Analysis of the Convergent and Discriminant Validity of the CBCL, TRF, and YSR in a Clinic-Referred Sample

Rapson Gomez[1]

University of Ballarat

Alasdair Vance & Rashika Miranjani Gomez

The University of Melbourne

[1]Correspondence concerning this article should be addressed to Rapson Gomez, School of Health Sciences, University of Ballarat, University Drive, Mount Helen, Ballarat, Victoria, 3353, Australia. Email: rapson.gomez@ballarat.edu.au

Analysis of the Convergent and Discriminant Validity of the CBCL, TRF, and YSR in a Clinic-Referred Sample

Abstract

This study used the correlated trait-correlated method minus one model to examine the convergent and discriminant validity of the Child Behavior Checklist (CBCL), Teacher's Report Form (TRF) and Youth Self-Report (YSR), based on ratings of clinic-referred adolescents. A total of 294 adolescents, comprising 70.5% males, provided self-ratings on the YSR. The adolescents were also rated by their mothers and teachers on the CBCL and TRF respectively. The findings indicated some support for convergent validity for the CBCL and TRF for Anxious/Depressed, Withdrawn/Depressed and Rule Breaking Behavior, and for the convergent validity for the CBCL and YSR for Rule Breaking Behavior. There was support for the discriminant validity between virtually all the traits. The findings are discussed in relation to the construct validity, cross-informant agreement, and clinical use of the CBCL, TRF and YSR.

Key Words: CBCL; TRF; YSR; correlated trait-correlated method minus one model; mother, teacher, and adolescent ratings.

The Achenbach System of Empirically Based Assessment (ASEBA; Achenbach & Rescorla, 2001) includes the Child Behavior Checklist (CBCL, completed by parents), the Teacher Report Form (TRF, completed by teachers) and the Youth Self-Report (YSR, completed by adolescents). Worldwide, these scales are probably the most widely used cross-informant measures of child and adolescent problem behaviors. The present study used a multi-trait multi-method (MTMM) approach, called the correlated trait-correlated method minus one ([CT-C(M-1)]; Eid, Lischetzke, Nussbeck, & Trierweile, 2003) model, to examine the convergent (the extent to which the same trait is measured by different methods) and discriminant (the extent to which different traits in a measure are distinct) validity of these scales, based on self-ratings of clinic-referred adolescents on the YSR, and the ratings by their mothers on the CBCL and teachers on the TRF.

The CBCL, TRF, and YSR have comparable items, subscales and factor structure. Based on first-order exploratory factor analyses (EFA) and confirmatory factor analyses (CFA) of ratings from community and clinic samples, the factor structure for all three ASEBA measures is an oblique 8-factor model, with factors for Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior, and Aggressive Behavior (Achenbach & Rescorla, 2001). Second order factor analyses of these factors in the CBCL, TRF and YSR have produced two correlated higher order factors, called Internalizing and Externalizing. The Anxious/Depressed, Withdrawn/Depressed, and Somatic Complaints factors loaded together on the higher order factor Internalizing, whereas the factors for Rule-Breaking Behavior and Aggressive Behavior loaded together on the higher order factor called Externalizing. The remaining factors, Attention Problems, Thought Problems and Social Problems, cross-loaded on both of these higher order factors. In terms of zero-order correlations, there is support for the discriminant validity of the eight factors for all three ASEBA measures (e.g., Achenbach

& Rescorla, 2001, Grigorenko, Geiser, Slobodskaya, & Francis, 2010). The construct validity for the CBCL, TRF and YSR model has generally been supported in subsequent studies CFA studies of these measures separately (for a review, see Achenbach et al., 2008).

The CBCL, TRF and YSR have been offered as parallel measures capable of providing similar cross-informant scores from parents, teachers and adolescents, respectively. This could imply that there will be at least adequate convergence of the scores of like factors across these three measures. Although studies have shown significant correlations between the equivalent factors of these measures, a consistent finding is that these correlations are usually low indicating only low levels of cross-informant agreement. Generally, there has been relatively more parent-teacher agreement than parent-child/adolescent or teacher-child/adolescent agreement (Achenbach, Dumenci, & Rescorla, 2002; Achenbach & Rescorla, 2001; Grigorenko et al., 2010). For example, Grigorenko et al. reported mean factor correlations of .26 for CBCL-TRF, .15 for CBCL-YSR, and .05 for the TRF-YSR, which constitute low to medium effect sizes based on guidelines proposed by Hemphill (2003) for correlation effect sizes: < .2 = small, .2 to .3 = medium or moderate, and >.30 = large. Such findings have also been reported for other child and adolescent measures (e.g., De Los Reyes & Kazdin, 2005; for a meta-analysis, see Achenbach, McConaughy, & Howell, 1987).

While correlations and factor analysis are often used to evaluate construct, convergent and discriminant validity, there are problems with these approaches. These problems are best understood within the context of Classical Test Theory (CTT; Lord & Novick, 1968). CTT suggests that a raw score comprises variance for three components: trait, method, and random measurement error. Trait variance is the variance for the construct being measured while method variance is the systematic variance specific to the method used to collect the information on the construct. Random measurement error variance includes all other variance and is unrelated to method or trait variance. Both method and random measurement effects

are considered problematic because these effects distort (either increase or decrease) the relations among the trait constructs (Campbell & Fiske, 1959). This means that the existing cross-informant findings across the CBCL, TRF and YSR based on correlation analyses could be confounded with both method and error variance, and the CFA findings could be confounded with method variance.

Relative to simple correlations and factor analysis, the convergent and discriminant validity of a measure can be more robustly evaluated using MTMM (Campbell & Fiske, 1959) when applied within a CFA framework (Lance, Noble, & Scullen, 2002). This procedure can account for both method, and random measurement error variance. In general, this approach involves data for two or more traits measured by two or more methods (or informants). The original Campbell and Fiske MTMM approach (1959) evaluates convergent and discriminant validity using observed scores. MTMM analysis can also be conducted with latent scores using CFA models (Lance et al., 2002). Such models allow evaluation of the convergence of a measure (trait variance), after taking into account the method variance (systematic variance that is specific to the method used to collect information) and error variance within it.

Several MTMM models have been proposed. Two commonly used models are the correlated trait–correlated method (CT-CM) and the correlated trait–correlated uniqueness (CT-CU) models. Eid and his associates have suggested that a key issue that has to be considered when selecting a CFA model for MTMM analysis is the type of methods in the model (Eid et al., 2003). Methods can be either interchangeable (all respondents have the same access to the target, and therefore rate the target from the same perspective) or structurally different (all raters have different access to the target, and as such respond from different perspectives). As the CBCL, TRF and YSR are parallel measures, completed by mothers, teachers and adolescents, respectively, the ratings on these measures for the same

individual constitute structurally different methods. Eid and his associates have recently developed a model called the correlated trait-correlated method minus one, CT-C(M-1), that is claimed to be able to provide a highly rigorous evaluation of the convergent and discriminant validity of measures that models structurally different methods (Eid et al., 2003).

A CT-C(M-1) model is a confirmatory factor analysis (CFA) model that involves two or more traits measured by two or more methods. In this model one of the methods is selected as the reference method (Eid, 2000; Eid et al., 2003). The reference method is modeled without a latent method factor. For each trait, the true score (which is its trait score) of its indicator in the reference method is used to predict the true scores of the same indicators in of the other methods or non-reference methods. The proportions of the true scores in the indicators in non-reference methods that are so predicted are interpreted as their trait scores. A method factor is defined as a residual factor that is common to all variables measured by the same method. Hence, a method factor represents that part of a trait measured by a non-reference method that is not predicted by the true-score of the indicator in the reference method. The trait variances in the reference and non-reference methods are referred to as consistency coefficients. Since the consistency coefficient of an indicator in a non-reference method indicates it convergence with the same indicator in the reference method, it provides a measure of convergent validity. As higher consistency coefficients indicate more convergence for the non-reference method with the reference method, more support for convergent validity can be interpreted with higher values of the consistency coefficients. The method variances in the non-references are referred to as method-specific coefficients, with higher values indicating higher method variances. The reliability of an indicator is the sum of its consistency coefficient and method-specific coefficient. The discriminant validity of the different traits (as measured by the reference method), and the discriminant validity of the

different non-reference method factors are demonstrated if there are low correlations between the different trait factors, and between the method factors.

To date, as far as it can be ascertained, only one study, by Grigorenko et al. (2010), has used MTMM to examine the convergent and discriminant validity of the CBCL, TRF, and YSR. It used the CT-C(M-1) model. The study involved ratings of a community sample of children and adolescents. Mother ratings on the CBCL were used as the reference method. This means that that the true scores for the CBCL factors (the reference method) were used to predict the variance in the corresponding TRF and YSR factors. In general, the results indicated low convergence for CBCL and TRF ratings, and even lower convergence for CBCL and YSR ratings. The amount of shared variance (consistency coefficients) for the CBCL and TRF for the different factors ranged from .05 to .11, and the amount of shared variance for CBCL and YSR factors ranged from were .01 to .07. In addition, the study found that agreement between CBCL and TRF, and CBCL and YSR were comparable across all subscales. It also found support for the discriminant validity of all eight factors (as reflected in the reference method or CBCL ratings), and also the method factors for the non-reference methods (TRF and YSR).

Although Grigorenko et al. (2010) have provided available CT-C(M-1) based convergent and discriminant validity estimates for the CBCL, TRF and YSR, limitations of the study, indicate more studies are needed. Firstly, Grigorenko et al. (2010) examined a community sample, and given the clinical nature and utilization of the CBCL, TRF and YSR CT-C(M-1) studies for clinic-referred children/adolescents are warranted. Existing data suggest that cross-informant agreements are higher when behaviors are more observable (De Los Reyes & Kazdin, 2005). Given that problem behaviors can be expected to be more observable in clinic-referred sample compared to community samples, more convergence can expected in CBCL, TRF and YSR ratings for a clinical sample. Also, relative to clinic-

referred samples, children/adolescents from the general community report more problematic externalizing behaviors about themselves than their parents do about them (e.g., Sawyer, Baghurst, & Mathias, 1992; van den Ende & Verhulst, 2005). There are data showing greater discrepancy between self-reports of children and teachers reports for clinic-referred samples than community samples, with child self-reports showing more internalizing and less externalizing problems than teacher reports (Epkins, 1993). Secondly, Grigorenko et al. (2010) used a multiple indicators CT-C(M-1) model, with two parcels for each trait-method unit. Parcelling involves using groups of item sets as observed indicators. Experts consider this a controversial practice that often leads to unstable parameter estimates, and not warranted if the goal of a study is to understand the nature of a set of constructs and not items (Little, Cunningham, Shahar, & Widaman, 2002; Marsh, Lüdtke, Nagengast, Morin, & Von Davier, 2013), as is the case when the convergence of the CBCL, TRF and YSR is examined. In this respect a single indicator CT-C(M-1) model would be more appropriate. Thirdly, Grigorenko et al. (2010) did not examine the external validity of their latent trait factors. Examining the external validity would provide an additional test of the CBCL, TRF and YSR's construct validity. If the external validity is examined in terms of the relations of the latent factors with relevant clinical disorders, it could provide valuable information on utilization of the CBCL, TRF and YSR in clinical diagnosis.

The major aim of the current study was to use a single indicator CT-C(M-1) model to examine the convergent and discriminant validity of the CBCL, TRF, and YSR, based on ratings of clinic-referred adolescents, with mother ratings as the reference method. As noted by Grigorenko et al. (2010) using mother as the reference method is appropriate because maternal biases tend to be smaller than those of other raters. Doing so also allows comparison of the findings in this study with that of Grigorenko et al. We also examined if the trait factors were related to age and sex, as there is evidence that CBCL, TRF and YSR scores

vary by age and sex (Achenbach & Rescorla, 2001). An additional aim of the study was to examine the external validity of the trait factors in the CT-C(M-1) model by examining their correlations with anxiety, depressive, ADHD and ODD/CD diagnoses. Given that we examined a clinic-referred sample, we expected support for the convergent validity for the CBCL, TRF and YSR ratings. Support for discriminant validity of the eight factors was also expected.

Method

Participants

The data for all participants were collected archivally from the Academic Child Psychiatry Unit (ACPU) of the Royal Children's Hospital, Melbourne, Australia. The ACPU is an out-patient psychiatric unit that provides services for children and adolescents with behavioral, emotional, and learning problems. Referrals are generally from other medical services, schools, and social and welfare organizations. All parents and children were informed that the clinic would provide diagnosis and appropriate treatment, and that assessment will be over two days, covering a range of tests involving the parents, child and teachers. They were informed that all data collected would be kept in an unidentifiable form in a secure database and (if consent was given) used to support future research.

For the current study we used the records of children and adolescents, aged between 11 and 18 years. The data comprised retrospective referrals between 2004 and 2010, who had been interviewed for clinical diagnosis. An individual was selected for inclusion in the study if that individual had ratings for the CBCL, TRF and YSR. Apart from this and the age criteria, no other inclusion/exclusion criterion was applied when selecting participants for the study. In all, there were 294 adolescents, comprising 70.5% males and 29.5% females, with rating on the CBCL, TRF and YSR. The overall mean age of participants was 12.95 years (*SD* = 1.69 years).

Table 1 shows the means and *SD* of the *T* scores of the all CBCL, TRF and YSR scales for all participants. As shown, all eight syndrome scores for parent ratings (CBCL) were either at the clinical range (*T* scores ≥ 70) or borderline clinical range (*T* scores between 65 and 69). For teacher ratings (TRF), all the mixed (Attention Problems, Thought Problems and Social Problems) and externalizing (Rule-Breaking Behavior and Aggressive Behavior) syndromes were at the borderline clinical range , and all the internalizing syndromes (Anxious/Depressed, Withdrawn/Depressed, and Somatic Complaints) were at the normal range (*T* scores less than 65). For adolescent ratings, all syndromes, except attention problems, were at the normal range. Attention problems were at the borderline clinical range.

All clinical diagnoses for adolescents were derived from the Anxiety Disorders Interview Schedule for Children, Parent Version (ADISC-IV-P; Silverman & Albano, 1996), described below. The ADISC-IV-P guidelines for diagnosis are that the child be given a diagnosis of all disorders meeting the diagnostic criteria. Table 2 shows the percentages of different categories of disorders (Any Anxiety Disorders, Dysthymia/Major Depressive Disorder, Attention Deficit/Hyperactivity Disorder, and Oppositional Defiant Disorder/Conduct Disorder) for the participants. As shown, there were high frequencies for all four categories of disorders, and also high comorbidity between the different categories.

The percentages of father employment status were as follows: employed = 82.0%, home duties = 2.4%, pensioner/retired = 5.8%, unemployed = 5.8%, others/unknown = 4.0%. The percentages of father highest education level were as follows: tertiary = 14.5%, high school/some years in secondary school or equivalent = 64.0%, technical certificate or equivalent = 18.3%, primary school = 2.8, and no schooling = 0.4%. Thus, most fathers of participants were employed, and more than two-third of participants had fathers who had attended at least secondary school. In terms of parental relationship, about 50% were living together and 43% were separated or divorced. More than two-thirds (68%) of participants

were from families with income of less than $50,000 per year. In terms of parental relationship, 47.1% of mothers and fathers were living together.

Measures

Anxiety Disorders Interview Schedule for Children, Parent Version (ADISC-IV-P; Silverman & Albano, 1996). The ADISC-IV-P was used for diagnosis, and these diagnoses were also used for examining the external validity of the trait factors in the CT-C(M-1) model. The ADISC-IV-P is a semi-structured interview, based on the DSM-IV-TR diagnostic system (American Psychiatric Association, 2000). It has been designed to facilitate the diagnosis of major childhood disorders. The scores of ADISC-IV-P have sound psychometric properties (Silverman, Saavedra, & Pina, 2001). Test-retest reliability for the ADISC-IV-P scores over a 7 to 14-day interval has shown good to excellent reliability. Kappa values for interview with children between 7 and 16 years ranged from 0.61–0.80 (Silverman et al., 2001).

For the current study, the two research assistants who administered the ADISC-IV-P were provided with extensive supervised training and practice by the two psychologists prior to them collecting data. Training of the ADISC-IV-P included observations of it being administered by the psychologists. The research assistants commenced administering the ACISC-IV-P only after they attained competence in its administration, as assessed by the two registered psychologists. There was adequate inter-rater reliability for the diagnoses made between the research assistants and the psychologists, and between research assistants (average kappa value across all diagnoses = .88). Using the categorical data from the parent ADISC-IV-P, clinical diagnosis was determined by two consultant child and adolescent psychiatrists who independently reviewed the data. The inter-rater reliability for diagnoses of the two psychiatrists was high (kappa = .90).

Achenbach System of Empirically Based Assessment (ASEBA; Achenbach & Rescorla

2001). The ASEBA includes the most recent editions of the CBCL (now designated as the CBCL/6–18), TRF, and YSR (Achenbach & Rescorla, 2001). The CBCL, completed by parents, has 113 items, while the TRF has 120 items for teacher completion. Both are used to rate children between 6 and 18 years of age. The YSR, completed by individuals between 11 and 18 years, has 112 items, worded in the first person. The CBCL, TRF, and YSR have scales for the eight syndromes of Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior, and Aggressive Behavior. There are a few item differences across the three measures.

For the CBCL, TRF, and YSR, respondents indicate the degree or frequency of each behavior described in the item on a scale of 0 (*not true*), 1 (*somewhat or sometimes true*), or 2 (*very true or often true*). The standard rating period is 6 months for the CBCL and YSR and 2 months for the TRF. Syndrome scores are derived by summing the responses of the items in the respective scales, and then converting these raw scores to *T* scores. The CBCL, TRF, and the YSR have excellent psychometric properties, and are summarized in Achenbach and Rescorla (2001) and Achenbach et al. (2008). The ASEBA measures were not used for facilitating diagnosis in this study.

Procedure

Children and parents participated in separate interviews and testing sessions with breaks over two days. Information was also obtained from teachers using various checklists and questionnaires. In all cases, parental and child consent forms were completed prior to the assessment. The consent from both parents and children gave permission for all relevant data collected by the ACPU of the RCH or provided by others to be used in future research, and was approved by the RCH ethics committee as part of our group's comprehensive examination of psychopathology in children and adolescents. The data collected covered a

comprehensive demographic, medical (primarily neurological and endocrinological), educational, psychological, familial, and social assessment of the child and his or her family. All psychological data were collected by research assistants, who were advanced doctoral students in clinical psychology, and under the supervision of two registered clinical psychologists. Standard procedures were used for the administration of all measures. Where necessary, researchers read the YSR items to participants (approximately 5% of the sample). Approximately 85% of the parent ADISC-IV-P interviews involved mothers only, and the rest involved fathers only or both fathers and mothers together.

Statistical procedures

The CFA model in the study was analyzed using M*plus* (Version 6) software (Muthen & Muthen, 2010). Robust maximum likelihood (MLM) was used for estimation. The robust scaled chi-square statistic (called Satorra-Bentler or S-Bχ^2) is reported, but as this statistic is affected substantially by sample size (Brown, 2006), the comparative fit index (CFI), and the root mean squared error of approximation (RMSEA) were used to ascertain the model fit. The guidelines suggested by Hu and Bentler (1998) are that CFI values close to .95 or more and RMSEA values close to 0.06 or less be taken as good fit. For the CFI, values above .90 are considered acceptable fit. For the RMSEA, values between 0.07 and 0.08 are considered acceptable fit. Model fit, including models that are not nested, can be guided by the Bayesian information criterion (BIC). Lower BIC values indicate greater parsimony and fit.

Figure 1 shows the path diagram of the CT-C(M-1) model tested in this study. As shown, we used the single indicator CT-C(M-1) model, with mother ratings for the CBCL as the reference method. Although, for comparison with the Grigorenko et al. (2010) study it would have been useful to also conduct a multiple indicators (involving parcels) CT-C(M-1) model, we were not able to do so as the archival data available to us had only total scale scores, and no item scores for the CBCL, TRF and TSR. In line with model specifications, all

scales of the reference method (CBCL ratings by mothers) were linked to their appropriate trait factors and not to any method factor. Also, these indicators for the non-reference methods (TRF ratings by teachers and YSR ratings by adolescents) were linked to the appropriate traits factors and to their method factors. The trait factors correlated with each other, and the method factors correlated with each other. The loadings of the indicators of the reference method were fixed to 1 to identify the metric of the latent reference factors, and the loadings of the first indicators (Withdrawn/Depressed) of the method factors were fixed to 1 to identify the metric of the method factors. This CT-C(M-1) model estimated 93 parameters. This means that with a sample size of 294, there were about three participants for every parameter estimated (ratio of 3:1). According to Hancock (2006), for an acceptable level of data-model fit discrepancy (i.e., .02 for the RMSEA), when models have $df \geq 60$ (as is the case in the current study), to achieve power of .80, sample sizes of around 300 are sufficient for testing overall data-model fit (assuming acceptable levels). Thus, although there were three participants for every parameter estimated, with 294 participants, the study had sufficient power to test model fit.

For the CT-C(M-1) model, the standardized factor loadings for the trait and method factors were used to compute the consistency coefficients and method-specific coefficients for the indicators. In brief, the consistency coefficient of an indicator is the square of the loadings on the trait factors, and its method-specific coefficient is the square of the loadings on the trait factors. Also, one minus its consistency coefficient and method-specific coefficient is its random measurement error.

Convergent validity of an indicator is inferred if it has significant consistency coefficient. If for instance, the indicator has a larger consistency coefficient than method-specific coefficient, then it means that there is good support for its convergent validity. If the opposite is the case, then the support for convergent validity is reduced. The degree of the

correlations of the trait factors indicate the discriminant validity of the traits (as reflected in the reference method), whereas the degree of the correlations between the method factors indicate the discriminant validity of the methods. In both cases, low values support their discriminant validity. The strength of all correlations was interpreted using the guidelines proposed by Hemphill (2003) for correlation effect sizes. The guideline suggested by Brown (2006) was used to assess discriminant validity between trait factors, and method factors (i.e. correlations less than .85 were supportive of discriminant validity).

Initially, we examined the relations of the trait factors in the CT-C(M-1) model with clinical disorders by extending the CT-C(M-1) model shown in Figure 1 to include correlation paths between each of the trait factors to anxiety disorders, depressive disorders, ADHD and ODD/CD. This model failed to converge. Consequently, we used SPSS to examine these relations by examining the correlations of the factors scores for all the trait (and also source) factors obtained from the postulated CT-C(M-1) model (Figure 1) with anxiety disorders, depressive disorders, ADHD and ODD/CD diagnoses. We also used this procedure to examine how the factors scores were related to age and sex.

Results

Missing Data

Out of a total of 7,056 scores (8 scales x 3 methods x 294 participants), there were 36 scores missing (i.e., around 0.051%). Maximum likelihood (direct ML) was used to handle missing data.

Model fit

Prior to the CT-C(M-1) analysis we applied the traditional CT-CM model to evaluate convergent and discriminant validity of the syndromes scales in the CBCL, TRF and YSR. In this analysis, the scales of the same syndromes from the three measures loaded on their own respective trait factors, and these factors were correlated with each. Also, all the syndromes

from the same questionnaire loaded on their own method factors, and the method factors were correlated with each other. Trait and method factors are not correlated. This model failed to converge. Indeed, this is a general problem with the CT-CM models (Eid et al., 2003).

The fit indices for the postulated CT-C(M-1) model (with mother ratings of the CBCL as the reference method) were S-Bχ^2 (df = 207) = 569.20, $p < .001$, CFI = .917, RMSEA = .077, and BIC = 49939.67. In order to examine if this model was better that alternate models with teacher ratings as the reference method and adolescent self-ratings as the reference method, we also computed CT-C(M-1) models with ratings from these respondents as reference methods. The model with teacher ratings as the reference method were S-Bχ^2 (df = 207) = 637.51, $p < .001$, CFI = .896 and RMSEA = .085, BIC = 50009.93. For the model with adolescent self-ratings as the reference method they were S-Bχ^2 (df = 207) = 703.53, $p < .001$, CFI = .886 and RMSEA = .090, and BIC = 500078.14. Overall, the CFA and RMSEA values showed at least adequate fit for the model with mother ratings as the reference method. They showed poor fit when teacher ratings or adolescent ratings were reference methods. Also, the BIC values with mother ratings as the reference method were lower than when either teacher or adolescent ratings were reference methods. These findings indicate better fit for the CT-C(M-1) model with mother ratings as the reference method compared to CT-C(M-1) models with either teacher ratings or adolescent self-ratings as reference methods, thereby further justifying the use of mother ratings as the reference method in this study.

Intercepts and reliabilities

Table 3 provides the estimated intercepts and reliabilities of the CBCL, TRF and YSR. The estimated intercepts indicate the model-implied mean scores of the different indicators. As will be noticed, these scores were comparable to their observed mean scores in Table 1. In a CT-C(M-1) model, the reliability of an indicator is its true score in this case the

sum of its consistency coefficient and method-specific coefficient. As show in Table 3, the reliabilities of all TRF scales were at least moderate (> .56). For the YSR, the reliabilities for Anxious/Depressed, Thought Problems, and Attention problems were low (< .50), whereas the other scales had at least moderate reliabilities.

Convergent Validity of the Trait Factors

As noted previously, the proportion of the true score in the reference method (CBCL in this study) and the proportions of the true scores in the non-reference methods (TRF and YSR in this study) that are predicted by the true score of the reference method are interpreted as their trait scores, and are called consistency coefficients. The residuals in the latter predictions (or variance not accounted for by the trait score of the reference method) are interpreted as method variances, and are called method-specific coefficients. Table 3 includes the consistency coefficients of all the scales for the CBCL, TRF and YSR, and the method-specific coefficients for the non-reference methods (TRF and YSR) in our postulated CT-C(M-1) model. It shows that for all CBCL, TRF and YSR scales, all consistency coefficients were significant. For the TRF, the consistency coefficients ranged from 18, to .47, and for the YSR they ranged from .15 to .34. These findings indicate support for the convergent validity of the factors in the CBCL and TRF, and the CBCL and TRF. Despite this, there was also considerable method variance in all the indicators, as the method-specific coefficients ranged from .22 to .49 for the TRF scales, and .32 to .63 for the YSR scales.

For TRF ratings, Anxious/Depressed, Withdrawn/Depressed, and Rule Breaking Behavior had larger consistency coefficients than method-specific coefficients, whereas Somatic Complaints, Social Problems, Thought Problems, Attention Problems, and Aggressive Behavior had larger method-specific coefficients than consistency coefficients. For YSR ratings, Rule Breaking Behavior had marginally greater consistency coefficient (.36) than method-specific coefficient (.32). Anxious/Depressed, Withdrawn/Depressed,

Somatic Complaints, Social Problems, Thought Problems, Attention Problems, and Aggressive Behavior had greater method-specific coefficients than consistency coefficients. These findings reduce the support for the convergent validity for the Somatic Complaints, Social Problems, Thought Problems, Attention Problems, and Aggressive Behavior factors in CBCL and TRF; and the Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, and Aggressive Behavior factors in CBCL and YSR.

As shown in Table 3, the consistency coefficients for all TRF scales, except Aggressive Behavior, were much higher that the corresponding YSR scales. The consistency coefficients for the TRF and YSR Aggressive Behavior were comparable. Collectively, these findings indicate more agreement for mother ratings of the CBCL and teacher ratings of the TRF than mother ratings of the CBCL and adolescent ratings of the YSR (except for Aggressive Behavior).

Discriminant Validity of the Trait and Method Factors

Table 4 shows the correlations between the trait factors, as reflected in the reference method (mother ratings). As shown, apart from the correlations for Rule-Breaking Behavior with Anxious/Depressed, Withdrawn/Depressed, and Somatic Complaints, all the trait correlations were significant. For these, the correlations for Aggressive Behavior with Anxious/Depressed and Withdrawn/Depressed were of medium effect sizes, and the correlation for Aggressive Behavior with Somatic Complaints was small effect size, based on Hemphill's (2003) guidelines. All other significant correlations had large effect sizes. Despite this, with the exception of the correlations between Rule Breaking Behavior and Aggressive Behavior, all correlations were less than .85, thereby supporting the discriminant validity, of all traits, except between Rule Breaking Behavior and Aggressive Behavior.

Table 4 also includes the correlation between the method factors for adolescent and teacher. As shown, this correlation was and not significant of low effect size (.11), thereby supporting their discriminant validity.

Correlations of the Latent Trait and Method Factors with Age and Sex

Table 5 shows the correlations of the latent trait and method factors scores in the CT-C(M –1) model with age and sex. As shown, none of the latent factors correlated significantly with age and sex, and were all of low effect sizes. Thus the findings indicated that age and sex were unrelated to the latent factors.

Correlations of the Latent Trait and Method Factors with Different Disorder Categories

Table 5 also shows the correlations of the latent trait and method factors scores in the CT-C(M –1) model with the different disorder categories. As shown, the trait factors for Anxious/Depressed and Withdrawn/Depressed were significantly correlated with large effect sizes with anxiety and depressive disorders, whereas these trait factors were not significantly correlated with ADHD and ODD/CD. Somatic Complaints was significantly correlated with medium effect sizes with anxiety and depressive disorders, whereas it was significantly related with only small effect size with ADHD and not significantly correlated with ODD/CD. Social Problems was significantly correlated with either large or medium effect sizes with all disorder categories. Thought Problems was significantly correlated with medium effect sizes with anxiety and depressive disorders, whereas these trait factors were significantly correlated with small effect sizes with ADHD and ODD/CD. Attention Problems was significantly correlated with large effect size with ADHD, whereas it was significantly correlated with small effect sizes with all other disorder categories. Rule Breaking Behavior was significantly correlated with large and medium effect sizes with ODD/CD and ADHD respectively, and was not significantly correlated with anxiety and depressive disorders. Aggressive Behavior was significantly correlated with large and

medium effect sizes with ODD/CD and ADHD respectively, whereas it was correlated with small effect sizes with anxiety and depressive disorders. Taken together these findings provide good support for the external validity of the trait factors in the postulated CT-C(M – 1) model.

Discussion

The results (significant consistency coefficients) in the study indicated some support for the convergent validity of the CBCL and TRF, and CBCL and YSR for all eight scales. For the CBCL and TRF, there was more support (higher consistency coefficients than method-specific coefficients) for the convergent validity of Anxious/Depressed, Withdrawn/Depressed and Rule Breaking Behavior, and for the CBCL and YSR, there was more support (higher consistency coefficients than method-specific coefficients) for the convergent validity of Rule Breaking Behavior. The findings also showed that the level of convergence was greater between the CBCL and the TRF, than the CBCL and the YSR (higher consistency coefficients for TRF than YSR) for all scales except the Aggressive Behavior scale. The findings (correlations between the scales of less than .85) also indicated support for the discriminant validity between all scales, except between Rule-Breaking Behavior and Aggressive Behavior . There was also support for the discriminant validity of the teacher and adolescent methods. In addition, the findings showed that none of the latent traits or methods was correlated or associated with age or sex, thereby by raising the possibility that the findings were unaffected by age or sex. There was also support for the external validity of the latent trait factors, with strong correlations or associations for Anxious/Depressed and Withdrawn/Depressed with anxiety and depressive disorders, Attention Problems with ADHD, and Rule Breaking Behavior and Aggressive Behavior with ODD/CD.

Our findings supporting the convergent validity of the CBCL and TRF, and CBCL and YSR for all eight scales is not consistent with interpretations that can be made from existing data for the relevant ASEBA measures (Achenbach et al., 2002; Achenbach & Rescorla, 2001; Grigorenko et al., 2010). Previous studies, based on zero-order correlations, have also shown low to moderate convergence for the CBCL and TRF scales for Anxious/Depressed, Withdrawn/Depressed and Rule Breaking Behavior, and for the CBCL and YSR scales for Rule Breaking Behavior. The study by Grigorenko et al. (2010) that examined the convergent validity of the CBCL, TRF and YSR using the CT-C(M-1) model also reported low to moderate convergence. The amount of shared variance for the CBCL and TRF for the different factors ranged from .05 to .11, and the amount of shared variance for CBCL and YSR factors ranged from were .01 to .07.

A number of reasons may have contributed to the differences in convergence in this and previous studies. The difference in relation to zero-order correlations could be related to differences in how convergence was examined, i.e., zero-order correlations in previous studies, compared to the CT-C(M-1) procedure in the current study. In relation to the Grigorenko et al. study, it is conceivable that the differences in findings may be related to the fact that while the current study examined a group of clinic-referred adolescents, the Grigorenko et al. (2010) study examined a general community sample. Another possibility is that unlike the Grigorenko et al. study that used a multiple indicators CT-C(M-1) model, with parcels as observed indicators, the current study used the total scores of the different scales as the indicators. Since parcelling can often lead to unstable parameter estimates (Little et al., 2002; Marsh et al., 2013), the findings in the current study can be regarded as more reliable. Also, as the primary concern is the convergence with the CBCL, TRF and YSR constructs, and as we focused on the total scores, our findings can be seen as providing more useful and relevant information than that provided by Grigorenko et al (2010).

As noted earlier, the findings in this study show that the level of agreement was greater between mother ratings on the CBCL and teacher ratings on the TRF, than mother ratings on the CBCL and adolescent ratings on the YSR for all except the Aggressive Behavior scale. The meta-analysis review by Achenbach et al. (1987), and the CT-C(M-1) study of the CBCL, TRF and YSR by Grigorenko et al. (2010) showed that agreement between mothers and teachers were comparable across internalizing and externalizing behavior problems. While the review by Achenbach et al. indicated that there was more agreement between parents and adolescents for externalizing than internalizing behavior problems, this was not found in the study by Grigorenko et al. In the current study, the consistency coefficients for all TRF scales, except Rule-Breaking Behaviors and Aggressive Behavior, were much higher than the corresponding YSR scales. The consistency coefficients for the TRF and YSR scales for Rule-Breaking Behavior and Aggressive Behavior were comparable. Taken together, these findings suggest that while there is no comparable level of agreement for mother and teacher ratings with mother and adolescent ratings for internalizing problem behaviors, there is a comparable level of agreement for mother and teacher ratings with mother and adolescent ratings for externalizing problem behaviors.

The general findings of relatively more mother-teacher agreement than mother-adolescent agreement for the internalizing problems is consistent with existing data involving the ASEBA measures (Achenbach et al., 2002; Achenbach & Rescorla, 2001; Grigorenko et al., 2010) as well as other measures (Achenbach et al., 1987; Choudhury et al., 2003; Grills & Ollendick, 2003). The finding of comparable mother-teacher agreement and mother-adolescent agreement for externalizing problems is also consistent with existing data. The review by Achenbach et al. (1987) concluded that for parallel instruments (as is the case for the CBCL, TRF and YSR), the correlations between parent-teacher and parent-adolescents were higher for externalizing than for internalizing behavior problems. Although this review

also concluded that for parallel instruments there was relatively more parent-teacher agreement than parent- adolescent agreement for externalizing problems, the current study found higher levels of agreement for mother and teacher ratings with mother and adolescent ratings for externalizing problem behaviors (as reflected in the consistency coefficients for Rule-Breaking Behavior and Aggressive Behavior). At one level, cross-informant differences have been linked to the degree to which behaviours are observable, with speculation that the more observable the behaviours are, the higher the likely agreement (De Los Reyes & Kazdin, 2005). This view is consistent with the findings here as there was comparable agreement for mother ratings with teacher and adolescent ratings for externalizing problem behaviours, which clearly are more observable compared to internalizing problem behaviors.

To date, a number of theoretical explanations have been proposed for the low to moderate cross-informant agreement for child and adolescent measures. In general these explanations relate to either real differences in children's and adolescents' behaviors at home, school, and other settings (situation specificity hypothesis), or differences in respondents' perceptions of children's and adolescents' behaviors (bias hypothesis). Existing data have provided more support for the situation specificity hypothesis (Achenbach et al., 1987). If so, the findings indicating more agreement for mother-teacher reports and mother-adolescent reports for externalizing problem behaviors suggests some degree of overlap in the way adolescents manifest their behavior problems at home and at school.

The findings in this study have important clinical implications. First, as this study showed support for convergence for all CBCL and TRF, and CBCL and YSR scales, it follows that generally there is some degree of agreement between mothers and teachers, and between mothers and adolescents for the scores for comparable syndromes. On the other hand as there was also considerable method variance in all the scales, it is also the case that the

observed scores for comparable constructs include considerable unique variance that is rater specific. The findings mean that reports from the CBCL, TRF and YSR are not totally interchangeable. Thus despite the support for convergence, this highlights the need for researchers and clinicians to be cautious when integrating and interpreting scores from ASEBA measures derived from a single source, and conversely the need for obtaining ratings from multiple sources (Achenbach & Dumenci, 2001; Grigorenko et al., 2010). The findings in the study suggest that this is especially so for mother-teacher reports of somatic problems, thought problems, social problems and aggressive behaviors; and mother-adolescent reports of anxiety, depression, somatic problems, thought problems, social problems and aggressive behaviors. It will be useful to keep in mind these findings when interpreting and integrating the scores of CBCL, TRF, and YSR. Also, it would be valuable if future research could develop empirical supported algorithms for integrating the scores across these measures. Second, as this study found no support for the discriminant validity between Rule-Breaking Behavior and Aggressive Behavior, it can be argued that these scales are confounded with each other. This also means that together, these scales may be better viewed as an overall measure for externalizing problem behaviors rather than separate measures for Rule-Breaking Behavior and Aggressive Behavior. Third, the study showed strong associations for Anxious/Depressed and Withdrawn/Depressed with anxiety and depressive disorders; Attention Problems with ADHD; and Rule Breaking Behavior and Aggressive Behavior with ODD/CD. This means that when the scores for the CBCL, TRF and YSR are considered together, scores for Anxious/Depressed and Withdrawn/Depressed scales are useful in diagnosis of anxiety and depressive disorders; scores for Attention Problems are useful in diagnosis of ADHD, and scores for Rule Breaking Behavior and Aggressive Behavior are useful in diagnosis of ODD and CD

The findings and interpretations made in this study have to be viewed with respect to the study limitations. First, the ratio of participants for every estimated parameter was 3:1. Although, based on Hancock's (2006) suggestions, we estimated that the study had sufficient power to test model fit, this could be questioned as other researchers suggest higher ratios (between 5:1 and 10:1; Brown, 2006) for stable parameter estimates. Second, in the CT-C(M-1) model the method selected as the reference has potential impact on the actual findings. It is important to keep this in mind when considering the findings. Although we planned to report CT-C(M-1) analyses with adolescent and teacher as reference methods, these models showed poor fit, and we therefore abandoned these analyses. Given that there were notable differences in the reliabilities of the scales across the different respondents, with questionable reliabilities for the YSR scales for Anxious/Depressed, Thought Problems, and Attention Problems it is highly probable that different results for the convergent validity of the CBCL, TRF and YSR would have emerged if either adolescent self-ratings or teacher ratings were modelled as the reference methods. Thus the findings here can be seen as specific to these measures only when the YSR and TRS are referenced to the CBCL. Third, although the CBCL, TRF and YSR provide scale scores for the same eight syndromes, they do not have identical items (Achenbach & Rescale, 2001).While it is possible that this difference could have confounded the findings, this is unlikely as there are only 15 items in the YSR not present in the 120 items in the CBCL and TRF. As the archival data we used did not have item scores for the CBCL, TRF and TSR, we were unable to confirm this. Despite these limitations, it is argued that the collective results in this and previous studies involving the CBCL, TRF, and the YSR, combined with the comparability of the findings with other child and adolescent measures, do provide a strong psychometric and empirical basis for the continued use of these separate scales in clinical and research work.

References

Achenbach, T. H., McConaughy, S. H., & Howell, C. Y. (1987). Child/adolescent behavior and emotional problems: Implications for cross-informant correlations for situational specificity. *Psychological Bulletin, 101*, 213-232.

Achenbach, T. M., Becker, A., Dopfner, M., et al. (2008). Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: research findings, applications, and future directions. *Journal of Child Psychology and Psychiatry 49*, 251–275.

Achenbach, T. M., & Dumenci, L. (2001). Advances in empirically based assessment: Revised cross-informant syndromes and new *DSM*-oriented scales for the CBCL, YSR, and TRF: Comment on Lengua, Sadowksi, Friedrich, and Fisher. *Journal of Consulting and Clinical Psychology, 69,* 699–702.

Achenbach, T. M., Dumenci, L., & Rescorla, L. A. (2002). Ten-year comparisons of problems and competencies for national samples of youth: Self, parent and teacher reports. *Journal of Emotional and Behavioral Disorders, 10,* 194–203.

Achenbach, T.M., & Rescorla, L.A. (2001). *Manual for the ASEBA School-Age Forms & Profiles*. Burlington,VT: University of Vermont, Research Center for Children, Youth, and Families.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders – IV text revised*. Washington, DC: Author.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 2*, 81–105.

Choudhury, M. S., Pimentel, S. S., & Kendall, P. C. (2003). Childhood anxiety disorders:

Parent-child (dis)agreement using a structured interview for the DSM–IV. *Journal of the American Academy of Child & Adolescent Psychiatry, 42,* 957–964.

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.

Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait–multimethod models: A multiple-indicator CT-C(*M*–1) model. *Psychological Methods, 8,* 38–60.

Epkins, C. C. (1993). A preliminary comparison of teacher ratings and child self-report of depression, anxiety, and aggression in inpatient and elementary school samples. *Journal of Abnormal Child Psychology, 21*, 649-661.

Grills, A. E., & Ollendick, T. H. (2002). Issues in parent– child agreement: The case of structured diagnostic interviews. *Clinical Child and Family Psychology Review, 5,* 57–83.

Grigorenko, E. L., Geiser, C., Slobodskaya, H. R., Francis, D. J. (2010). Cross-Informant symptoms from CBCL, TRF, and YSR: Trait and method variance in a normative sample of Russian youths. *Psychological Assessment, 22*, 893–911

Hancock, G. R. (2006). Power analysis in covariance structure modelling. In G. H. Hancock and R. O. Mueller (Eds.), Structural equation modeling: A second course. Greenwich, Connecticut. IAP Publication.

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist, 58*, 78–79.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3,* 424-453.

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.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002).To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9,* 151–173.

Lord, F. N., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Marsh, H. W., Lüdtke, O., Nagengast, B., Morin, A. J. S., & Von Davier. M. (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right-camouflaging misspecification with item parcels in CFA models. *Psychological Methods, 18*, 257–284.

Muthen, L. K., & Muthen, B. O. (2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthen & Muthen.

Sawyer, M. G., Baghurst, P., & Mathias, J. (1992). Differences between informants' reports describing emotional and behavioural problems in community and clinic-referred children: A research note. *Journal of Child Psychology and Psychiatry, 33*, 441-449.

Silverman, W. K., & Albano, A. M. (1996). *Manual for the ADIS-IV C/P*. New York: Psychological Corporation.

Silverman, W. K., Saavedra, L. M., & Pina, A. A. (2001). Test–retest reliability of anxiety symptoms and diagnoses with the Anxiety Disorders Interview Schedule for DSM-IV: child and parent versions. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 937–944.

van den Ende, J., & Verhulst, F. C. (2005). Informant, gender and age differences in ratings of adolescent problem behaviour. *European Child and Adolescent Psychiatry, 14,* 117–126.

Table 1

Means and *SD* of the *T* Scores of the CBCL, TRF and YSR Syndromes

	CBCL		TRF		YSR	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Anxious/Depressed	68.95	11.03	62.70	10.42	61.69	11.14
Withdrawn/Depressed	67.64	10.88	62.90	9.95	60.31	9.56
Somatic Complaints	67.45	10.16	56.95	8.83	61.35	10.04
Social Problems	68.80	10.32	65.70	10.11	62.29	10.21
Thought Problems	69.84	8.91	62.25	10.39	62.23	10.24
Attention Problems	70.21	11.27	65.78	10.51	65.10	11.84
Rule Breaking Behavior	67.71	9.65	65.66	10.92	60.30	9.30
Aggressive Behavior	73.75	13.27	67.97	13.98	63.93	11.60

Note. CBCL = Child Behavior Checklist; TRF = Teacher's Report Form; YSR = Youth Self-Report.

Table 2

Frequency and Percentage of Different Categories and Comorbidity of Disorders for Participants

	Frequency	Percentage
Disorder Categories		
Anxiety Disorders	228	77.3
Dysthymia/Major Depressive Disorder	153	51.9
Attention Deficit/Hyperactivity Disorder	217	73.6
Oppositional Defiant/Conduct Disorder	206	69.8
Comorbidity		
No diagnosis	11	3.7
1 disorder	35	11.9
2 disorders	61	20.7
3 disorders	105	35.6
4 disorders	83	28.1

Note. Any Anxiety Disorder includes Separation Anxiety, Social Phobia, Specific Phobia, Panic, Agoraphobia, Generalized Anxiety, Obsessive Compulsive and/or Post-Traumatic Stress disorders. For individuals with three or four disorder categories, the highest comorbidity was for anxiety disorders, ADHD and ODD/CD. For those with two disorder categories, the highest comorbidity was between ADHD and ODD/CD.

Table 3

Results of the CT-C(M –1) Model: Standardized Factor Loadings, Reliability of Non-Reference Methods, and Variance Components for Observed Scores

	Loading				Variance		
Ratings	Intercept	Trait	Method	Reliability	Consistency Coefficient	Method-Specific	Random Error
			Anxious/Depressed				
CBCL	64.45	1.00		.63	.63		
TRF	56.94	0.51	0.65	.65	.35	.30	.35
YSR	61.35	0.55	0.72	.75	.15	..60	.25
			Withdrawn/Depressed				
CBCL	68.94	1.00		.65	.65		
TRF	62.69	0.79	1.00	.59	.35	.24	.41
YSR	61.68	0.52	1.00	.67	20	.47	.33
			Somatic Complaints				
CBCL	67.64	1.00		.61	.61		
TRF	62.89	0.74	0.80	.40	.18	.22	.60
YSR	60.31	0.51	0.77	.56	.17	.39	.44
			Social Problems				
CBCL	68.79	1.00		.87	.87		
TRF	64.69	0.62	1.13	.78	.30	.48	.22
YSR	62.28	0.56	0.89	.74	.23	.51	.26
			Thought Problems				
CBCL	69.84	1.00		.50	.50		
TRF	62.25	0.71	1.15	.67	.21	.46	.33

YSR	62.23	0.53	0.88	.66	.13	.53	.34
Attention Problems							
CBCL	70.21	1.00		.63	.63		
TRF	64.78	0.70	1.10	.79	.32	.47	.21
YSR	65.09	0.64	1.14	.82	.19	.63	.18
Rule Breaking Behavior							
CBCL	67.70	1.00		.80	.80		
TRF	64.65	0.88	0.93	.79	.47	.32	.21
YSR	60.29	0.69	0.62	.68	.36	.32	.32
Aggressive Behavior							
CBCL	73.74	1.00		.79	.79		
TRF	67.96	0.69	1.48	.80	.31	.49	.20
YSR	63.92	0.63	0.95	.80	.34	.46	.20

Note. Consistency = proportion of variance that is shared with the reference method (mother); method-specific = proportion of variance that is specific to the non-reference method (and not shared with the reference method); reliability = proportion of reliable observed variance (sum of consistency and method-specific for observed scores); CBCL = Child Behavior Checklist; TRF = Teacher's Report Form; YSR = Youth Self-Report. All loadings, and consistency coefficients and method-specific coefficients were significant ($p < .001$).

Table 4

Correlations of the Latent Trait and Method Factors in the CT-C(M –1) Model

	1	2	3	4	5	6	7	8	9
Trait									
Anxious/Depressed (1)	-								
Withdrawn/Depressed (2)	.81***	-							
Somatic Complaints (3)	.82***	.69***	-						
Social Problems (4)	.69***	.53***	.56***	-					
Thought Problems (5)	.78***	.68***	.70***	.80***	-				
Attention Problems (6)	.38***	.30***	.39***	.71***	.71***	-			
Rule Breaking Behavior (7)	.06	.12	.06	.36***	.40***	.62***	-		
Aggressive Behavior (8)	.27***	.21**	.15**	.60***	.67***	.75***	.89***	-	
Method									
Teacher (9)								-	
Youth (10)									.11

*p<.05, **p<.01, ***p<.001

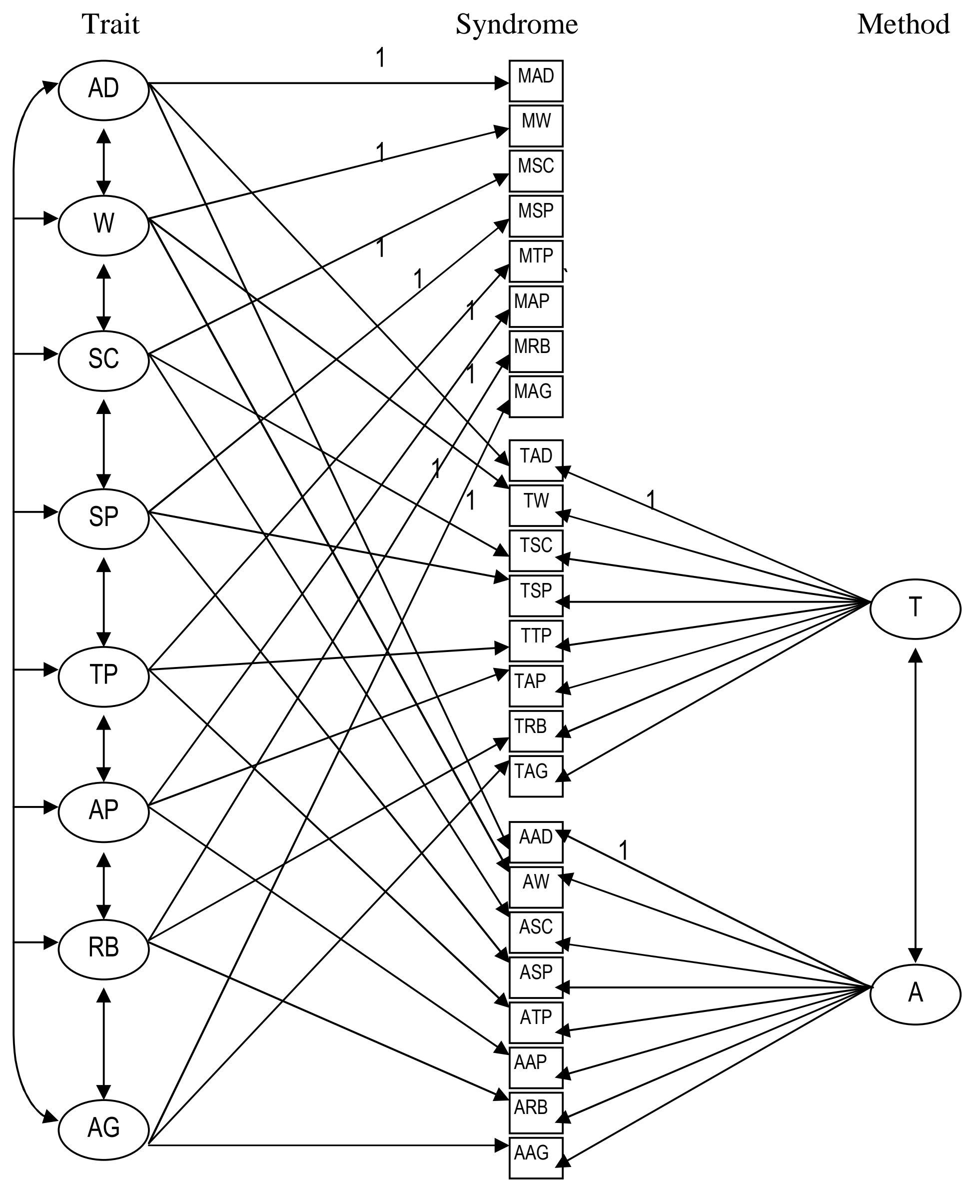
Table 5

Correlations of the Latent Trait and Method Factors Scores in the CT-C(M –1) Model with Age and Sex, and Different Disorder Categories,

	Age	Sex	Anxiety	Depression	ADHD	ODD/CD
			Disorders			
Anxious/Depressed	.,03	-.06	.37**	.36**	-.01	.09
Withdrawn/Depressed	-.06	-.10	.31**	.38**	.04	.07
Somatic Complaints	.04	.01	.26**	.26**	.13*	.03
Social Problems	.05	.07	.30**	.26**	.23**	.27**
Thought Problems	-.01	.01	.29**	.27**	.16**	.19**
Attention Problems	.08	.04	.14*	.12*	.44**	.17**
Rule Breaking Behavior	.04	.11	.07	.10	.29**	.53**
Aggressive Behavior	.05	.03	.17**	.18**	.25**	.60**
			Sources			
Teacher	.00	.01	.07	.10	.29**	.53**
Youth	.00	.00	.17**	.18**	.25**	.60**

Note. For the variable sex, girls are coded with 0, and boys are coded 1.

*p<.05, **p<.01.



Note. AD = Anxious/Depressed, W = Withdrawn/Depressed, SC = Somatic Complaints, SP = Social Problems, TP = Thought Problems, AP = Attention Problems, RB = Rule Breaking Behavior, AG = Aggressive Behavior; M = mother; A = adolescent self-rating; T = teacher.

Figure 1. The correlated trait-correlated method minus one model used in the study.