CrossMark

The Validity of the Child Problematic Trait Inventory in 6–12 Year Old Italian Children: Further Support and Issues of Consistency Across Different Sources of Information and Different Samples

Antonella Somma [1,2] · Henrik Andershed [3] · Serena Borroni [4,2] · Andrea Fossati [1,2]

Abstract In the present study, teacher's ratings, mother's ratings, and father's ratings on the Child Problematic Trait Inventory (CPTI) were obtained on two independent samples ($N=381$ and $N=406$, respectively) of Italian school age children (age range: 6–12 years). Cronbach alpha values greater than 0.80 were observed for all CPTI scale scores based on teacher's ratings, mother's ratings, and father's ratings respectively. Although intraclass correlation values suggested poor parent-teacher agreement, robust weighted least square Confirmatory Factor Analyses (WLSMV CFAs) suggested that the three-factor model of the CPTI was consistently replicated as the best fitting model across all informants. Logistic regression analysis conducted in one of the samples showed that teacher's ratings and mother's ratings of the CPTI, but not CPTI father's ratings, significantly predicted child's problem behavior at school. Multiple regression analyses showed that CPTI scores were significantly predicted by selected temperament dimensions, but with different relationships for different sources of information. Overall, our findings suggest that the CPTI is a reliable and valid measure of problematic traits with a three-factor structure, although this study indicates that multiple sources of information may prove necessary when assessing psychopathy in school-age children with the CPTI.

Keywords Child problematic traits inventory · Psychopathy · Childhood · Parents and teachers' ratings

Although different conceptions of psychopathy exist, psychopathy usually refers to a dysfunctional personality syndrome which is characterized by prominent behavioral deviancy in the presence of distinctive emotional and interpersonal features (Patrick et al. 2009). Influenced by a number of prominent theorists including Cleckley (1941), Hare (1991), Karpman (1941), and Lykken (1957, 1995), current conceptions of psychopathy include reference to features such as superficial charm, manipulativeness, egocentricity, callousness, a lack of remorse or empathy, impulsivity and irresponsibility, along with a marked risk for violence and other criminal behaviors (Cooke and Michie 2001; Marcus et al. 2013), although criminality may represent a correlate (or a consequence) of psychopathy rather than a core feature of this maladaptive personality syndrome (e.g., Cleckley 1941; Cooke and Michie 1997; Lykken 1957, 1995). Research findings have quite consistently documented associations between psychopathy and a wide range of externalizing behaviors such as crime and aggression (e.g., Gretton et al. 2004; Porter et al. 2001), criminal recidivism (e.g., Walters et al. 2008), substance use (e.g., Gustavson et al. 2007; Kennealy et al. 2007), and sexual offending (Caldwell et al. 2008).

The existing scientific literature does not support the notion that psychopathy comprises a unitary diagnostic construct. For example, factor analytic studies of the widely-used Psychopathy Checklist-Revised (PCL-R; Hare 1991, 2003) in adult criminal samples indicate that it taps distinguishable components, although alternative models with differing

✉ Andrea Fossati
fossati.andrea@hsr.it

[1] Department of Human Studies, LUMSA University, Rome, Italy

[2] Clinical Psychology and Psychotherapy Unit, San Raffaele Turro Hospital, via Stamira d'Ancona 20, 20127 Milan, Italy

[3] Center for Criminological and Psychosocial Research, Örebro University, Örebro, Sweden

[4] Faculty of Psychology, Vita-Salute San Raffaele University, Milan, Italy

numbers of factors have been proposed (Patrick et al. 2009). The best-known structural model of the PCL-R is the two-factor model (Harpur et al. 1988; Hare et al. 1990), in which Factor 1 encompasses the interpersonal and affective features of psychopathy and Factor 2 encompasses the antisocial deviancy features. However, Cooke and Michie (2001) proposed an alternative three-factor model of the PCL-R psychopathy item set. Notably, three-factor models have also been reported for other psychopathy inventories in adult nonoffender samples (Benning et al. 2003), in youth offender samples (Frick et al. 2000), and in youth nonoffender samples (Andershed et al. 2002). A four-factor model of the PCL-R was subsequently proposed by Hare and Neumann (2006) as an alternative to the three-factor model.

Research on the latent structure of psychopathy has shown it may be not categorical, but rather dimensional (e.g., Marcus et al. 2004; Murrie et al. 2007), like the wide majority of normative and dysfunctional personality features. Therefore, this would suggest that there are not differences in kind per se, but rather differences in degree. However, when individuals are high on psychopathy, they may suffer (and cause) severe personal, social, and legal problems, and psychopathy treatment is required. Treatability of psychopathy has been a controversial issues for many years, although a growing number of clinicians believes that psychopathic subjects can be treated (Tennent et al. 1993). Salekin et al. (2010) carried out a systematic review on the efficacy of treatments for psychopathy with adults and youth; their findings suggested that treatment for adults shows low to moderate success with three of eight studies demonstrating treatment gains, whereas treatment of youth appeared to be more promising with six of eight studies showing treatment benefits. As a whole, Salekin et al. (2010) review did not support the notion that psychopathy is untreatable; rather, Salekin et al. (2010) findings indicated that the earlier psychopathic subjects are treated, the better the treatment gains.

Psychopathy in Youth

These considerations lead to a growing interest in recognizing and assessing psychopathy in adolescents and children. Indeed, investigating psychopathy in these age groups may help clinicians and researchers to gain insights into the different pathways toward severe antisocial behavior, and to understand the etiology of this severe adult personality disorder, as well as to offer preventive interventions or early treatment programs (van Baardewijk et al. 2011). To date, studies of psychopathic traits in youth have generally yielded results which are strikingly similar to those in adults in terms of stability, relations to conduct disorder, aggressive behavior, and emotional and cognitive functioning (for reviews see, e.g., Andershed 2010; Lynam 2010). Recently there has been

an increased interest in investigating the personality traits related to psychopathy in non-criminal populations of children and adolescents (Ribeiro da Silva et al. 2012; Roose et al. 2012; Salekin 2010; Salekin and Lynam 2010; van Baardewijk et al. 2010). Moreover, Andershed et al. (2002) have shown that psychopathic traits in non-referred youths manifest similarly to how they are manifested among incarcerated offenders, as indicated by similarities in factor structures.

The Assessment of Psychopathic Traits in Children and the Child Problematic Traits Inventory

Although the Psychopathy Checklist: Youth Version (PCL:YV; Forth et al. 2003) is often considered to be the most reliable and valid measure of psychopathic-like traits among forensic youth (Andershed et al. 2007), measures for assessing psychopathy dimensions in community samples of adolescents and children are needed (Colins et al. 2012). Although self-report instruments are deemed useful for research on non-incarcerated adolescents because they are easy to use and fast to administer (Andershed et al. 2007), assessing psychopathy in non-incarcerated young children requires specific instruments based on observer ratings. Indeed, there is reason to assume that preadolescents themselves are capable of rating these traits through self-report (van Baardewijk et al. 2011); however before that age, children are too young to provide self-report (e.g., Goldberg 2001). Then, children's psychopathic traits are usually assessed by teachers and caregivers (e.g., Frick et al. 2000).

Several instruments that were specifically designed to assess psychopathic traits in children and adolescents are currently available (for a review see: Kotler and McMahon 2010). Yet, none of these instruments were designed for use in young children. Only one of them, the Antisocial Process Screening Device (APSD; Frick and Hare 2001) allows for assessment of psychopathic traits in relatively young children (from age 6), whereas most other instruments were developed for use in late childhood (e.g., Youth Psychopathic Traits Inventory—Child Version; van Baardewijk et al. 2008) and/or adolescence (e.g., Childhood Psychopathy Scale; Lynam 1997). Previous studies reported unsatisfactory reliability estimates for the ASPD scales (Dadds et al. 2005; Kimonis et al. 2012). Available studies gave mixed findings also for the Inventory of Callous Unemotional Traits (ICU; Essau et al. 2006). On the one hand, the proposed bifactorial factor model of the ICU was not replicated in a sample of 3-year-old children; on the other hand, callous-unemotional (CU) traits were generally associated with variables of interest (e.g., effortful control, executive functioning; Ezpeleta et al. 2013). Moreover, the ICU does not assess features relating to the

interpersonal dimension nor the behavioral dimension of the psychopathic personality (Colins et al. 2014).

In recent years, these controversial findings prompted the development of several assessment tools with the aim to measure psychopathic traits in childhood (e.g., Dadds et al. 2005; Waller et al. 2012). In order to assess the three-factor model of psychopathy in children, Colins et al. (2014) recently proposed the Child Problematic Traits Inventory (CPTI; Colins et al. 2014). In developing the CPTI, the authors aimed to assess a construct of psychopathic personality in childhood, including early childhood, which closely resembled how it is usually conceptualized and assessed in adolescence and adulthood (e.g., Andershed et al. 2002; Cooke and Michie 2001). Only those psychopathic traits that have theoretical and/or empirical support for being applicable and assessable in children from 3 to 12 years of age were included in the CPTI (Colins et al. 2014). Moreover, the CPTI did not include traits or behaviors that bear resemblance to rule-breaking, conduct problems, and antisocial behavior; in fact, the CPTI was designed to measure psychopathic traits that in turn can be used to understand and predict conduct problems (Colins et al. 2014). In particular, Colins et al. (2014) developed the CPTI to measure a childhood version of the three-factor model of psychopathic personality (e.g., Andershed et al. 2002; Cooke and Michie 2001) rather than a single dimension such as the CU dimension. Also, the CPTI was developed as a measure to be primarily rated by preschool teachers/teachers. Teacher assessments have been found to be relevant for measuring personality traits of children too young to provide self-reports (e.g., Hampson and Goldberg 2006). Teachers are familiar with the child across a variety of classroom and other settings (e.g., recess), and they can be argued to be in a very good position to make normative judgments because of their extensive experience with children of the same ages.

Based on their literature review, Colins et al. (2014) developed 28 items which were explicitly designed to assess three broad dimensions of psychopathy: 1. an interpersonal dimension characterized by grandiose sense of self-worth, lying, and deceitfulness (the so-called Grandiose-Deceitful [GD] factor); 2. an affective dimension characterized by lack of remorse or guilt and callousness/lack of empathy (the so-called Callous-Unemotional [CU] factor); and 3. an impulsive/need for stimulation dimension characterized by need for stimulation/sensation-seeking and proneness to boredom, and impulsivity (the so-called Impulsive-Need for Stimulation [INS] factor). The reliability and validity of the CPTI was initially tested in a Swedish general population sample of 2056 3- to 5-year-olds (Colins et al. 2014). The CPTI items loaded distinctively on the three theoretically proposed factors, i.e., GD, CU and INS factors. The three CPTI factors showed adequate internal consistency reliability estimates (all CPTI scales exhibited Cronbach α values≥0.89) and external validity, in terms of expected correlations with theoretically relevant constructs, such as fearlessness and conduct problems.

The Present Study

We aimed at extending previous data on the psychometric properties of the CPTI to older children, from a different cultural context, and using different sources of information.

In particular, in the present study, we administered the Italian translation of the CPTI to two independent samples of 6–12 year-old Italian children from the Italian general population attending elementary school. A recent development in the discussion concerning psychopathic traits in children pertains to the use of informants; indeed, correlations between scores on measures of psychopathy using different informants have generally been low, possibly indicating that a single (external) source of information is not covering the full manifestation of the construct (van Baardewijk et al. 2011). Therefore, in the present study, in both samples, each child was independently rated on the CPTI by his/her teacher and parents; thus, for each child we had a set of teacher-rated CPTI scores, a set of mother-rated CPTI scores, and a set of father-rated CPTI scores.

In the first sample, we aimed to evaluate the internal consistency, as well as the mean-level and absolute consistency of the CPTI scores based on teacher's ratings, mother's ratings, and father's ratings, respectively. Moreover, we aimed to evaluate the factor structure of the CPTI items validity of the CPTI scales across the two independent data sets.

In the second sample, we aimed to assess the replicability of Sample 1 findings concerning the reliability and factor structure of the CPTI scores across the three sources of information that participated in this study. In order to evaluate the nomological network validity of the CPTI, in the second sample we gathered official grades of children's conduct from school records in order to identify children with problem behavior at school; behavior grades were regressed on teacher-rated CPTI scores, mother-rated CPTI scores, and father-rated CPTI scores, respectively. Finally, teachers and parents were also asked to rate Sample 2 children on measures of theoretically-relevant temperament dimensions (i.e., fearlessness and easy temperament). In particular, we expected the CPTI total score to be negatively predicted by measures of easy temperament and positively predicted by measures of fearlessness. Similarly, we expected a negative relationship between the CPTI total score and measures of effortful control and behavior inhibition, and a positive association between the CPTI total score and measures of approach/positive affect and anger/irritability. This allowed us to evaluate if the CPTI scores were significantly predicted by theoretically-relevant temperament features both within-source of information (e.g., teacher-rated CPTI scores were regressed on teacher-

rated personality measures) and between-sources of information (e.g., teacher-rated CPTI scores were regressed on mother-rated personality measures).

Method

Participants

Sample 1 was composed of 381 children attending a public elementary school in Sora, a town of 26,589 inhabitants in Central Italy, near Rome. One hundred ninety nine (52.2 %) children were female, and 182 children (47.8 %) were male; the children's age was on average 8.78 years, $SD=1.64$, ranging from a minimum of 6 years to a maximum of 12 years. The mean age of the mothers of Sample 1 children was 40.08 years, $SD=4.91$ years; the fathers of Sample 1 children were on average 43.78 years old, $SD=5.68$ years. According to their parents' reports, two hundred eighty nine (75.9 %) children were from married couples, and nine children (2.4 %) were from unmarried couples; seventy-six (19.9 %) children were from divorced families, two (0.5 %) children had a widow parent, whereas five (1.3 %) parents did not report their civil status. The average socio-economic status (SES), which was assessed using Hollingshead's (2011) four-point index based on parents' education, occupation, sex, and marital status, was 39.02, $SD=12.18$. Based on the four-point index of SES (Hollingshead 2011), 11 (2.9 %) parents belonged to "unskilled laborers, menial service workers" category, 64 (16.8 %) parents belonged to the "machine operators, semi-skilled workers", 117 (30.7 %) parents belonged to the "skilled craftsmen, clerical, sales workers" category, 113 (29.7 %) parents belonged to the "medium business, minor professional, technical" category, and 33 (8.7 %) parents belonged to the "major business and professional" category; 43 (11.3 %) parents did not report information necessary to compute the four-point SES index.

Sample 2 was composed of 406 children who were attending a public elementary school in Rome, Italy. Two hundred nine (51.5 %) children were female, and 197 children (48.5 %) were male; the children's age was on average 8.49 years, $SD=1.56$, ranging from a minimum of 6 years to a maximum of 12 years. The mean age of the mothers of Sample 2 children was 39.43 years, $SD=5.67$ years; the fathers of Sample 2 children were on average 43.27 years old, $SD=6.57$ years. According to their parents' reports, 303 (74.6 %) children were from married couples, and 30 children (7.4 %) were from unmarried couples; 57 (14.0 %) children were from divorced families, 5 (1.2 %) children had a widowed parent, whereas 11 (2.7 %) parents did not report their civil status. The parents' average four-point SES index (Hollingshead 2011) was 32.77, $SD=16.24$. Based on the four-point index of SES (Hollingshead 2011), 102 (25.1 %) parents belonged to "unskilled laborers, menial service workers" category, 73 (18.0 %) parents belonged to the "machine operators, semi-skilled workers", 64 (15.8 %) parents belonged to the "skilled craftsmen, clerical, sales workers" category, 88 (21.7 %) parents belonged to the "medium business, minor professional, technical" category, and 49 (12.1 %) parents belonged to the "major business and professional" category; 30 (7.4 %) parents did not report information necessary to compute the four-point SES index.

Children's gender did not significantly differentiated the two samples, $\chi^2 (1)=0.05$, $p>.80$, $\varphi=0.01$, although Sample 2 children were significantly, albeit slightly older than Sample 1 children, $t(785)=2.60$, $p<.01$, $d=0.19$. No significant difference between Sample 1 children and Sample 2 children in mother's mean age, $t(785)=1.70$, $p>.05$, $d=0.12$, and in father's mean age, $t(785)=1.09$, $p>.20$, $d=0.08$. Parents' civil status significantly differentiated Sample 1 from Sample 2, $\chi^2 (3)=15.18$, $p<0.01$, Cramer $V=0.14$; according Bonferroni-corrected z-tests for comparing proportions, a significantly lower rate of unmarried parents (2.4 % vs. 7.6 %) and a significantly higher rate of divorced parents (20.2 % vs. 14.4 % was observed in Sample 1 when compared to Sample 2. The number of parents who did not report their civil status in Sample 1 ($n=5$, 1.3 %) was not significantly different from the frequency of missing values on civil status in Sample 2 ($n=11$, 2.7 %), $\chi^2 (1)=1.93$, $p>.10$, $\varphi=0.05$. Parents who did not report their civil status did not significantly differ on CPTI scale scores from parents who reported their civil status both in Sample 1 (min. t (379; INS scale, father's ratings)$=-1.01$, $d=-0.10$, max. t (379; GD scale, teacher's ratings)$=1.23$, $d=0.13$, all $ps>0.20$) and in Sample 2 (min. t (404; CU scale, teacher's ratings)$=-0.27$, $d=-0.03$, max. t (404; CPTI total score, father's ratings)$=1.47$, $d=0.15$, all $ps>0.10$).

Considering parents who reported information that are necessary to compute SES index ($N=714$), on average, parents of Sample 1 children showed a significantly higher SES score than parents of Sample 2 children, separate-variance t (690)$=5.85$, $p<.001$, $d=0.45$. In line with this finding, the five SES categories were not homogeneously distributed in Sample 1 and in Sample 2, $\chi^2 (4)=93.87$, $p<.001$, Cramer $V=0.36$. According to Bonferroni-corrected z test for comparing proportions, "skilled craftsmen, clerical, sales worker" (34.6 % vs. 17.0 %) and "medium business, minor professional, technical" (33.4 % vs. 23.4 %) categories were significantly more represented in Sample 1 parents than in Sample 2 parents; at the opposite, "unskilled laborers, menial service worker" category was significantly more frequent among Sample 2 parents ($n=102$, 27.1 %) than among Sample 1 parents ($n=11$, 3.3 %). Sample 1 parents ($n=43$, 11.3 %) did not significantly differ from Sample 2 parents ($n=30$, 7.4 %) in the frequency of parents who did not report information on their SES, $\chi^2 (1)=3.55$, $p>.05$, $\varphi=0.07$. Parents who did not report information on their SES did not significantly differ on CPTI

scale scores from parents who reported their SES both in Sample 1 (min. t (379; GD scale, mother's ratings)=−0.09, d=−0.01, max. t (379; INS scale, father's ratings)=0.99, d= 0.10, all ps>0.20) and in Sample 2 (min. t (404; INS scale, father's ratings)=0.31, d=0.03, max. t (404; GD scale, teacher's ratings)=1.09, d=0.11, all ps>0.20).

In order to avoid cultural and linguistic bias in the evaluation of psychopathic traits, only children from parents who spoke Italian as a first language were included in both Sample 1 and Sample 2. In both samples, parents also had to sign written informed consent forms to allow their children to participate in the study, as well as to participate themselves in the study. A separate informed consent form was filled by all the teachers who participated in the study. Before contacting personally the children's teachers and parents, the study was formally approved by the principals and school boards of the two elementary schools which participated in the research. Thus, only children who received a consent to participate in the study by their parents and their teachers were included in the final samples; in both Sample 1 and Sample 2 all children were assessed independently by their teachers and by their parents. In the case of divorced parents, only the parent with whom the child was living most of the time was asked to rate the child's behavior.

In both samples, teachers and parents used the Italian translation of the CPTI in order to assess the children's behavior; each teacher was asked to use to CPTI to assess the behavior of all children in her/his classroom, whereas parents were asked to use the CPTI to rate only their own children's behavior; in order to confirm and extend the construct validity of the CPTI, different outcome measures were assessed by the teachers and/or parents in Sample 2. In particular, in Sample 2 the official grade for behavior was reported for each child by his/her teacher and was used as a measure of the child's possible problem behavior at school; moreover, both teachers and parents were asked to rate measures of the child's easy temperament and fearlessness (Colins et al. 2014), and measures of approach/positive affect, irritability/anger, fear/behavioral inhibition, and effortful control (Rothbart et al. 2000). In both samples, parents were asked to rate independently their own children's behavior after routine meeting with the teachers at school. In order to ensure the child's anonymity, the CPTI and all additional measures were assessed using forms reporting only an alphanumeric code based on school records.

Measures

Both Sample 1 and Sample 2 children were administered the CPTI by their teachers and parents, respectively. In both samples, parents' SES was assessed by computing Hollingshead's (2011) four-point SES index.

Child Problematic Traits Inventory (CPTI; Colins et al. 2014) The CPTI was developed as a research instrument that could enable longitudinal studies in order to test developmental theories and stability of psychopathic traits and psychopathic personality across different developmental phases. In particular, the CPTI was designed to assess psychopathic traits from age 3 to 12. The items of the CPTI were developed using a theory-driven approach in the sense that they were developed based on the three-factor model of psychopathy (Andershed et al. 2002; Cooke and Michie 2001); CPTI item content was assumed to represent the childhood manifestations of the psychopathic traits included in the three-factor model (Colins et al. 2014). Based on these considerations, the 28 CPTI items were assigned to three scales which are thought to assess a corresponding psychopathic personality trait: a) Grandiose-Deceitful (GD) scale which is composed of 8 items measuring grandiose sense of self-worth, lying, and deceitfulness; b) Callous-Unemotional (CU) scale which comprises 10 items assessing lack of remorse, callousness, and lack of empathy; c) Impulsive-Need for Stimulation (INS) scale which is composed of 10 items measuring sensation seeking, proneness to boredom, and impulsivity.

In line with the Youth Psychopathic traits Inventory (Andershed et al. 2002), each CPTI item is measured on a four-point Likert scale, *1= Does not apply at all; 2= Does not apply well; 3= Applies fairly well; and 4= Applies very well*. For each scale, the item sum is divided by the number of items composing the scale in order to obtain the scale score. In addition to providing a score for each of the three psychopathic personality traits, the CPTI yields also a total score which is a measures of the general level of psychopathy; the higher the CPTI score, the higher the psychopathy level.

On the first page of the CPTI, information is given concerning the overall content of the items and that the rater should assess each item based on how the child usually and typically behaves rather than based on how he or she behaves at the moment. The reliability and validity of the original version of the CPTI in a sample of preschool children have been documented (Colins et al. 2014).

Parents' Socio-Economic Status The four-point SES index (Hollingshead 2011) was used to evaluate the parents' SES. The four-point SES index takes into account that SES is a multidimensional construct; the four factors used in the new index are: education, occupation, sex, and marital status (Hollingshead 2011). Information on each of the four factors is easily gathered in an empirical study. In the absence of theoretical and empirical evidence, a rule of thumb is followed, that is, education and occupation scores for the husband and wife are summed and divided by two (Hollingshead 2011). Research has indicated that the prestige of occupations is similar for males and females and that education is essentially the same for males and females in the same occupation

(Treiman and Terrell 1975). The years of school a respondent has completed are scored on a seven-point scale, premised upon the assumption that men and women who possess different levels of education have different tastes and tend to exhibit different behavior patterns (Hollingshead 2011). The occupation a person ordinarily pursues during gainful employment is graded on a nine-step scale (Hollingshead 2011). The status score of an individual or a nuclear family unit is estimated by combining information on sex, marital status, education, and occupation. The status score of an individual is calculated by multiplying the scale value for occupation by a weight of five and the scale value for education by a weight of three. To calculate the status score for a nuclear family it is necessary to determine the education, occupation, and marital status of its head or heads and their relationship to the labor force in the present, or for retired persons in the past (Hollingshead 2011). The four-point SES index scores range from a high of 66 to a low of 8; the higher score of a family or nuclear unit, the higher the status its members are accorded by other members of our society (Hollingshead 2011).

In addition to the CPTI, Sample 2 participants were administered also the Easy Temperament measure, the Fearlessness measure, and measures based on Mary Rothbarth's (Rothbart et al. 2000) temperament dimensions of approach/positive affect, irritability/anger, fear/behavioral inhibition, and effortful control. For each child who participated in Sample 2, behavior grades from the school records were also obtained.

Easy Temperament Measure (ETM; Colins et al. 2014)
This three-item measure was developed by Colins et al. (2014) and it is designed to be rated by teachers on the following response scale, *1= Does not apply at all, 2= Applies poorly, 3= Applies fairly well, 4= Applies well*. The three items are: "He/she is happy and positive", "He/she is a child who other children want to play with", and "He/she is easy and nice to be with". The items were framed to comprise the behavior of the child during the last 6 months. The Easy temperament measure total score is computed as the mean of the three items. In a previous study, the Easy temperament measure showed adequate internal consistency reliability (Colins et al. 2014). In the present study, the Easy temperament measure was rated by parents and teachers, respectively.

Fearlessness Measure (FM; Colins et al. 2014) The FM measure is a six- item scale that was designed to be primarily rated by teachers. Examples of items are: "He/she does not seem to be afraid of anything", "He/she does not seem to be afraid when someone is trying to frighten him/her", and "He/she never seems to get scared when someone is mad at him/her." Each item is measured on a four-point Likert scale, *1= Does not apply at all, 2= Applies poorly, 3= Applies fairly well, 4= Applies well*. The items were framed to comprise the behavior of the child during the last 6 months. The FM

total score is simply computed as the mean of the six items. The original version of the FM showed adequate internal consistency reliability in a sample of preschool children (Colins et al. 2014).

Child's Temperament Measure Although several measures of child's temperament for school-age children are currently available (for instance, see Rothbart and Mauro 1990), they were generally too time-consuming to be used in the present study by the teachers. Thus, we developed a short measure of child's temperament that was based on Rothbart et al.' (2000) model of temperament and personality in children. In particular, we aimed at assessing four major temperament dimensions, a) Approach/Positive affect (AP), which is related to the development of positive anticipation and outgoing activity, but it is also linked to later potential problems with control, including impulsivity, anger/frustration and lower inhibitory control (Kochanska 1993, 1997; Rothbart et al. 2000); b) Fear (F), which is an important control system for the development of conscience; Rothbart and Bates (1998) have also reviewed several additional instances of earlier fear as a protective factor for later-developing aggression. Behavioral inhibition, a component of fear, represents an important control of behavior and can be used in socialization to support children's control of their actions through fear of punishment or disappointment (Kochanska 1991); c) Irritability/Anger (IA), which seems to be a factor that is predisposing to later externalizing negative affect but not to fear. How others react to these tendencies toward anger and frustrative distress, especially in the development of mutually coercive cycles, is also important in the development of later externalizing outcomes (for a review, see Rothbart and Bates 1998); d) Effortful Control (EC), which is characterized by high inhibitory control and attentional capacities. Consistent data indicates that children in the United States who are high in effortful control tend to be low in negative affectivity (Ahadi et al. 1993) and that adults high in self-reported attentional control are likely to be low in negative affect (Derryberry and Rothbart 1988; Evans and Rothbart 1999).

Based on Rothbart et al.' (2000) model of child temperament, a set of three adjectives was extracted from the description of each temperamental dimension (Rothbart et al. 2000); each adjective was measured on a five-points Likert scale, *1= Not at all or very little, 2= A little, 3= Moderately, 4= A lot, 5= Extremely*. Both teachers and parents were instructed to rate how accurately each adjective described the child's usual way to feel and behave; parents received identical indications on how to use the temperament measure. The AP scale was composed by the following adjectives, "Excited", "Curious", and "Active"; the F scale included the following adjectives, "Fearful", "Scared", and "Frightened". The IA scale was composed of the following adjectives, "Edgy", "Irritable", and "Hostile". Finally, the EC scale included the following

adjectives, "Alert", "Attentive", and "Persistent". For each scale, the item scores were summed and averaged to yield the scale total score.

Behavior Grade as Indicator of Problem Behavior at School School grades for behavior that were obtained from official school records were used to evaluate the presence of problem behavior at school. In Italian elementary school, the grades for behavior are given independently from school performance and are not expressed using a numerical/letters grade scale; rather, teachers have to resort to detailed evaluations and summary judgments based on four categories, *Insufficient, Sufficient, Good,* and *Excellent.* Behavior grades are assigned by the classroom board and take into account the child's usual behavior at school. In order to receive *Insufficient* in behavior, children should a) have received disciplinary actions; and b) have shown no positive behavioral changes in response to these actions. In the case of *Insufficient* in behavior, failure may occur.[1] In the case of disciplinary actions with at least partial behavioral changes, *Sufficient* in behavior is usually considered. In other terms, both *Insufficient* and *Sufficient* in behavior suggest the presence of problem behaviors at school.

Measure Translation Procedures

Equivalence with the original meaning of the items was the guiding principle in the translation process (Denissen et al. 2008). First, the CPTI was translated into Italian by one of the authors (A.F.). Then a second author who was fluent in English (A.S.) reviewed the translation independently. After reaching a consensus, a professional translator with English as his first language translated the Italian version back into English. If the latest version differed from the English original, the first author, the second author, and the professional translator came to an agreement on the definitive Italian translation (Cha et al. 2007; Van de Vijver and Hambleton 1996; Geisinger 1994). The authors followed the same procedure of translation concerning the ETM and FM.

Data Analyses

Cronbach's alpha was used to assess the internal consistency of the CPTI (as well as of all the other questionnaires that were used in this study). Mean-level consistency of the CPTI scores across teacher's ratings, mother's ratings, and father's ratings was tested using repeated-measure generalized; in the case of

---

[1] Italian school system allows children not to pass the year because of problem behavior. In this case, failure is completely independent from the child's school performance; failure is due to the fact that the child's behavior is incompatible with passing the school year. Indeed, in order to pass the year, each child is supposed to have both a sufficient performance in each subject and a sufficient behavior grade.

significance of Mauchley's sphericity test, Pillai V was used as omnibus test for mean equality. In case of significance, the multivariate test was followed by Bonferroni-protected paired-sample t-tests. In order to explore if the mean-level consistency of the CPTI scores could be replicated in the two samples, belonging to Sample 1 or Sample 2 was entered as between-group effect in the repeated-measure analyses. Intraclass correlation (ICC) coefficient for absolute rater agreement based on one-way random effects ANOVA (Nunnaly and Bernstein 1994) was used to evaluate the consistency of the CPTI scores across different sources of information (i.e., teacher, mother, and father). Alsawalmeh and Feldt's (1992) T statistic was used to evaluate the presence of significant differences between ICC values computed in sample 1 and the corresponding ICC values computed in Sample 2; Alsawalmeh and Feldt's (1992) T is distributed as an independent F variable. Within each sample, the null hypothesis equality of dependent ICCs was assessed using Donner and Zou's (2002) confidence intervals. The significance of the difference between independent correlation coefficients was tested using z test; the Silver et al.'s (2006) z test was used to evaluate the difference between nonoverlapping dependent correlation coefficients.

Following McCrae et al.' (1996) suggestions, dimensionality analyses were carried out before performing CFA. In this study, parallel analysis (Buja and Eyuboglu 1992) was used to evaluate if a three-factor solution could adequately explain to correlations among the CPTI items. Parallel analysis entails comparing the eigenvalues of a correlation matrix of P random uncorrelated variables with those of the data set in question, based on an equivalent sample size (Zwick and Velicer 1986). Quasi-inferential parallel analysis makes use of the full null distribution; using upper 5 or 1 % null quantiles instead of means may reduce some liberal tendencies of parallel analysis – i.e., extraction of minor components (Buja and Eyuboglu 1992). Following Buja and Eyuboglu's (1992) suggestions, parallel analysis was based on 1000 random permutation of the original data; a given eigenvalue was considered to explain nonrandom variance if it exceeded the 95th upper percentile of the distribution of the corresponding 1000 random eigenvalues (Buja and Eyuboglu 1992). Performing dimensionality analyses before carrying out CFA may sound an excess of accuracy; rather, it seems to represent a useful approach. Indeed, CFA may easily lead to rejecting correct models when nonsimple factor structure are at issue, as it is frequently the case in personality research (e.g., McCrae et al. 1996). Thus, knowing that the number of factors underlying the variable correlation (or variance–covariance) matrix matches the hypothesized number of factors helps the researcher to identify the reasons of poor fit in CFA without unduly resorting to models with extra dimensions.

In order to evaluate the factor structure of the CPTI, Confirmatory Factor Analyses (CFAs) of the CPTI item

polychoric correlation matrices were carried out separately in Sample 1 and Sample 2. Based on Flora and Curran (2004) suggestions, robust WLS (i.e., weighted least squares mean and variance adjusted; WLSMV) method was used in order to compute the weight matrix. Since WLSMV algorithm does not allow for testing the significance in improvement of fit between two nested models using the regular procedure of chi-square difference, we used the appropriate chi-square difference testing procedure (i.e., DIFFTEST; Muthén and Muthén 1998–2012). CFAs were carried out separately for teacher's CPTI scores, mother's CPTI score, and father's CPTI scores.

For each source of information and within each sample, the following factor models of the CPTI items were tested in CFAs: a) a unidimensional model (i.e., one-factor model), in which all CPTI items were assigned to a single latent dimension. b) a two factor model, which was based on the results of a promax-rotated two-factor exploratory factor analysis solution which was suggested by the MAP statistic. CPTI items 1, 3, 5, 6, 7, 9, 10, 12, 14, 15, 16, 18, 19, 21, 23, 24, 26, 28, were assigned to Factor 1, whereas CPTI items 2, 4, 8, 11, 13, 17, 20, 22, 25, 27, were assigned to Factor 2. Based on item content, Factor 2 was closely akin to a CU dimension, whereas Factor 1 collapsed indicators of Grandiosity/Deceitfulness and Need for Stimulation/Irresponsibility. Factor 1 and Factor 2 were allowed to be correlated, and the errors were forced to be independent; c) a three-factor model, in which the CPTI items were assigned to GD (CPTI items 5, 7, 9, 15, 18, 21, 24, and 26), CU (CPTI items 2, 4, 8, 11, 13, 17, 20, 22, 25, and 27), and INS (CPTI items 1, 3, 6, 10, 12, 14, 16, 19, 23, and 28) dimensions. It should be noted that this model corresponds to the a priori theoretical structure of the CPTI items. Factors/dimensions were allowed to correlate with each other, and the errors were forced to be independent; d) a three-factor model, in which as in model c) the CPTI items were assigned to GD, CU, and INS dimensions; however, in Model d) some error parameters were allowed to correlate each other. It is known that error correlation may arise in scale development from several sources, such as measurement method (e.g., similar number of reverse scored items in the scales), multidimensionality of the scales, and similarity in wording among the items that compose different scales (see, for instance, Bollen 1989). The relevant number of psychometric articles based on CFA that included correlated error terms documents the importance of these sources of error correlation in psychometric literature (Schweitzer 2012). Because strong a priori hypotheses about the structure of the error correlations of the CPTI scales would be premature, in the present study model modification indices were used to refine Model d) error matrix. However, to reduce the risk of capitalizing on chance, the same pattern of error correlation was cross-validated in Sample 2.

Considering the multidimensional nature of fit assessment in structural equation models and the need for multiple fit indices (Tanaka 1993), we used several measures to identify model fit, including the WLSMV χ^2 goodness-of-fit statistic, Browne and Cudeck (1993) root mean square error of approximation (RMSEA), the Tucker–Lewis index (TLI), Bentler's (1990) comparative fit index (CFI). Following Hu and Bentler (1999) suggestions, TLI, CFI, and IFI values ≥0.95, and RMSEA values close to 0.06 were considered as indicating good model fit, whereas TLI and CFI values of 0.90 and higher, and an RMSEA of 0.08 and lower are indications of an adequate fit (but see Marsh et al. 2004, for a comment on these commonly applied cutoff values for assessing model fit).

In Sample 2, logistic regression analyses were carried out in order to evaluate the predictive role on behavior grade of the CPTI scale scores. McFadden generalized R^2 statistic (R^2_{McF}) and Cox and Snell pseudo-R^2 ($R^2_{C\&S}$) were used to evaluate the predictive efficiency of CPTI scores. Separate logistic regression analyses were carried out using teacher's CPTI ratings, mother's CPTI ratings, and father's CPTI ratings, respectively.

In Sample 2, multiple hierarchical regression analyses were performed in order to evaluate if temperament measure scores significantly predicted the CPTI scales and total score, respectively. Potential confounders (e.g., parents' SES, child's gender, etc.) were entered in step 1, whereas temperament scales were entered in step 2 of the hierarchical regression model.

Hierarchical regression analyses were carried out within-source of information (for instance, using teacher's ratings of fearlessness and easy temperament to predict teacher's CPTI total score) as well as between-source of information (for instance, using mother's ratings of fearlessness and easy temperament to predict teacher's CPTI total score).[2] Changes in R^2 and adjusted R^2 values when child's temperament scales were entered in step 2 of the regression equation were computed as effect size measures in each regression model. The Variance Inflation Factor (VIF) was computed in order to assess collinearity. Usually, VIF values greater than five or even ten are considered to suggest problems with collinearity (Kutner et al. 2004).

SMH z tests, Donner and Zou's (2002) 95 % confidence intervals for equality of dependent ICCs, and polychoric correlations were computed using R statistical software (R Development Core Team 2014). Dimensionality analyses were computed using FACTOR 9.2 (Lorenzo-Seva and Ferrando 2006). CFAs were carried out using Mplus 7.3 (Muthen and Muthén 2014). All other statistical analyses were performed using SPSS 22.

---

[2] We chose not to enter all informants' ratings (i.e., teachers' ratings, mothers' ratings, and fathers' ratings) in step 2 because we were interested in testing the contribution of each predictor in terms of amount of variance explained in the dependent variable, and we preferred to avoid capitalizing on predictor collinearity in estimating regression coefficients (and their standard errors).

Results

CPTI Descriptive Statistics and Internal Consistency Reliability Results

CPTI scale descriptive statistics, internal consistency reliability estimates (i.e., Cronbach's α values), and gender comparisons in Sample 1 and in Sample 2 are listed in Table 1. In Sample 1, children's age did not correlate significantly with any of the CPTI ratings, median $r=0.02$, $SD=0.04$, min. r value (CPTI GD scale, teacher's ratings)$=-0.06$, max. r value (CPTI INS scale, teacher's ratings)$=0.08$, all ps>0.10. Similar findings for the correlations between CPTI ratings and children's age were observed also in Sample 2, with r values ranging from -0.08 (CPTI INS scale, mother's ratings) to 0.04 (CPTI GD scale, teacher's ratings), median $r=-0.03$, $SD=0.04$. In Sample 1 parents' SES score showed significant, albeit weak correlations only with mother's ratings on the CPTI CU scale, $r=-0.18$, $p<.001$, with mother's ratings on the CPTI total score, $r=-0.15$, $p<.01$, with father's ratings on the CPTI CU scale, $r=-0.20$, $p<.001$, and with father's ratings on the CPTI total score, $r=-0.16$, $p<.01$; none of the correlations between parents' SES and teacher's CPTI ratings reached statistical significance. Almost identical findings were observed in Sample 2; SES scores correlated significantly, albeit weakly with mother's ratings of the CPTI CU scale, $r=-0.18$, $p<.001$, with mother's ratings of the CPTI total score, $r=-0.13$, $p<.05$, with father's ratings of the CPTI CU scale, $r=-0.18$, $p<0.001$, and with father's ratings of the CPTI total score, $r=-0.12$, $p<.05$.

On average (see Table 1), teacher's CPTI ratings were not significantly different in Sample 1 and in Sample 2, with Student t (785) values ranging from -0.56 ($d=-0.04$; CU scale) to 1.90 ($d=0.14$; GD scale), all ps>0.05. Similar considerations held also for mother's and father's CPTI ratings, respectively, with t (785) values ranging from -0.37 ($d=-0.03$; GD scale, mother's ratings) to 1.68 ($d=0.12$; CU scale, mother's ratings). When the effect of parents' SES was held constant in ANCOVA analyses, no significant difference was observed between on Sample 1 and Sample 2 on any of the CPTI mean scores that are reported in Table 1, min. $F(1, 784)=0.00$ ($\eta^2=0.00$; CPTI total score, mother's ratings), max. $F(1, 784)=1.47$ ($\eta^2=0.00$; CPTI GD scale, teacher's ratings), all ps>0.20.

In Sample 1, when teacher's ratings were considered, Pearson r values for bivariate association between CPTI scales were 0.56, 0.74, and 0.60, all ps<0.001, for the correlations between GU and CU scales, GU and INS scales, and CU and

Table 1 Child Problematic Traits Inventory factors and total score: Descriptive statistics, reliability coefficients, and gender comparisons in Sample 1 ($N=381$) and Sample 2 ($N=406$) participants

	Sample 1 ($N=381$)													Sample 2 ($N=406$)												
	Females ($n=199$)			Males ($n=182$)				Whole sample ($N=381$)			Females ($n=209$)			Males ($n=197$)				Whole sample ($N=406$)		
	M	SD	α	M	SD	α	d	M	SD	α	M	SD	α	M	SD	α	d	M	SD	α
CPTI teacher ratings																				
CPTI GD factor	1.23	0.37	0.86	1.53	0.57	0.90	−0.63*	1.37	0.50	0.90	1.4O	0.46	0.93	1.52	0.54	0.94	−0.25	1.46	0.51	0.93
CPTI CU factor	1.45	0.55	0.93	1.67	0.62	0.94	−0.38*	1.56	0.59	0.94	1.69	0.60	0.94	1.76	0.65	0.91	−0.25	1.72	0.63	0.93
CPTI INS factor	1.65	0.49	0.87	1.93	0.63	0.90	−0.49*	1.78	0.58	0.89	2.04	0.51	0.91	2.18	0.63	0.93	−0.42*	2.11	0.57	0.93
CPTI Total Score	1.46	0.39	0.93	1.72	0.54	0.96	−0.56*	1.59	0.48	0.95	1.73	0.44	0.95	1.84	0.52	0.97	−0.35*	1.79	0.48	0.96
CPTI mother ratings																				
CPTI GD factor	1.46	0.50	0.84	1.50	0.51	0.86	−0.07	1.48	0.50	0.85	1.46	0.51	0.84	1.53	0.55	0.85	−0.24	1.50	0.53	0.85
CPTI CU factor	1.61	0.59	0.87	1.70	0.55	0.86	−0.16	1.65	0.57	0.86	1.76	0.64	0.86	1.76	0.6	0.89	−0.12	1.76	0.62	0.88
CPTI INS factor	2.10	0.59	0.86	2.13	0.56	0.84	−0.05	2.12	0.58	0.85	2.06	0.53	0.79	2.17	0.65	0.86	−0.23	2.11	0.59	0.84
CPTI Total Score	1.74	0.46	0.92	1.79	0.46	0.92	−0.12	1.77	0.46	0.92	1.78	0.47	0.91	1.84	0.51	0.93	−0.22	1.81	0.49	0.92
CPTI father ratings																				
CPTI GD factor	1.47	0.50	0.84	1.48	0.53	0.87	−0.01	1.47	0.51	0.85	1.38	0.55	0.86	1.53	0.69	0.87	−0.13	1.45	0.62	0.86
CPTI CU factor	1.67	0.63	0.89	1.73	0.60	0.87	−0.09	1.70	0.62	0.88	1.46	0.51	0.89	1.61	0.66	0.87	−0.01	1.53	0.59	0.88
CPTI INS factor	2.08	0.56	0.84	2.14	0.57	0.84	−0.09	2.11	0.57	0.84	1.71	0.58	0.81	1.99	0.75	0.88	−0.17	1.85	0.68	0.85
CPTI Total Score	1.76	0.47	0.92	1.80	0.48	0.92	−0.09	1.78	0.48	0.92	1.53	0.48	0.92	1.72	0.64	0.93	−0.12	1.62	0.57	0.92

CPTI Child Problematic Traits Inventory, *GD* CPTI Grandiose-Deceitful factor, *CU* CPTI Callous-Unemotional Factor, *INS* CPTI Impulsive-Need for Stimulation factor; within each sample, nominal significance level ($p<.05$) for gender comparisons (Student t-tests) on CPTI variables was corrected according to the Bonferroni procedure and set at $p<.0042$; *d*: Cohen *d* effect size measure; asterisks mark mean differences and *d* values that were significant at Bonferroni-corrected nominal significance level

INS scales, respectively. When mother's ratings were taken into account, the CPTI GD scales showed substantial, significant correlations with both CU, $r=0.53$, $p<.001$, and INS scales, $r=0.64$, $p<.001$, whereas the CPTI CU scale was significantly correlated with the INS scale, $r=0.46$, $p<.001$. Finally, when father's ratings were considered, Pearson r values for bivariate association between CPTI scales were 0.52, 0.69, and 0.47, all ps<0.001, for the correlations between GU and CU scales, GU and INS scales, and CU and INS scales, respectively.

None of the correlation coefficients between CPTI scales scores based on mother's ratings differed significantly from the corresponding correlation based on father's ratings, with SHM z-values ranging from -1.65 (correlations between CPTI GD and INS scales) to 0.31 (correlations between CPTI GD and CU scales), all ps>0.05. The correlation coefficients between GD and INS scales based on teacher's ratings and on mother's ratings, respectively, were significantly different, SMH $z=2.54$, $p<.05$; similar findings were observed also for the difference in the values of correlation coefficients between CU and INS scales, SMH $z=2.65$, $p<.01$. The correlation between GD and INS scales based on teacher's ratings did not differ significantly from the corresponding correlation based on mother's ratings, SMH $z=0.50$, $p>.60$. Only the correlation between CPTI CU and INS scales based on teacher's ratings was significantly different from the corresponding correlation based on father's ratings, SMH $z=2.33$, $p<.05$.

In Sample 2, according to teacher's ratings, the CPTI GD scales showed substantial, significant correlations with both CU, $r=0.78$, $p<.001$, and INS scales, $r=0.73$, $p<.001$, whereas the CPTI CU scale was significantly correlated with the INS scale, $r=0.68$, $p<.001$. Based on mother's ratings, Pearson r values were 0.55, 0.66, and 0.47, all ps<0.001 for the association between GU and CU scales, GU and INS scales, and CU and INS scales, respectively. Finally, according to father's ratings, the CPTI GD scales showed substantial, significant correlations with both CU, $r=0.57$, $p<.001$, and INS scales, $r=0.62$, $p<.001$, whereas the CPTI CU scale was significantly correlated with the INS scale, $r=0.48$, $p<.001$.

In line with Sample 1 results, none of the correlations between the CPTI scales based on mother's ratings differed significantly from the corresponding correlation based on father's ratings, with SMH z values ranging from -0.58 (correlations between CPTI GD and CU scales) to 1.25 (correlations between CPTI GD and INS scales), all ps>0.20 The correlation coefficients between GD and CU scales based on teacher's ratings and on mother's ratings, respectively, were significantly different, SMH $z=5.98$, $p<.001$; similar findings were observed also for the difference in the values of correlation coefficients between CU and INS scales, SMH $z=4.51$, $p<.001$. Consistent with Sample 1 findings, in Sample 2 the correlation between GD and INS scales based on teacher's ratings did not

differ significantly from the corresponding correlation based on mother's ratings, SMH $z=1.92$, $p>.05$. Different from what was observed in Sample 1, in Sample 2 all correlations between CPTI scales based on teacher's ratings differed significantly from the corresponding correlations based on father's ratings, with SMH z values ranging from 2.85 (correlations between CPTI GD and INS scales) to 5.54 (correlations between CPTI GD and CU scales), all ps<0.005.

Considering the comparison between correlations among CPTI scales that were observed in Sample 1 and those that were observed in Sample 2, with the exception of the correlation between GD and CU based on teacher's ratings, $z=5.78$, $p<.001$, none of the remaining differences between correlation coefficients reached statistical significance.

CPTI Scale Mean-Level Consistency Across Different Sources of Information: Repeated-Measure Analysis Results

Since Mauchley's sphericity test was highly significant, we relied on multivariate omnibus test to identify the presence of any significant difference among mean CPTI total scores in repeated-measures analyses, $W=0.51$, χ^2 (2)= 419.97, $p<.001$. Repeated-measure MANOVA showed that the mean CPTI total scores based on teacher's ratings, mother's ratings, and father's ratings, respectively, were not homogeneous, Pillai $V=0.11$, $p<.001$; no significant sample-by-rating interaction was observed, Pillai $V=0.00$, $p>.90$. The nominal significance level (i.e., $p<.05$) for post hoc comparisons was corrected according to the Bonferroni procedure and set at $p<.016$ (i.e., $p_{Bonferroni}<0.05$). Bonferroni-corrected paired-sample t-test showed no significant difference between mothers and fathers on mean CPTI total score, $d=-0.07, p_{Bonferroni}>0.05$; rather, the mean CPTI total score based on teacher's ratings significantly differed from mean CPTI total scores based on both mather's ratings, $d=-0.30$, $p_{Bonferroni}<0.05$, and father's ratings, $d=-0.33$, $p_{Bonferroni}<0.05$.

Similar considerations held also for the differences in mean scores among our three sources of information on the CPTI GD, within-subject Pillai $V=0.02$, $p<.005$, sample-by-rating interaction Pillai $V=0.01$, $p>.10$, CU, within-subject Pillai $V=0.06$, $p<.001$, sample-by-rating interaction Pillai $V=0.01$, $p>.10$, and INS scales, within-subject Pillai $V=0.22$, $p<.005$, sample-by-rating interaction Pillai $V=0.00$, $p>.50$, respectively. According to Bonferroni-protected post hoc comparisons, no significant difference was observed between mother's ratings and father's ratings on the CPTI scales, with the exception of a small, albeit significant difference on the CU scale, $d=-0.10$, $p_{Bonferroni}<0.05$. Rather, teacher's ratings were significantly lower than both mother's ratings (min. d [GD scale]$=-0.10$, max. d [INS scale]$=-0.43$, all $p_{Bonferroni}$ <0.05) and father's ratings (min. d [GD scale]$=-0.13$, max. d

[INS scale]$=-0.43$, all $p_{Bonferroni}<0.05$) on all CPTI dimensions, although the effect size estimates for these differences were all in the small-to-moderate range. Based on Donner and Zou's (2002) confidence intervals for the equality of dependent ICCs, in both samples ICCs for the agreement between mother's ratings and father's ratings were significantly larger than the corresponding ICC values for the agreement between teacher's ratings, and mother's and father's ratings.

Intraclass correlation coefficients (ICCs) for absolute agreement between pairs of raters are listed in Table 2. Alsawalmeh and Feldt's (1992) statistic showed that ICC values that were observed in Sample 1 were significantly lower than the corresponding Sample 2 ICC values, min. $T=1.21$, max. $T=1.36$, all $ps<0.01$, with the only exception of the ICC values for father's ratings and mother ratings on the CPTI total score, $T=1.12$, $p>.05$, and for father's ratings and teacher's ratings on the CPTI INS scale, $T=1.07$, $p>.10$. As a whole, the ICC estimates for the agreement between mother's ratings and father's ratings were larger than those between teacher's ratings, and mother's and father's ratings, respectively; Donner and Zou (2002) 95 % confidence intervals for the difference between dependent ICCs ranged from (lower limit/upper limit) 0.49/0.69 (INS scale, father's ratings) to 0.56/0.78 (CU scale, father ratings) in Sample 1, and from 0.37/0.54 (GD scale, both mother's and father's ratings) to 0.51/0.70 (CU scale, fathers ratings) in Sample 2. According to Donner and Zou (2002) 95 % confidence intervals, none of the ICC values based on mother's ratings differed significantly from the corresponding ICC values based on father's ratings.

Dimensionality Analysis, and Confirmatory Factor Analysis Results

In Sample 1, the values of the Kaiser-Meyer-Olkin (KMO), statistic for the CPTI item polychoric correlation matrices were 0.94, 0.92, and 0.92 for teacher's, mother's ratings, and father's ratings, respectively. Results of quasi-inferential parallel analyses are summarized in Table 3.

The following CPTI factor models were assessed in CFAs: a) a unidimensional model in which all CPTI items load on a common latent dimension; b) a two-factor model in which CU items define Factor 1 and the remaining CPTI items define Factor 2; and c) the theoretical three-factor model of CPTI item assignment.

Goodness-of-fit statistics and incremental fit indices for WLSMV CFAs in Sample 1 and in Sample 2, respectively, are listed in Table 4. In both samples the three-factor model of proved to be the best fitting model for CPTI item polychoric correlation matrices based on teacher's ratings, mother's ratings, and father's ratings. With the exception of RMSEA value for teacher's ratings, all other fit statistics supported the theoretical three-factor model of CPTI items. When we considered teacher's ratings in Sample 1, modification indices suggested significant error correlations between item 4 ("Usually does not seem to share others' joy and sorrow") and item 2 ("Seldom expresses sympathy for others"), item 7 ("Seems to see himself/herself as superior compared to others") and item 24 ("Thinks that he/she is better than everyone at almost everything"), and item 11 ("Often seems to be completely indifferent when other children are upset") and item 13 ("Does not become upset when others are being hurt"), respectively. Interestingly, this pattern of error correlation was

Table 2 Child Problematic Traits Inventory: Intraclass correlation coefficients for absolute agreement among raters in Sample 1 and Sample 2, respectively

	Sample 1 (N=381)		Sample 2 (N=406)	
	Teachers' rating	Mothers' rating	Teachers' rating	Mothers' rating
CPTI Grandiose-Deceitful factor				
Mothers' ratings	0.12	–	0.36*	–
Fathers' ratings	0.13	0.77*	0.36*	0.81*
CPTI Callous-Unemotional				
Mothers' ratings	0.06	–	0.22*	–
Fathers' ratings	0.03	0.70*	0.19*	0.79*
CPTI Impulsive-Need Stimulation				
Mothers' ratings	0.18*	–	0.34*	–
Fathers' ratings	0.21*	0.80*	0.26*	0.84*
CPTI total score				
Mothers' ratings	0.13	–	0.34*	–
Fathers' ratings	0.14	0.81*	0.29*	0.83*

CPTI Child Problematic Traits Inventory; within each sample, nominal significance level ($p<.05$) was corrected according to the Bonferroni procedure and set at $p<.0028$; asterisks mark intraclass correlation coefficients that were significant at Bonferroni-corrected nominal significance level

Table 3 Dimensionality analyses of the child problematic traits inventory polychoric correlation matrices: parallel analysis results based on 1000 random permutations of sample 1 original data (N=381)

	Teacher's ratings		Mother's ratings		Father's ratings	
	Real data eigenvalue	95th percentile of random eigenvalues	Real data eigenvalue	95th percentile of random eigenvalues	Real data eigenvalue	95th percentile of random % of variance
1	15.43[a]	2.43	11.69[a]	2.03	11.69[a]	2.03
2	2.76[a]	1.74	2.72[a]	1.71	2.72[a]	1.71
3	1.87[a]	1.62	1.68[a]	1.60	1.68[a]	1.60
4	1.21	1.53	1.29	1.51	1.29	1.51
5	0.94	1.45	1.05	1.44	1.05	1.44
6	0.63	1.39	0.94	1.38	0.94	1.38
7	0.59	1.33	0.81	1.33	0.81	1.33
8	0.50	1.28	0.71	1.28	0.71	1.28
9	0.46	1.23	0.66	1.23	0.66	1.23
10	0.40	1.18	0.64	1.18	0.64	1.18
11	0.39	1.13	0.57	1.14	0.57	1.14
12	0.35	1.09	0.55	1.10	0.55	1.10
13	0.32	1.04	0.49	1.06	0.49	1.06
14	0.28	1.00	0.46	1.02	0.46	1.02
15	0.26	0.96	0.44	0.98	0.44	0.98
16	0.22	0.92	0.41	0.94	0.41	0.94
17	0.20	0.88	0.37	0.90	0.37	0.90
18	0.18	0.85	0.36	0.87	0.36	0.87
19	0.16	0.81	0.34	0.83	0.34	0.83
20	0.15	0.77	0.32	0.79	0.32	0.79
21	0.14	0.73	0.30	0.76	0.30	0.76
22	0.12	0.69	0.27	0.72	0.27	0.72
23	0.11	0.66	0.22	0.69	0.22	0.69
24	0.11	0.62	0.20	0.65	0.20	0.65
25	0.09	0.58	0.17	0.61	0.17	0.61
26	0.07	0.54	0.14	0.57	0.14	0.57

[a] Eigenvalue/percentage of variance exceeding the 95th percentile of random data distribution

consistently reproduced across the three sources of information that were considered in this study, as well as across our two independent samples. Indeed, including these correlated error terms in model d) significantly improved model fit.

In Table 5, robust weighted least square factor loadings in Sample 1 and in Sample 2, respectively, are presented.

Nomological Network Validity of the CPTI

In Sample 2, only five children (1.2 %) had "*insufficient*" in behavior according to official school records; thus, they were pooled with children (n=39, 9.6 %) having poor grades, albeit sufficient in behavior, to generate a "problem behavior at school" group (n=44, 10.8 %); children with "adequate" or "excellent" in behavior according to official school records were pooled to compose a "adequate behavior at school" group (n=362, 89.2 %). This dichotomous variable was used as dependent variable in binary logistic regression analyses in which the CPTI scales based on different sources of information were entered as predictors. In all logistic regression models, child's gender and parents' SES was entered as covariate.

Child's male gender (odds ratio=4.21, 95 % confidence interval=1.92, 9.22) and parents' SES (odds ratio=0.97, 95 % confidence interval=0.95, 0.99) significantly predicted problem behavior at school in the first step of logistic regression models, McFadden R^2=0.06, Cox and Snell pseudo-R^2= 0.05, omnibus test for coefficient significance χ^2 (2)=19.32, p<.001 (% of correctly classified children=65.1). When the CPTI total score based on teacher's ratings was entered in the logistic regression equation, McFadden R^2 value was 0.22 and Cox and Snell pseudo-R^2 value was 0.15, Wald χ^2 (1)=34.44, p<.001; the odds ratio for the effect of CPTI score based on teacher's rating in predicting child's problem behavior at school in logistic regression analysis was 5.76, 95 % confidence interval=3.21, 10.33 (% of correctly classified children=80.8).

Table 4 Weighted least square mean and variance adjusted confirmatory factor analysis results of the items of the Child Problematic Traits Inventory in Sample 1 (N=381), and Sample 2 (N=406)

CPTI item models	Sample 1 (N=381)						
	χ^2	df	$\Delta\chi^2$	RMSEA	90 % CI for RMSEA	CFI	TLI
Teachers' ratings (n= 351)							
a) 1 factor	2292.537***	350	–	0.126	0.121–0.131	0.888	0.879
b) 2 correlated factors	1280.018***	349	124.174***	0.087	0.082–0.092	0.947	0.942
c) 3 correlated factors	1238.148***	347	30.260***	0.086	0.080–0.091	0.949	0.944
d) 3 correlated factors with correlated error terms	1072.289***	344	159.393***	0.078	0.072–0.083	0.958	0.954
Mothers' ratings (n=373)							
a) 1 factor	1656.204***	350	–	0.100	0.095–0.105	0.834	0.821
b) 2 correlated factors	978.488***	349	115.002***	0.070	0.064–0.075	0.920	0.913
c) 3 correlated factors	817.655***	347	64.390***	0.060	0.055–0.066	0.940	0.935
d) 3 correlated factors with correlated error terms	763.381***	344	61.690***	0.057	0.052–0.063	0.947	0.941
Fathers' ratings (n=352)							
a) 1 factor	1584.272***	350	–	0.100	0.095–0.105	0.842	0.829
b) 2 correlated factors	938.930***	349	109.125***	0.069	0.064–0.075	0.924	0.918
c) 3 correlated factors	859.276***	347	49.711***	0.065	0.059–0.070	0.934	0.928
d) 3 correlated factors with correlated error terms	792.747***	344	72.637***	0.061	0.055–0.066	0.942	0.937
CPTI item models	Sample 2 (N=406)						
	χ^2	df	$\Delta\chi^2$	RMSEA	90 % CI for RMSEA	CFI	TLI
Teachers' ratings (n= 392)							
a) 1 factor	2233.265***	350	–	0.117	0.113–0.122	0.923	0.916
b) 2 correlated factors	1787.003***	349	100.264***	0.103	0.098–0.107	0.941	0.936
c) 3 correlated factors	1402.676***	347	109.190***	0.088	0.083–0.093	0.957	0.953
d) 3 correlated factors with correlated error terms	1166.604***	344	125.845***	0.078	0.073–0.083	0.966	0.963
Mothers' ratings (n=395)							
a) 1 factor	1711.989***	350	–	0.099	0.095–0.104	0.855	0.844
b) 2 correlated factors	1018.560***	349	134.150***	0.070	0.065–0.075	0.929	0.923
c) 3 correlated factors	875.582***	347	64.373***	0.062	0.057–0.067	0.944	0.939
d) 3 correlated factors with correlated error terms	818.107***	344	75.089***	0.059	0.054–0.064	0.950	0.945
Fathers' ratings (n=336)							
a) 1 factor	1586.453***	350	–	0.103	0.097–0.108	0.852	0.840
b) 2 correlated factors	1102.946***	349	103.234***	0.080	0.075–0.086	0.910	0.902
c) 3 correlated factors	876.373***	347	273.095***	0.067	0.062–0.073	0.936	0.931
d) 3 correlated factors with correlated error terms	812.739***	344	80.722***	0.064	0.058–0.069	0.944	0.938

CPTI Child Problematic Traits Inventory, $\Delta\chi^2$ chi-square difference between nested models (DIFFTEST analysis results), RMSEA root mean square error of approximation, CI confidence interval, CFI comparative fit index, TLI Tucker-Lewis index;—: statistic not computed

** $p<.01$; *** $p<.001$

Similar results were observed when the CPTI total score based on mother's ratings was used as predictor in the second step of logistic regression analysis (child's male gender and parents' SES were entered in the first step), McFadden R^2= 0.11, Cox and Snell pseudo-R^2=0.08, Wald χ^2 (1)=10.31, $p<.005$, odds ratio=3.11, 95 % confidence interval=1.56, 6, 21 (% of correctly classified children=72.0). Rather, logistic regression analyses evidenced no significant effect of the

CPTI total score based on father's ratings in predicting child's problem behavior at school, at least when the effect of child's male gender and parents' SES was controlled for, Wald χ^2 (1)=2.64, $p>.10$, odds ratio=1.82, 95 % confidence interval=0.88, 3.76 (% of correctly classified children=66.0).

Descriptive statistics, Cronbach's α values and scale inter-correlations for temperament measures in Sample 2 are listed in Table 6. With the possible exception of Fearlessness scale

Table 5 Child Problematic Traits Inventory confirmatory factor analysis results: Robust weighted least square factor loadings in Sample 1 (*N*=381) and in Sample 2 (*N*=406)

CPTI items	Sample 1 (*N*=381)									Sample 2 (*N*=406)								
	Teacher' ratings			Mother' ratings			Father' ratings			Teacher' ratings			Mother' ratings			Father' ratings		
	GD	CU	INS	GD	CU	INS	GD	CU	INS	GD	CU	INS	GD	CU	INS	GD	CU	INS
1. Likes change and that things…			0.24			0.11			0.18			0.39			0.18			0.24
2. Seldom expresses sympathy…		0.67			0.51			0.48			0.67			0.57			0.58	
3. Often has difficulties with…		0.80			0.63			0.63			0.88			0.62			0.64	
4. Usually does not seem to share…		0.75			0.66			0.66			0.73			0.68			0.61	
5. Lies often to avoid problems.	0.84			0.76			0.68			0.92			0.77			0.77		
6. Seems to do certain things just…		0.88			0.79			0.73			0.91			0.69			0.67	
7. Seems to see himself/herself as…	0.71			0.70			0.67			0.77			0.72			0.72		
8. Never seems to have bad…		0.87			0.74			0.78			0.88			0.72			0.72	
9. Often lies to get what he/she…	0.92			0.82			0.83			0.95			0.82			0.82		
10. Provides himself/herself with…		0.83			0.72			0.78			0.93			0.75			0.82	
11. Often seems to be completely…		0.86			0.78			0.76			0.86			0.69			0.68	
12. Often does things without…		0.87			0.79			0.74			0.91			0.68			0.75	
13. Does not become upset when…		0.83			0.73			0.72			0.83			0.67			0.64	
14. Often consumes things…		0.88			0.67			0.66			0.82			0.65			0.67	
15. Seems to lie more than other…	0.95			0.71			0.79			0.95			0.76			0.79		
16. Seems to have a great need for…		0.79			0.68			0.67			0.79			0.71			0.74	
17. Seldom remorseful when…		0.92			0.79			0.78			0.92			0.82			0.81	
18. Is often superior and arrogant…	0.83			0.78			0.75			0.89			0.75			0.77		
19. Does not like waiting.		0.81			0.67			0.67			0.87			0.69			0.67	
20. Often does not seem to care…		0.84			0.67			0.70			0.75			0.80			0.77	
21. To get people to do what…	0.80			0.77			0.78			0.91			0.78			0.81		
22. Sometimes seems to…		0.94			0.83			0.88			0.90			0.80			0.88	
23. Seems to get bored quickly.		0.75			0.59			0.61			0.82			0.66			0.64	
24. Thinks that he/she is better…	0.63			0.70			0.71			0.75			0.74			0.75		
25. Never expresses feelings of…		0.93			0.79			0.82			0.91			0.76			0.75	
26. To frequently lie seems to be…	0.94			0.82			0.81			0.95			0.79			0.84		
27. Does not express guilt and…		0.88			0.77			0.77			0.92			0.79			0.82	
28. Quickly gets tired of things and…			0.72			0.78			0.69			0.82			0.65			0.75

CPTI Child Problematic Traits Inventory, *GD* CPTI Grandiose-Deceitful factor, *CU* CPTI Callous-Unemotional factor, *INS* CPTI Impulsive-Need for Stimulation factor; within each sample, bold highlights weighted least square mean and variance adjusted confirmatory factor analysis loadings which were significant at Bonferroni-corrected *p*-level (i.e., *p*<.0018)

based on father's ratings, all other temperament measures showed adequate internal consistency reliability estimates (i.e., Cronbach's α values>0.70). Considering Colins et al.' (2014) fearlessness, none of the teacher's ratings correlated with mother's ratings and father's ratings; rather, significant correlations were observed between mother's ratings and father's ratings. Significant correlations between teacher's ratings, and mother's ratings and father's ratings, respectively were observed for all child's temperament scales based on Rothbart et al.' (2000) model. For all sources of information that were considered in this study, measures of easy temperament and positive affect correlated significantly, albeit moderately with effortful control scales, with the partial exception of father's ratings of child's easy temperament scale. Mother's ratings and father's ratings on child's fear scale correlated significantly with both mother's ratings and father's ratings on child's anger/irritability scale. For all sources of information that were involved in this study, the child's easy temperament scale correlated negatively with teacher's ratings and mother's ratings of anger/irritability. Surprisingly, no significant correlations were observed between fearlessness ratings and fear ratings.

Paired-sample *t*-tests with Bonferroni-corrected nominal *p*-level (i.e., *p*<.0019) showed that average scores on child's

Table 6 Descriptive statistics, correlation coefficients (below diagonal) and Cronbach alpha coefficients (Main Diagonal) of child's temperament scales in Sample 2 (*N*=406)

	M	SD	Temperament scale Cronbach's α coefficients (main diagonal) and inter-correlations (below diagonal)																
			1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
1. F-Teacher	2.40	0.72	0.95																	
2. F-Mother	2.58	0.60	0.00	0.88																
3. F-Father	2.62	0.61	−0.01	0.62	0.80															
4. ET-Teacher	3.26	0.68	0.21	−0.04	−0.08	0.94														
5. ET-Mother	3.56	0.43	−0.02	0.23	0.15	0.21	0.84													
6. ET-Father	3.59	0.43	−0.03	0.14	0.20	0.10	0.66	0.67												
7. PA-Teacher	3.27	0.88	0.28	0.06	0.02	0.28	0.17	0.01	0.73											
8. PA-Mother	3.72	0.70	0.10	0.11	0.06	0.13	0.30	0.26	0.33	0.75										
9. PA-Father	3.72	0.73	0.04	0.06	0.07	0.10	0.30	0.33	0.26	0.75	0.76									
10. AI-Teacher	1.64	0.85	0.04	0.01	0.01	−0.38	−0.25	−0.28	0.12	−0.03	−0.05	0.94								
11. AI-Mother	1.74	0.77	0.00	0.02	0.01	−0.24	−0.24	−0.21	−0.01	−0.02	−0.03	0.46	0.90							
12. AI-Father	1.74	0.71	−0.03	0.07	0.06	−0.20	−0.17	−0.15	−0.05	−0.03	0.00	0.35	0.70	0.85						
13. Fe-Teacher	1.33	0.56	−0.07	0.08	0.07	−0.08	−0.02	−0.04	−0.02	−0.08	−0.02	0.12	0.07	0.04	0.95					
14. Fe-Mother	1.44	0.64	0.00	0.00	−0.02	0.02	−0.13	−0.15	0.01	0.02	0.03	0.14	0.33	0.25	0.19	0.91				
15. Fe-Father	1.43	0.58	0.00	0.05	−0.04	0.05	−0.10	−0.19	0.02	0.03	0.03	0.12	0.22	0.25	0.20	0.70	0.90			
16. EC-Teacher	3.27	0.95	0.17	0.00	−0.02	0.38	0.25	0.09	0.68	0.27	0.23	−0.06	−0.12	−0.10	−0.03	−0.07	−0.03	0.82		
17. EC-Mother	3.67	0.82	0.06	−0.01	0.00	0.24	0.28	0.18	0.29	0.47	0.40	−0.08	−0.15	−0.11	−0.05	0.00	0.00	0.41	0.73	
18. EC-Father	3.73	0.80	0.08	−0.05	−0.02	0.24	0.18	0.15	0.31	0.43	0.50	−0.10	−0.14	−0.15	−0.04	−0.03	0.03	0.42	0.73	0.76

F Fearlessness, *ET* Easy Temperament, *PA* Positive Affectivity, *AI* Anger/irritability, *Fe* Fear, *EC* Effortful control. The nominal significance level (i.e., *p*<.05) was corrected according to the Bonferroni procedure and set at *p*<.0003. Correlation coefficients greater than 0.18 in absolute value are significant at *p*<.0003. For ease of presentation, bold highlights significant *r* coefficients

temperament measures assessing positive affectivity (e.g., effortful control, positive affect) were significantly higher than temperament measures assessing negative affectivity (i.e., anger/irritability and fear) for all sources of information that were considered in this study, with Cohen d values ranging from 1.23 (attention vs. irritability, teacher's ratings) to 2.49 (positive affect vs. fear, father's ratings), all ps<0.0019. No significant differences were observed between positive affect measures and effortful control measures, independent from the source of information, min. d value (teacher's ratings)= 0.00, max. d value (mother's ratings)=0.06, all ps>0.20. This finding suggested that all sources of information described Sample 2 children as reasonably well-functioning children, at least on average. With the exception of teacher's ratings of child's anger/irritability, which were significantly higher in male children ($M=1.84$, $SD=0.99$) than in female children ($M=1.46$, $SD=0.64$), t (404)=4.52, $p<.001$, $d=0.45$, and teacher's ratings of easy temperament, which were significantly higher in female children ($M=3.34$, $SD=0.64$) than in male children ($M=3.18$, $SD=0.70$), t (404)=2.44, $p<.05$, $d=0.024$, no other significant effect of child's gender on temperament scale scores was observed.

Significant, albeit small correlations were observed between child's age and teacher's ratings of child's fearlessness, $r=0.15$, $p<.01$, mother's ratings of child's fearlessness, $r=0.12$, $p<.05$, father's ratings of child's fearlessness, $r=0.11$, $p<.05$, teacher's ratings of child's positive affect, $r=0.21$, $p<.001$, mother's ratings of child' anger/irritability, $r=0.14$, $p<.01$, and father's ratings of child's anger/irritability, $r=0.12$, $p<.05$. Parents' SES correlated significantly, albeit weakly with teacher's ratings of child's easy temperament, $r=0.19$, $p<.001$, mother's rating of child's anger/irritability, $r=-0.12$, $p<.05$, teacher's ratings of child's effortful control, $r=0.16$, $p<.01$, mother's ratings of child's effortful control, $r=0.12$, $p<.05$, and father's ratings of child's effortful control, $r= 0.13$, $p<.05$.

The results of regression analyses of CPTI total scores based on teacher's ratings, mother's ratings, and father's ratings, respectively, on the child's measures of fearlessness and easy temperament, as well as on the measures of child's positive affect, anger/irritability, fear, and effortful control are summarized in Table 7. With the aim to examine the role of the three CPTI scales in relation to child's temperament, a series of multiple regression analyses were conducted. Table 8 showed the results of multiple regression analysis in which measures of fearlessness and easy temperament, and measures of positive affect, anger/irritability, fear, and effortful control, respectively, were entered as predictors of CPTI scale scores based on teacher's ratings, mother's ratings, and father's ratings, respectively. In all regression models child's gender, child's age, and parents' SES were entered as covariate in step 1. None of the VIF values suggested problems with collinearity.

Discussion

As a whole, we conclude that our findings confirm and extend previous data on the CPTI (Colins et al. 2014). Results suggest that core problematic traits of psychopathy can be reliably and validly assessed also in school-age children living in a different cultural context than Sweden (namely, living in Italy).

Teachers', Mothers' and Fathers' CPTI Ratings

Although the CPTI was developed as a measure to be primarily rated by teachers, in both samples that were involved in our study internal consistency reliability estimates for mother's ratings and father's ratings of the CPTI were of comparable size as those that were observed for teacher's ratings. According to all sources of information that participated in this studies, the average ratings of the child's behavior on the CPTI scales were on average roughly 2 (i.e., *Does not apply well*) or lower, indicating that the school age children that our samples were composed on average by well-functioning children. Interestingly, parents' ratings were significantly higher than teacher's ratings in both samples, thus suggesting that mothers and fathers were not lenient in rating their children's behavior on the CPTI. The lack of a consistent association between CPTI scores and child's male gender could be explained by the absence in the CPTI of items assessing aggressive behavior or delinquency, and by the fact that in our study the CPTI was used to rate the behavior of generally well-adapted children coming from average SES families.

CPTI Rating Agreement Across Different Sources of Information

In both samples, ICC values suggested that mothers and fathers showed an acceptable level of agreement in reporting their children's behavior using the CPTI even when they had to independently rate their children's behavior. Rather, in our study poor agreement was observed between teacher's ratings of child's behavior on the CPTI and both mother's and father's ratings of child's behavior on the same scale; by the way, a poor-to-modest convergence between teacher's ratings and parents' ratings was observed also for all child's temperament measures that were administered to Sample 2 participants. This finding was consistent with a number of studies documenting inconsistent findings on the correspondence between parent ratings and observer ratings of child's temperament; moreover, when significant correlations were found they were generally in the weak to modest range (r's<0.30; Stifter et al. 2008).

Thus, our data suggest that in assessing child's psychopathy traits using the CPTI teachers and parents may show substantial disagreement, which is likely to be due to the

Table 7 Measures of child's temperament as predictor of the Child Problematic Trait Inventory total score in Sample 2 ($N=406$): Multiple regression analysis summary table

Independent variables	Teacher's CPTI total score		Mother's CPTI total score		Father's CPTI total score	
	β	VIF	β	VIF	β	VIF
Fearlessness-teacher's ratings	0.37***	1.09	0.04	1.09	0.02	1.07
Easy temperament-teacher's ratings	−0.39***	1.13	−0.28***	1.12	−0.27***	1.05
Change in R^2 (change in $R^2_{adjusted}$)	0.21***(0.20)		0.07***(0.06)		0.06***(0.06)	
Fearlessness-mother's ratings	0.12*	1.08	0.15**	1.08	0.13*	1.08
Easy temperament-mother's ratings	−0.28***	1.08	−0.19***	1.07	−0.17**	1.07
Change in R^2 (change in $R^2_{adjusted}$)	0.08***(0.07)		0.04***(0.04)		0.03**(0.03)	
Fearlessness-father's ratings	0.13*	1.06	0.07	1.07	0.13*	1.10
Easy temperament-father's ratings	−0.27***	1.06	−0.14*	1.06	−0.19***	1.17
Change in R^2 (change in $R^2_{adjusted}$)	0.07***(0.07)		0.02*(0.02)		0.04***(0.04)	
Positive affect-teacher's ratings	0.14**	1.83	0.00	1.91	−0.05	1.90
Anger/irritability teacher's ratings	0.64***	1.12	0.26***	1.13	0.25***	1.12
Fear-teacher's ratings	−0.08*	1.03	−0.10*	1.02	−0.07	1.03
Effortful control-teacher's ratings	−0.15**	1.84	−0.08	1.91	−0.08	1.94
Change in R^2 (change in $R^2_{adjusted}$)	0.42***(0.41)		0.08***(0.06)		0.07***(0.06)	
Positive affect-mother's ratings	0.04	1.29	0.06	1.30	0.01	1.28
Anger/irritability mother's ratings	0.33***	1.19	0.32***	1.18	0.24***	1.17
Fear-mother's ratings	0.00	1.13	0.03	1.12	0.01	1.12
Effortful control-mother's ratings	−0.05	1.34	−0.04	1.35	−0.11	1.34
Change in R^2 (change in $R^2_{adjusted}$)	0.11***(0.10)		0.12***(0.10)		0.07***(0.06)	
Positive affect-father's ratings	0.08	1.32	0.10	1.36	0.08	1.33
Anger/irritability father's ratings	0.19***	1.12	0.30***	1.12	0.41***	1.11
Fear-father's ratings	0.00	1.09	0.11*	1.08	0.05	1.08
Effortful control-father's ratings	−0.14*	1.38	−0.20***	1.44	−0.15*	1.40
Change in R^2 (change in $R^2_{adjusted}$)	0.06**(0.05)		0.16***(0.15)		0.22***(0.21)	

In all regression models, child's age, child's gender, and parents' SES were entered in step 1 as covariates; *VIF* Variance inflation factor. For ease of presentation, only significant standardized regression coefficients are displayed

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

reasons that Rothbart and Bates (1998) acutely pointed out as potential sources of measurement error: (1) rater characteristics that are relatively independent of child behavior; (2) bias as a function of child behavior or child–rater interaction; and (3) method factors. Several studies have documented systematic associations between parent personality, psychopathology, and stress with temperament ratings (Mebert 1991; Sameroff et al. 1982; Vaughn et al. 1981). Such factors may alter parent perceptions of their children's behavior, interfering with the ability to accurately identify and report on their actions and emotional responses. However, it should be observed that the context within which parents and teachers are rating the child's behavior may also explain their lack of agreement; for instance, parents view their children across a variety of situations, while teachers assess the child's behavior in the normative environment of the classroom (Stifter et al. 2008). Moreover, it is known that parents and observers converge, to a degree, on ratings of positive temperament but tend

to diverge on the negative temperament dimension (Stifter et al. 2008), and psychopathy may be hardly considered as a positive personality dimension.

In summary, our findings suggest that obtaining ratings on the CPTI by different sources of information may prove useful in assessing psychopathy traits in school-age children not at risk for delinquency because it may help to differentiate child's maladaptive behavior which is consistently present in a number of contexts from child's context-specific manifestations, although care should be used in combining information since teacher's ratings and parent's ratings on the CPTI do not show substantial agreement.

CPTI Factor Structure

Notwithstanding the poor absolute agreement between parents' ratings and teacher's ratings on the CPTI, the factor structure of the CPTI items was consistently replicated within

Table 8 Measures of child's temperament as predictors of the Child Problematic Trait Inventory scale scores in Sample 2 ($N=406$): Multiple regression analysis summary table

Independent variables	Teacher's ratings			Mother's ratings			Father's ratings			VIF
	CPTI GD β	CPTI CU β	CPTI INS β	CPTI GD β	CPTI CU β	CPTI INS β	CPTI GD β	CPTI CU β	CPTI INS β	
Fearlessness-teacher's ratings	0.31***	0.33***	0.35***	0.05	0.06	0.13*	0.05	0.04	0.07	1.11
Easy temperament-teacher's ratings	−0.42***	−0.43***	−0.24***	−0.25***	−0.18***	−0.28***	−0.27***	−0.14*	−0.28***	1.17
Change in R^2 (change in $R^2_{adjusted}$)	0.20*** (0.19)	0.21*** (0.20)	0.13*** (0.12)	0.05*** (0.05)	0.04*** (0.03)	0.07*** (0.06)	0.06*** (0.06)	0.02* (0.01)	0.07*** (0.06)	
Fearlessness-mother's ratings	0.09	0.11*	0.11*	0.10	0.16**	0.10	0.10	0.12*	0.10	1.08
Easy temperament-mother's ratings	−0.22***	−0.29***	−0.24***	−0.20***	−0.15**	−0.14**	−0.19***	−0.12*	−0.13*	1.08
Change in R^2 (change in $R^2_{adjusted}$)	0.05*** (0.04)	0.08*** (0.07)	0.06*** (0.05)	0.04*** (0.03)	0.04*** (0.03)	0.02* (0.02)	0.04* (0.03)	0.02* (0.02)	0.02* (0.02)	
Fearlessness-father's ratings	0.13*	0.15**	0.08	0.06	0.10	0.02	0.15**	0.12*	0.07	1.07
Easy temperament-father's ratings	−0.25***	−0.24***	−0.22***	−0.16**	−0.10	−0.12	−0.22***	−0.16**	−0.12*	1.07
Change in R^2 (change in $R^2_{adjusted}$)	0.07*** (0.06)	0.07*** (0.06)	0.05*** (0.04)	0.02* (0.02)	0.02* (0.02)	0.01(0.01)	0.06*** (0.05)	0.03*** (0.02)	0.02* (0.01)	
Positive affect-teacher's ratings	0.06	0.02	0.27	0.06	0.03	0.02	−0.04	−0.02	−0.06	1.91
Anger/irritability teacher's ratings	0.66***	0.57***	0.53***	0.27***	0.14***	0.26***	0.31***	0.10	0.26***	1.13
Fear-teacher's ratings	−0.08*	−0.07	−0.07	−0.11*	−0.04	−0.12*	−0.10	−0.03	−0.05	1.03
Effortful control-teacher's ratings	−0.09*	−0.13*	−0.18***	−0.02	−0.10	−0.06	−0.10	−0.06	−0.06	1.94
Change in R^2 (change in $R^2_{adjusted}$)	0.42*** (0.41)	0.32*** (0.31)	0.33*** (0.32)	0.07*** (0.06)	0.03*** (0.03)	0.08*** (0.07)	0.11*** (0.10)	0.02* (0.01)	0.07*** (0.06)	
Positive affect-mother's ratings	0.01	0.03	0.08	0.01	−0.01	−0.06	−0.02	−0.04	0.06	1.27
Anger/irritability mother's ratings	0.31***	0.27***	0.31***	0.32***	0.16**	0.35***	0.30***	0.05	0.30***	1.19
Fear-mother's ratings	−0.03	0.03	−0.01	0.04	0.10	0.01	−0.03	0.09	−0.06	1.14
Effortful control-mother's ratings	0.02	−0.06	−0.08	0.03	−0.05	−0.03	−0.05	−0.05	−0.12	1.33
Change in R^2 (change in $R^2_{adjusted}$)	0.09*** (0.08)	0.08*** (0.07)	0.10*** (0.09)	0.11*** (0.10)	0.05*** (0.04)	0.13*** (0.12)	0.09** (0.08)	0.02* (0.03)	0.10*** (0.09)	
Positive affect-father's ratings	0.06	0.04	0.11	0.07	0.00	0.17**	0.02	0.01	0.16**	1.35
Anger/irritability father's ratings	0.19***	0.12*	0.20***	0.36***	0.12*	0.35***	0.44***	0.20***	0.50***	1.12
Fear-father's ratings	0.00	0.03	−0.02	0.05	0.15**	0.04	0.05	0.11	−0.04	1.09
Effortful control-father's ratings	−0.09	−0.15*	−0.13	−0.13*	−0.13	−0.22***	−0.08	−0.08	−0.18**	1.41
Change in R^2 (change in $R^2_{adjusted}$)	0.05* (0.05)	0.04** (0.03)	0.06** (0.05)	0.16*** (0.15)	0.07*** (0.05)	0.19*** (0.18)	0.21*** (0.21)	0.07*** (0.06)	0.26*** (0.25)	

In all regression models, child's age, child's gender, and parents' SES were entered in step 1 as covariates; *VIF* Variance inflation factor. For ease of presentation, only maximum VIF values for each predictor are displayed

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

each sample across teacher's ratings, mother's ratings, and father's ratings respectively. For all sources of information that were considered in our study, in both samples quasi-inferential parallel analyses supported a three-factor structure of the CPTI items. Confirming and extending Colins et al.' (2014) findings, our WLSMV CFA results indicated that the a priori model of CPTI item assignment to scales showed at least marginally acceptable fit indices and showed the best fit among the competing models that were tested in this study.

However, in our study significant error correlations were observed between three couples of items; interestingly, this pattern or error correlation was consistently across our two samples and our three sources of information. Akin to the concept of local dependence in item response theory (Thissen and Steinberg 2010), error correlation indicates that items are more strongly correlated than can be accounted for by the general factor; in other terms, error correlation hints at item multidimensionality (i.e., item responses are influenced by more than a single latent dimension). Although error correlation is likely to arise by common wording or other method factors, it may be also the results of theoretically relevant latent dimensions which could not be held constant in the study. For instance, error correlations between item 4 ("Usually does not seem to share others' joy and sorrow") and item 2 ("Seldom expresses sympathy for others"), and between item 11 ("Often seems to be completely indifferent when other children are upset") and item 13 ("Does not become upset when others are being hurt"), respectively, may be the result of the fact that these items are influenced by child's CU which is directly assessed by the CPTI as well as by child's low E (i.e., a general personality trait whose low scores are associated with emotional coldness) which is not assessed by the CPTI and was not controlled for in this study. Similar considerations hold also for item 7 ("Seems to see himself/herself as superior compared to others") and item 24 ("Thinks that he/she is better than everyone at almost everything"), which may reflect child's GD as well as child's low A (i.e., Antagonism) or child's defiant attitude towards adults.

When these specific error correlations were considered in the three-factor model based on the priori model of CPTI item assignment to scales, all fit indices suggested at least adequate fit for this model. Although these data were somewhat encouraging, they also seemed to suggest the necessity of further studies using full-information Item Response Theory in order to further investigate the CPTI item latent structure and to address the issue of local dependence of selected CPTI items.

With the partial exception of CPTI item 1 ("Likes change and that things happen all the time"), all other CPTI items showed significant and substantial loadings on their corresponding factor even in the Italian translation of the CPTI. The moderate size of the samples and the presence of systematic differences between our two samples (e.g., significant between-group differences in parents' SES and occupation)

prevented from formally testing for nested factor invariance hypotheses. Thus, we were forced to test CFA models separately in the two study groups in order to see if the a priori three-factor model of the CFA items could be safely replicated in both samples. Descriptively, the CFA factor loading matrices that were obtained in our two study groups looked pretty similar, as it was suggested by congruence coefficients values greater than 0.995 for all between-group comparison based on the same source of information (e.g., GD factor loading based on teacher ratings in Sample 1 vs. GD factor loading based on teacher ratings in Sample 2). However, future studies should address the issue of CPTI item factor invariance across independent samples, as well as across groups based on participants' gender.

Thus, our findings landed some support to the hypothesis that the three-factor structure of the CPTI items could be safely reproduced across different sources of information and across different samples. Of course, saying that the hypothesis of the three-factor structure of CPTI items has found some support in our study does not imply that our data support the three-factor structure of psychopathy per se. For instance, we were not able to test a four-factor model of psychopathy simply because the CPTI does not list any item assessing behavioral deviance; thus, we cannot rule out the hypothesis that if we had relied on a different measure of psychopathy, a different latent structure may have been found.

CPTI Nomological Network Validity

In our study, teacher's ratings of the CPTI total score were significant, relevant predictors of child's problem behavior at school, at least as it was indexed by official behavior grade, even when the effects parents' SES, and child's age and gender were controlled for in logistic regression analyses. Interestingly, mother's ratings of the CPTI total scores also significantly predicted the child's problem behavior at school, whereas the CPTI total score based on father's ratings did not significantly discriminated children with problem behavior at school from children with appropriate school behavior. This finding suggest that mother's ratings and father's ratings may have different external validity with theoretically-relevant external constructs (e.g., poor behavior at school), although mothers and fathers showed adequate agreement indices concerning the CPTI total score; thus, their ratings should be considered separately, rather than pooled to compose a general parent CPTI score of child's behavior.

Multiple regression analysis results showed that child's temperament measures significantly predicted the CPTI scales and total scores across all sources of information that were considered in this study. Consistent with previous findings (Colins et al. 2014), child's difficult temperament (i.e., low rating on the easy temperament scale) predicted the CPTI total score across all sources of information which participated in

our study. Less consistent findings were observed for the fearlessness dimension; indeed, mother's ratings of the fearlessness scale significantly predicted the CPTI total score across all sources of information. However, teacher's ratings of child's fearless behavior predicted the teacher's ratings of the CPTI total score, but they did not significantly predict neither mother's ratings nor father's ratings of the CPTI total score; father's ratings of child's fearlessness did not significantly predict mother's ratings of the CPTI total score. Notwithstanding the partial inconsistency of the findings concerning the association between measures of child's fearlessness and child's psychopathy, our findings were largely supported the nomological network validity of the CPTI even when different sources of information are considered.

According to our regression analyses results, anger/irritability consistently represented a significant predictor of child's psychopathy personality traits, at least as operationalized by the CPTI total score, across all sources of information that were considered in this study (i.e., teacher's ratings, mother's ratings, and father's ratings of the CPTI, respectively). Father's ratings of child's effortful control were significantly and negatively associated with all ratings of the CPTI total score; at the opposite, mother's ratings of child's effortful control did not significantly predicted the CPTI total score. This finding gave further support to the hypothesis that mother's ratings of the CPTI and father's ratings of the CPTI should perhaps not be pooled to compose a general parent score. Marginally, teacher's ratings of child's poor effortful control were significantly associated only with teacher's ratings of the CPTI total score. Similar considerations held also for positive affect scores which showed the expected significant association with the CPTI total score, but only when teacher's ratings were considered. Finally, only teacher's ratings of child's fearful temperament, at least as operationalized by the Fear scale, showed a significant, albeit weak relationship in the negative direction with the CPTI total score based on teacher's ratings and mother's ratings, respectively.

As a whole, our multiple regression analysis results gave at least partial support to the nomological network validity of the CPTI in school-age children, and highlighted the importance of examining the three dimensions of psychopathy (Salekin 2015). In our opinion, the weak associations between Fear scale scores and CPTI rating, as well as the relatively inconsistent findings concerning the relationship between ratings of the CPTI total score and ratings of child's fearlessness, were likely to reflect the poor convergent validity of the measures that were used in this study, rather than suggesting problems with the nomological network validity of the CPTI.

Finally, our temperament-related regression analyses findings are particularly notable in reference to DeLisi and Vaughn's (2014) temperamental theory. Indeed, DeLisi and Vaughn (2014) recently reviewed over 300 studies to introduce a temperament-based theory of antisocial conduct with criminal justice system implications. Two temperamental constructs, i.e., effortful control and negative emotionality, were identified as significantly predictive of self-regulation deficits and behavioral problems in infancy, in toddlerhood, in childhood, in adolescence, and across adulthood. DeLisi and Vaughn's (2014) study found that it is not until a poorly controlled, negative child enters school (i.e., a normative environment) that his/her temperamental deficits come to light. Children who are aggressive, demanding, and angry, rather quickly irritated their classmates; teachers also quickly identify children who are the most behaviorally challenging to school order (DeLisi and Vaughn 2014). Our findings seemed to suggest that the CPTI may represent a particularly useful instrument in these situations. Indeed, the structure of the classroom environment is often difficult for children with self-regulation problems to handle (DeLisi and Vaughn 2014), and in these situations the CPTI may represent a useful instrument to assess psychopathic traits in children.

Limitations

We feel that our findings should be considered in the light of a number of limitations. None of the samples was composed of randomly selected participants; rather, they were all volunteers. Thus, we relied on convenience study groups rather than on samples actually representative of the Italian population of 6–12 years old children. This inherently limits the generalizability of our findings. The children who were rated in this study using the CPTI were on average well-adapted children, living in average SES families; this limits the possibility to generalize our findings on the CPTI to children at high risk for behavioral deviance and delinquency. We were not able to evaluate inter-rater reliability of teacher's ratings, since we relied only on CPTI ratings of the teacher who spent the highest amount of time with the classroom. Although intraclass correlation coefficient values for parents' ratings were adequate, we cannot exclude that poor inter-rater reliability of teacher's ratings on the CPTI may have biased the correlation between the CPTI and external criterion measures. All child's temperament measures that were used in this study were ad hoc measures, and none of them received extensive validity studies. Moreover, fear/fearlessness measures showed severe problems with convergent validity. All these consideration strongly recommend further studies before accepting our regression analysis results.

Conclusion

This study to a large extent suggest that the CPTI represent a reliable and valid measure of children's problematic traits which can be used in assessing psychopathy traits in school-age children, particularly when different sources of information are used to rate the child's behavior.

References

Ahadi, S. A., Rothbart, M. K., & Ye, R. M. (1993). Child temperament in the U.S. and China: similarities and differences. *European Journal of Personality, 7*, 359–378.

Alsawalmeh, Y. M., & Feldt, L. S. (1992). Test of the hypothesis that the intraclass reliability coefficient is the same for two measurement procedures. *Applied Psychological Measurement, 16*(2), 195–205.

Andershed, H. (2010). Stability and change of psychopathic traits. What do we know? In R. T. Salekin & D. R. Lynam (Eds.), *Handbook of child and adolescent psychopathy* (pp. 233–250). New York: The Guilford Press.

Andershed, Hodgins, S., & Tengstrom, A. (2007). Convergent validity of the Youth Psychopathic traits Inventory (YPI)-association with the Psychopathy Checklist: Youth Version (PCL: YV). *Assessment, 14*(2), 144–154.

Andershed, H., Kerr, M., Stattin, H., & Levander, S. (2002). Psychopathic traits in non-referred youths: A new assessment tool. In E. Blau & L. Sheridan (Eds.), *Psychopaths: Current international perspectives* (pp. 131–158). Amsterdam: Elsevier.

Benning, S. D., Patrick, C. J., Hicks, B. M., Blonigen, D. M., & Krueger, R. F. (2003). Factor structure of the psychopathic personality inventory: validity and implications for clinical assessment. *Psychological Assessment, 15*, 340–350.

Bentler, P. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park: Sage.

Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research, 27*, 509–540.

Caldwell, M. F., Ziemke, M. H., & Vitacco, M. J. (2008). An examination of the Sex offender registration and notification act as applied to juveniles: evaluating the ability to predict sexual recidivism. *Psychology, Public Policy, and Law, 14*, 89–114.

Cha, E. S., Kim, K. H., & Erlen, J. A. (2007). Translation of scales in cross-cultural research: issues and techniques. *Journal of Advanced Nursing, 58*(4), 386–395.

Cleckley, H. (1941). *The mask of sanity* (1st ed.). St. Louis: Mosby.

Colins, O., Noom, M., & Vanderplasschen, W. (2012). Youth psychopathic traits inventory - short version: a further test of the internal consistency and criterion validity. *Journal of Psychopathology and Behavioral Assessment, 34*(4), 476–486.

Colins, O. F., Andershed, H., Frogner, L., Lopez-Romero, L., Veen, V., & Andershed, A. K. (2014). A new measure to assess psychopathic personality in children: the child problematic traits inventory. *Journal of Psychopathology and Behavioral Assessment, 36*(1), 4–21.

Cooke, D. J., & Michie, C. (1997). An item response theory analysis of the hare psychopathy checklist-revised. *Psychological Assessment, 9*, 3–14.

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: towards a hierarchical model. *Psychological Assessment, 13*(2), 171–188.

Dadds, M. R., Fraser, J., Frost, A., & Hawes, D. J. (2005). Disentangling the underlying dimensions of psychopathy and conduct problems in childhood: a community study. *Journal of Consulting and Clinical Psychology, 73*, 400–410.

DeLisi, M., & Vaughn, M. G. (2014). Foundation for a temperament-based theory of antisocial behavior and criminal justice system involvement. *Journal of Criminal Justice, 42*, 10–25.

Denissen, J. J., Geenen, R., van Aken, M. A., Gosling, S. D., & Potter, J. (2008). Development and validation of a Dutch translation of the Big Five Inventory (BFI). *Journal of Personality Assessment, 90*, 152–157.

Derryberry, D., & Rothbart, M. K. (1988). Arousal, affect, and attention as components of temperament. *Journal of Personality and Social Psychology, 55*, 958–966.

Donner, A., & Zou, G. Y. (2002). Testing the equality of dependent intraclass correlation coefficients. *The Statistician, 51*, 367–379.

Essau, C. A., Sasagawa, S., & Frick, P. J. (2006). Callous-unemotional traits in a community sample of adolescents. *Assessment, 13*, 454–469.

Evans, D. E., & Rothbart, M. K. (1999). *Relationships between self-report temperament measures and personality.* Unpublished manuscript, University of Oregon.

Ezpeleta, L., de La Osa, N., Granero, R., Penolo, E., & Domenech, J. M. (2013). Inventory of callous-unemotional traits in a community sample of preschoolers. *Journal of Clinical Child & Adolescent Psychology, 22*(1), 91–105.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods, 9*, 466–491.

Forth, A. E., Kosson, D. S., & Hare, R. D. (2003). *The psychopathy checklist: Youth version*. Toronto: Multi-Health Systems.

Frick, P. J., & Hare, R. D. (2001). *The antisocial process screening device*. Toronto: Multi-Health Systems.

Frick, P. J., Bodin, S. D., & Barry, C. T. (2000). Psychopathic traits and conduct problems in community and clinic-referred samples of children: further development of the psychopathy screening device. *Psychological Assessment, 12*, 382–393.

Geisinger, K. F. (1994). Cross-cultural normative assessment: translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment, 6*, 304–312.

Goldberg, L. R. (2001). Analyses of Digman's child-personality traits: derivation of big-five factor scores from each of six samples. *Journal of Research in Personality, 69*, 709–743.

Gretton, H. M., Hare, R. D., & Catchpole, R. E. (2004). Psychopathy and offending from adolescence to adulthood: a ten year follow-up. *Journal of Clinical and Consulting Psychology, 72*, 636–645.

Gustavson, C., Stahlberg, O., Sjodin, A., Forsman, A., Nilsson, T., & Anckarsater, H. (2007). Age at onset of substance abuse: a crucial covariate of psychopathic traits and aggression in adult offenders. *Psychiatry Research, 153*, 195–198.

Hampson, S. E., & Goldberg, L. R. (2006). A first large cohort study of personality trait stability over the 40 years between elementary school and midlife. *Journal of Personality and Social Psychology, 91*, 763–779.

Hare, R. D. (1991). *Manual for the revised psychopathy checklist* (1st ed.). Toronto: Multi-Health Systems.

Hare, R. D. (2003). *The hare psychopathy checklist-revised* (2nd ed.). Toronto: Multi-Health Systems.

Hare, R. D., & Neumann, C. S. (2006). The PCL-R assessment of psychopathy: Development, structural properties, and new directions. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 58–88). New York: Guilford Press.

Hare, R. D., Harpur, T. J., Hakstian, A. R., Forth, A. E., Hart, S. D., & Newman, J. P. (1990). The revised psychopathy checklist: reliability and factor structure. *Psychological Assessment, 2*, 338–341.

Harpur, T. J., Hakstian, A. R., & Hare, R. D. (1988). Factor structure of the psychopathy checklist. *Journal of Consulting and Clinical Psychology, 56*, 741–747.

Hollingshead, A. B. (2011). Four factor index of social status. *Yale Journal of Sociology, 8*, 21–51.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.

Karpman, B. (1941). On the need for separating psychopathy into two distinct clinical types: symptomatic and idiopathic. *Journal of Criminology and Psychopathology, 3*, 112–137.

Kennealy, P. J., Hicks, B. M., & Patrick, C. J. (2007). Validity of the factors of the psychopathy checklist- revised in female prisoners. *Assessment, 14*, 323–340.

Kimonis, E. R., Frick, P. J., Cauffman, E., Goldweber, A., & Skeem, J. L. (2012). Primary and secondary variants of juvenile psychopathy differ in emotional processing. *Development and Psychopathology, 24*, 1091–1103.

Kochanska, G. (1991). Socialization and temperament in the development of guilt and conscience. *Child Development, 62*, 1379–1392.

Kochanska, G. (1993). Toward a synthesis of parental socialization and child temperament in early development of conscience. *Child Development, 64*, 325–347.

Kochanska, G. (1997). Multiple pathways to conscience for children with different temperaments: From toddlerhood to age 5. *Developmental Psychology, 3*, 228–240.

Kotler, J. S., & McMahon, R. J. (2010). Assessment of child and adolescent psychopathy. In R. T. Salekin & D. R. Lynam (Eds.), *Handbook of child and adolescent psychopathy*. New York: The Guilford Press.

Kutner, M. H., Nachtsheim, C. J., & Neter, J. (2004). *Applied linear regression models* (4th ed.). New York: McGraw-Hill- Irwin.

Lorenzo-Seva, U., & Ferrando, P. J. (2006). FACTOR: a computer program to fit the exploratory factor analysis model. *Behavioral Research Methods, Instruments, & Computers, 38*(1), 88–91.

Lykken, D. T. (1957). A study of anxiety in the sociopathic personality. *Journal of Abnormal and Clinical Psychology, 55*, 6–10.

Lykken, D. T. (1995). *The antisocial personalities*. Hillsdale: Erlbaum.

Lynam, D. R. (1997). Pursuing the psychopath: capturing the fledgling psychopath in a nomological net. *Journal of Abnormal Psychology, 106*, 425–438.

Lynam, D. R. (2010). Child and adolescent psychopathy and personality. In R. T. Salekin & D. R. Lynam (Eds.), *Handbook of child and adolescent psychopathy* (pp. 179–201). New York: Guilford Press.

Marcus, D. K., John, S. L., & Edens, J. F. (2004). A taxometric analysis of psychopathic personality. *Journal of Abnormal Psychology, 113*(4), 626–635.

Marcus, D. K., Fulton, J. J., & Edens, J. F. (2013). The two-factor model of psychopathic personality: evidence from the psychopathic personality personality inventory. *Personality Disorders Theory, Research, and Treatment, 4*(1), 67–76.

Marsh, H. W., Hau, K.-. T., & Wen, Z. (2004). In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*(3), 320–341.

McCrae, R. R., Zonderman, A. B., Costa, P. T., Jr., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the revised NEO personality inventory: confirmatory factor analysis versus procrustes rotation. *Journal of Personality and Social Psychology, 70*, 552–566.

Mebert, C. J. (1991). Dimensions of subjectivity in parents' ratings of infant temperament. *Child Development, 62*(2), 352–361.

Murrie, D. C., Boccaccini, M. T., McCoy, W., & Cornell, D. G. (2007). Diagnostic labelling in juvenile court: how do psychopathy and conduct disorder findings infuence judges? *Journal of Clinical Child & Adolescent Psychology, 36*, 228–241.

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus User's Guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén.

Muthen, L. K., & Muthén, B. O. (2014). *Mplus 7.3 [computer software]*. Los Angeles: Muthén & Muthén.

Nunnaly, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Patrick, C. J., Fowles, D. C., & Krueger, R. F. (2009). Triarchic conceptualization of psychopathy: developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology, 21*, 913–938.

Porter, S., Birt, A. R., & Boer, D. P. (2001). Investigation of the criminal and conditional release histories of Canadian federal offenders as a function of psychopathy and age. *Law and Human Behavior, 25*, 647–661.

R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Ribeiro da Silva, D., Rijo, D., & Salekin, R. T. (2012). Child and adolescent psychopathy: a state-of-the art reflection on the construct and etiological theories. *Journal of Criminal Justice, 40*, 269–277.

Roose, A., Bijttebier, P., Claes, L., Lilienfeld, S. O., De Fruyt, F., & Decuyper, M. (2012). Psychopathic traits in adolescence and the five factor model of personality. *Journal of Psychopathology and Behavioral Assessment, 34*(1), 84–93.

Rothbart, M. K., & Bates, J. E. (1998). Temperament. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Vol 3. Social, emotional and personality development* (5th ed., pp. 105–176). New York: Wiley.

Rothbart, M. K., & Mauro, J. A. (1990). Questionnaire approaches to the study of infant temperament. In J. W. Fagen & J. Colombo (Eds.), *Individual differences in infancy: Reliability, stability and prediction* (pp. 411–429). Hillsdale: Erlbaum.

Rothbart, M. K., Ahadi, S. A., & Evans, D. E. (2000). Temperament and personality: origins and outcomes. *Journal of Personality and Social Psychology, 78*(1), 122–135.

Salekin, R. T. (2010). Treatment of child and adolescent psychopathy: Focusing on change. In R. T. Salekin & D. R. Lynam (Eds.), *Handbook of child and adolescent psychopathy* (pp. 343–373). New York: Guilford Press.

Salekin, R. T. (2015). Psychopathy in Childhood: Toward Better Informing the DSM–5 and ICD-11 Conduct Disorder Specifiers. *Personality Disorders: Theory, Research, and Treatment*. doi:10.1037/per0000150

Salekin, R. T., & Lynam, D. R. (2010). Child and adolescent psychopathy: An introduction. In R. T. Salekin & D. R. Lynam (Eds.), *Handbook of child and adolescent psychopathy* (pp. 1–12). New York: Guilford Press.

Salekin, R. T., Worley, C., & Grimes, R. D. (2010). Treatment of psychopathy: a review and brief introduction to the mental model approach for psychopathy. *Behavioral Sciences & the Law, 28*(2), 235–266.

Sameroff, A. J., Seifer, R., Elias, P. K. (1982). Sociocultural variability in infant temperament ratings. *Child Development*, 164–173.

Schweitzer, K. (2012). On correlated errors. *European Journal of Psychological Assessment, 28*, 1–2.

Silver, N. C., Hittner, J. B., & May, K. (2006). A FORTRAN 77 program for comparing dependent correlations. *Applied Psychological Measurement, 30*, 152–153.

Stifter, C. A., Willoughby, M. T., & Towe-Goodman, N. (2008). Agree or agree to disagree? Assessing the convergence between parents and observers on infant temperament. *Infant and Child Development, 17*(4), 407–426.

Tanaka, J. S. (1993). Multifaceted concepts of fit in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 10–39). Newbury Park: Sage.

Tennent, G., Tennent, D., Prins, H., & Bedford, A. (1993). Is psychopathic disorder a treatable condition? *Medicine, Science and the Law, 33*(1), 63–66.

Thissen, D., & Steinberg, L. (2010). Using item response theory to disentangle constructs at different levels of generality. In S. Embretson (Ed.), *Measuring psychological constructs: Advances in model-based approaches* (pp. 123–144). Washington, DC: American Psychological Association.

Treiman, D. J., & Terrell, K. (1975). Sex and the process of status attainment: A comparison of working women and men. *American Sociological Review*, 174–200.

van Baardewijk, Y., Stegge, H., Andershed, H., Thomaes, S., Scholte, E., & Vermeiren, R. (2008). Measuring psychopathic traits in children through self-report. The development of the Youth Psychopathic traits Inventory-Child version. *International Journal of Law and Psychiatry, 31*(3), 199–209.

van Baardewijk, Y., Andershed, H., Stegge, H., Nilsson, K. W., Scholte, E., & Vermeiren, R. (2010). Development and tests of short versions of the Youth Psychopathic traits Inventory and the Youth Psychopathic traits Inventory-Child version. *European Journal of Psychological Assessment, 26*(2), 122–128.

van Baardewijk, Y., Vermeiren, R., Stegge, H., & Doreleijers, T. (2011). Self-reported psychopathic traits in children: their stability and concurrent and prospective association with conduct problems and aggression. *Journal of Psychopathology and Behavioral Assessment, 33*(2), 236–245.

Van de Vijver, F., & Hambleton, R. K. (1996). Translating tests: some practical guidelines. *European Psychologist, 1*(2), 89.

Vaughn, B. E., Taraldson, B. J., Crichton, L., & Egeland, B. (1981). The assessment of infant temperament: a critique of the Carey infant temperament questionnaire. *Infant Behavior and Development, 4*, 1–17.

Waller, R., Gardner, F., Hyde, L. W., Shaw, D. S., Dishion, T. J., & Wilson, M. N. (2012). Do harsh and positive parenting predict parent reports of deceitful-callous behavior in early childhood? *Journal of Child Psychology and Psychiatry, 53*(9), 946–953.

Walters, G. D., Knight, R. A., Grann, M., & Dahle, K. (2008). Incremental validity of the psychopathy checklist facet scores: predicting release outcome in six samples. *Journal of Abnormal Psychology, 117*, 396–405.

Zwick, W. R., & Velicer, W. F. (1986). Factors influencing five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432–442.