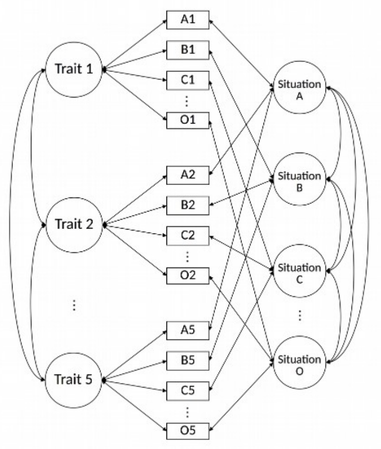


Figure 1. Correlated traits, correlated methods model. Letters correspond to different situations; numbers correspond to traits. Ellipses indicate omitted variables.

Figure 2

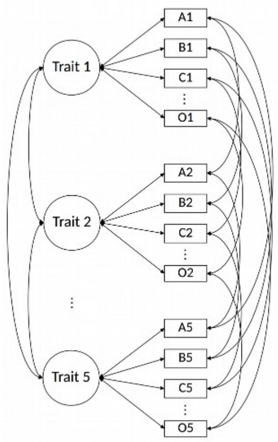


Figure 2. Correlated traits, correlated uniquenesses model. Uniquenesses are correlated among measures from the same situation. Letters correspond to different situations; numbers correspond to traits. Ellipses indicate omitted variables.