Published in final edited form as:

J Psychopathol Behav Assess. 2019 March ; 41(1): 144–159. doi:10.1007/s10862-018-9703-5.

Reliability and validity of the Thinking Skills Inventory, a screening tool for cross-diagnostic skill deficits underlying youth behavioral challenges

Lu Wang^{#1,2}, Alisha R. Pollastri^{#1,2}, Pieter J. Vuijk^{1,2,3,4}, Erin N. Hill^{1,2}, Brenda A. Lee^{1,2,3,4}, Anna Samkavitz^{1,2,3,4}, Ellen B. Braaten^{1,2}, J. Stuart Ablon^{1,2}, and Alysa E. Doyle^{1,2,3,4}

¹Department of Psychiatry, Massachusetts General Hospital, Boston, MA.

²Harvard Medical School, Boston, MA.

³Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA.

⁴The Stanley Center for Psychiatric Research, Broad Institute, Cambridge, MA.

These authors contributed equally to this work.

Abstract

Deficits in a range of skill domains (including executive functioning, emotion regulation, social cognition and language/communication) are associated with disrupted youth behavior and functioning across mental health diagnoses. The identification of skill deficits are important for effective treatment planning, particularly for personalized interventions. While there are multiple ways to assess these skills, parent/caregiver reports represent an important information source. To date, no single, brief measure has been developed that gathers parent/caregiver ratings across this range of constructs. We have developed a short caregiver-report questionnaire (the Thinking Skills Inventory; TSI), to screen for skill deficits. Here, we examine the reliability and validity of this rating scale in 384 youth who were consecutively referred for neuropsychiatric evaluation. A primary caregiver completed the TSI as well as other established measures. Exploratory and confirmatory factor analyses support five subscales on the TSI: Attention and Working Memory, Language and Communication, Emotion Regulation, Cognitive Flexibility, and Social Thinking Skills. The subscales showed moderate to high internal consistency (Cronbach's alphas range from 0.84 to 0.91). Correlations with established caregiver-report measures confirm their convergent and discriminant validity, and associations with multiple clinical diagnoses and cross-diagnostic aggressive behavior further support the utility of the scale for our intended purpose. In sum, this free, brief measure is a valid and reliable way to identify variation in skill domains relevant to a range of psychopathology. The TSI may be useful in youth mental health settings to assist with treatment planning and to inform referral for further evaluation.

Keywords

cross-diagnostic traits; child psychopathology; assessment; Collaborative Problem Solving

The most common reasons youth are referred for mental health evaluations or services include some form of challenging behavior exhibited at home and/or at school (Rushton, Bruckman, & Kelleher, 2002). Challenging behavior, which might include tantrums, withdrawal, verbal or physical aggression, or rule-breaking, can occur in the context of a range of clinical diagnoses, including depression, mania, psychosis, attention-deficit/hyperactivity disorders (ADHD), or autism spectrum disorders (ASD). Conventional treatment for behavioral symptoms like these tend to focus on reinforcing appropriate behavior primarily through operant strategies that target motivation as a mechanism of behavior change. However, recent neuropsychiatric research has documented that youth who exhibit challenging behaviors have impairments in an array of skill domains, such as language and communication (e.g. Zadeh, Im-Bolter, & Cohen, 2007), emotion and self-regulation (e.g. Hill, Degnan, Calkins, & Keane, 2006; Stieben et al. 2007), attention and working memory (e.g. Doyle et al. 2017), cognitive flexibility (e.g. Van Goozen et al. 2004) and social thinking (e.g. Dodge et al. 2003; Matthys, Cuperus, & Engeland, 1999). Indeed, these cross-diagnostic dimensions overlap with several of the constructs featured in the NIMH's Research Domain Criteria (RDoC), which aims to redirect focus in psychopathology research away from traditional, dichotomous diagnostic categories towards traits that can be more directly linked to underlying neural systems (Morris & Cuthbert, 2012).

Although conventional treatments for externalizing behaviors tend to be based on operant conditioning, newer personalized therapeutic approaches have been developed that recognize the benefit of targeting the skill deficits that contribute to children's behavioral struggles. One example is the Collaborative Problem Solving approach (CPS; Greene, 1998; Greene & Ablon, 2005; Pollastri, Epstein, Heath, & Ablon, 2013), which proposes that children's challenging behaviors are a result of a mismatch between environmental demands and the child's ability to access or implement the thinking skills required to meet those demands. According to these approaches, the way to reduce challenging behavior is to support the child in building the lagging skills that will be necessary for meeting those demands in the future. A valid assessment of skills in domains that potentially underlie cross-diagnostic challenging behaviors, and that can be targeted therapeutically, is important for this type of personalized intervention. Such information can assist with the development of a treatment plan, and may also serve as a basis for recommending further, formal assessment in specific areas.

Caregiver ratings on inventories assessing youth's functioning, symptoms, and skills are a cost-efficient means of gathering information. Standard caregiver-report measurements have been developed and tested for each of the skill domains mentioned above (e.g. the Behavior Rating Inventory of Executive Functions; BRIEF-Second Edition, Gioia, Isquith, Guy, & Kenworthy, 2015). However, no short, freely-available, single, validated measure exists to survey the range of domains such as executive functioning, language, emotion regulation, and social thinking skills, that can be under-developed in youth with emotional and behavioral disorders and thus therapeutic targets. Currently, an array of different instruments must be used to assess these areas, often requiring considerable monetary cost and time for completion and scoring. Thus, a brief parent-rated measure to screen for domains of weakness could be useful in treatment planning.

To address this gap, licensed clinical psychologists at [MASK 1] have developed the Thinking Skills Inventory (TSI). The TSI was developed to assess the aforementioned skill domains and has been revised over time in its function as a clinical tool. The goals of the current study are to introduce the revised TSI and evaluate its psychometric properties. Specifically, we aimed to: (1) confirm the items and latent structure of the TSI and assess its internal consistency; (2) assess the convergent and discriminant validity of the TSI in relation to other established measures; and (3) assess the criterion-related validity of the TSI by examining whether skill deficits are, as hypothesized, relevant to multiple diagnoses and cross-diagnostic aggressive behavior.

Method

Participants

Participants were 384 youth consecutively referred for neuropsychiatric evaluation who were administered the TSI along with other measures in a larger, source study. This source study, the [MASK 2], collects cognitive, psychiatric, and genomic data on youth evaluated at a pediatric assessment clinic within the psychiatry department at an academic teaching hospital. Reasons for referral include assistance with differential diagnosis or with treatment or school planning, as well as need for documentation of cognitive strengths and weaknesses. Participants in the current study were recruited from May 2013 to May 2015, during a period in which all participants received the TSI. The sample included 252 males and 132 females, age 6 to 18 years ($M = 11.15$, $SD = 3.22$). Range of IQ estimated by the Wechsler General Ability Index was 71 – 146 ($M = 103.44$, $SD = 14.30$). The racial composition of the sample was 85.4% White, 3.9% African American, 2.9% Asian, and 7.3% Other, with 10.4% of the total reporting Hispanic/Latino ethnicity.

Psychiatric diagnoses were assigned to each participant after a comprehensive evaluation by, or under the supervision of, licensed clinical psychologists. A diagnosis was assigned within the primary diagnostic record if full DSM-IV-TR (American Psychiatric Association, 2000) criteria were unambiguously met based on information from: 1) clinical interviews with a caregiver/ guardian and the patient; 2) review of available medical records; and 3) review of omnibus and targeted behavioral rating scales. One exception to this rule was that a diagnosis of ADHD was allowed in the presence of an ASD, as is now true in DSM-5 (American Psychiatric Association, 2013). Additional diagnoses of oppositional defiant and conduct disorders (ODD/ CD) that were considered secondary to particular concerns (i.e. resulting from difficulties with executive functions or expressive language) were also available via symptom-level parent-reports from the clinical record. Diagnostic information and parental education (as a proxy for the family's socio-economic status) for our sample are shown in Table 1.

To enroll in [MASK 2], guardians provided written informed consent and youth provided written assent, and the study, including administration of the TSI, was approved by the Institutional Review Board of the hospital in which the participants were ascertained.

Measures

Thinking Skills Inventory (TSI).—The TSI was developed by licensed clinical staff at [MASK 3], a program in the Psychiatry Department at [MASK 1], in the context of providing mental health treatment to youth exhibiting challenging behaviors. The first iteration of the clinical tool that would become the TSI was first published more than a decade ago (Greene & Ablon, 2005), and was refined over time for the purpose of providing clinically useful information during treatment. The original measure has 27 items, with each item reflecting a skill deficit hypothesized by clinical staff as a possible cause of children's behavioral struggles (e.g., "Difficulty considering a range of solutions to a problem"). In subsequent versions of the scale, the items were rephrased to reflect skill abilities rather than deficits (e.g., "Considers a range of solutions to a problem"), and a dichotomous rating scale was expanded to a 5-point Likert scale, where 1 = Consistent Strength; 2 = Sometimes a Strength; 3 = Depends; 4 = Sometimes Difficult; and 5 = Consistently Difficult. At the start of the present study, four licensed psychologists at Think:Kids agreed upon splitting three items and other minor wording changes for the purpose of improving clarity and reducing ambiguity. This resulted in a 30-item scale organized in five skill domains, which were labeled as Executive Functioning Skills, Language Processing Skills, Emotion Regulation Skills, Cognitive Flexibility Skills, and Social Skills (see Appendix A).

Measures used to establish validity.—Measures used to examine validity are summarized in Table 2. To establish convergent validity, we selected caregiver-report measures from our source study that represented the five skill domains that emerged from the factor analyses. **Attention and Working Memory** was measured by the Attention Problems subscale of the Child Behavior Checklist (CBCL/6–18; Achenbach & Rescorla, 2001), and the Working Memory subscale of the Behavior Rating Inventory of Executive Functions - Parent version (BRIEF-P, Gioia, Isquith, Guy, & Kenworthy, 2000). **Language and Communication** was measured by the Communication subscale of the Adaptive Behavior Assessment System (ABAS-II; Harrison & Oakland, 2000). **Emotion Regulation** was measured by the Emotional Control subscale of the BRIEF-P (Gioia et al., 2000). **Cognitive Flexibility** was measured by the Shift subscale of the BRIEF-P (Gioia et al., 2000). **Social Thinking** was measured by the Social Problems subscale of the CBCL/6–18 (Achenbach & Rescorla, 2001). These measures are commonly used in both research and clinical settings, with good psychometric properties (for CBCL/6–18, Cronbach's $\alpha = 0.78$ – 0.97 , test-retest reliability = 0.90 , Achenbach & Rescorla, 2001; for BRIEF parent form, Cronbach's $\alpha = 0.80$ – 0.98 , test-retest reliability = 0.82 , Gioia, 2000; for ABAS-II, Cronbach's α and test-retest reliability > 0.90 , Harrison & Oakland, 2003; internal reliabilities of these measures were not calculated with this sample, because item-level scores were not recorded). Subscales of both CBCL/6–18 and BRIEF are presented as T-scores, with higher scores suggesting more severe problems in the assessed domain. ABAS-II is a measure of adaptive behavior, with lower scores representing more problems in the skill areas.

To assess divergent validity of the TSI, and to determine the impact of common method variance from caregivers' reports (see Lindell, & Whitney, 2001; Podsakoff, MacKenzie, Lee, & Podsakoff, 2003), the neuropsychiatric dimension of somatization was assessed with

the Somatic Complaints subscale of the CBCL. Somatization was chosen because it is separate from other aspects of psychopathology based on large scale phenotypic modeling (Kotov et al., 2011). Finally, criterion-related validity was explored using the Aggressive Behavior subscale of the CBCL/6–18 (Achenbach, 2009) in regression models (as explained in the analytic plan).

Analytic Plan

Because no empirical analyses on the TSI had been completed for item selection or structure exploration, we first used exploratory factor analysis (EFA) with half of the data to select/delete/move items (if applicable), and to find possible latent structures of the TSI; then we applied confirmatory factor analysis (CFA) to check the structure using the other half of the sample. Next, we examined the psychometric properties of each subscale of the TSI identified in factor analyses, including its internal consistency, convergent and divergent validity, and criterion-related validity.

All analyses were done in RStudio (2016; version 1.0.143), using appropriate statistical packages: (1) Mardia's Multivariate Normality Test implemented using the MVN library (Korkmaz, Goksuluk, & Zararsiz, 2014) suggested that the data were not multivariate normally distributed. Then using the MissMech library (Jaamshidian, Jalal, & Jansen, 2014), non-parametric test confirmed that missing values were missing (completely) at random (MCAR), so that missing entries were deleted listwise for the factor analysis (resulting in $N = 326$). (2) The original sample was randomly split into two, and non-parametric multivariate analysis implemented by the MNM library (Nordhausen & Oja, 2011) suggested that the two random samples were equivalent in terms of participants' responses on the TSI items. Sample A ($n = 164$) was used for exploratory factor analysis (EFA) using polychoric correlations (Dolan, 1994) and principal axis factoring with oblique rotation to allow correlations among factors (Russell, 2002) in Revelle's psych library (2017). Sample B ($n = 162$) was used for confirmatory factor analysis (CFA) using maximum likelihood estimator with robust standard errors (MLR) for estimations in Rosseel's lavaan library (2012). CFA was implemented on the TSI in its entirety, to first compare possible structures suggested by EFA. Second order factor model was explored with highly correlated factors. After the structure was confirmed, CFA was applied to each subscale separately. Specific paths from items to more than one latent variable or residual errors of highly correlated items were added sequentially based on the inspection of modification indexes.

We applied pair-wise deletion of missing values for reliability and validity assessments. Reliability of each domain was assessed by internal consistency (Cronbach's alpha), and average inter-item correlations (Allen & Yen, 2002). Convergent and divergent validity of the TSI were examined by correlating scores of the TSI with the corresponding scores from standard caregiver-report measures for each cognitive domain as well as for a negative control.

To explore skills by diagnostic categories, and to examine the cross-diagnostic relevance of the scale, we evaluated the relevance of weaknesses in skill domains across different diagnostic groupings. We tested whether skills assessed by the TSI were selectively associated with the presence of a diagnosis from any of the four major diagnostic categories

in the primary diagnostic record (including autism spectrum disorders, mood disorders, ADHD, and anxiety disorders). We did not incorporate ODD/CD in these analyses because the majority of these diagnoses were obtained from a separate clinical source, as they were considered secondary to other issues. The presence of each diagnosis was treated as a dummy variable and entered in regression models, together with youth's IQ, sex, and age (later dropped for no significant impact), to predict skills as measured by the TSI.

To assess criterion-related validity, we examined the association of TSI scores to crossdiagnostic behavioral concerns. Since the TSI is hypothesized to assess lagging skills common in children and youth exhibiting challenging behavior, we tested whether lagging skills as measured by the TSI predicted aggressive behavior in our sample. We tested regression models with scores on the Aggressive Behavior subscale of the CBCL as the dependent variable. Youth's age and IQ estimated by the Wechsler General Ability Index were entered in the first block of regression models as covariates (youth's sex, ethnicity and each parent's highest education were not significant according to preliminary analyses), and TSI subscale scores were entered individually and separately to the second block, to examine whether and to what extent each lagging skill could explain youths' aggressive behavior beyond the effects of the covariates. Then, all lagging skills were entered simultaneously to the second block to assess the relative importance of the lagging skills in explaining variance in aggressive behavior across diagnostic categories. Missing values in all regression models were handled by listwise deletion.

Results

Exploratory Factor Analysis

Factors were selected based on the scree plot (see Figure 1), the variance threshold of 5% for any one factor, and the theoretical meaning of each factor. Either a four-or five-factor model was compatible with the data. Informed by item contents and the factor loadings of both models, we deleted Items 8 and 23 for loadings on multiple factors in both models, and items 18 and 19, which loaded weakly on all domains. The final scale contains 26 items. Item 1 was moved from its original place in the Executive Functioning domain to the Cognitive Flexibility domain, because it was more closely associated with cognitive flexibility than with the remaining executive function items, which were recognized as better measuring attention and working memory, and thus the latter was renamed. The factor loadings of the final scale are presented in Table 3. Based on the face validity of the items in each subscale, the five-factor model includes: Attention and Working Memory Skills, Language and Communication Skills, Emotion Regulation Skills, Cognitive Flexibility Skills, and Social Thinking Skills, which cumulatively explain 53.5% of the variance. The four-factor model explained 53.6% of the variance, and the pattern was similar, except that the subscales for Emotion Regulation and Cognitive Flexibility were merged into one factor.

Confirmatory Factor Analysis

Based on results of the EFA and the theory underlying the scale, two equivalent models were considered and compared in CFA, with four-or five first-order factors. Overall, modifications informed by the Modification Index (MI) were the same among all models¹. As shown in

Table 4, after modifications, the model fits were adequate to good according to the indices. More importantly, the five-factor model fit better than the four-factor model ($\chi^2 = 85.98$, $df = 4$, $p < .001$), both with and without modifications. Given the high correlation between the latent variables for Emotion Regulation and Cognitive Flexibility, and since these two factors have been found to be highly correlated on the BRIEF (Emotional Control and Shift, together with Inhibition, make up the BRIEF's Behavioral Regulation Index), a second-order model with five first-order factors and one second-order factor combining Emotion Regulation and Cognitive Flexibility was also evaluated. With the same modifications, the second-order model (3b) resulted in excellent fit, similar to the first-order five-factor model (2b). But statistical comparisons favor the first-order model ($\chi^2 = 12.57$, $df = 2$, $p < .01$; smaller AIC and BIC).

Overall, a five-factor solution was more compatible with the data compared to four factors. Therefore, a CFA was applied to each of the five domains separately. For each subscale, we specified one factor; and given the CFA outcomes above, error correlations between Items 10 & 11 (expressive language), and between Items 24 & 25 (non-verbal social cues) were added to the subscales of Language & Communication, and Social Thinking Skills, respectively.

Within the skill domains of Emotion Regulation, Cognitive Flexibility, and Attention & Working Memory, the fit of a one-factor model is acceptable. A one-factor model for Social Thinking is acceptable, as long as the errors of the two items on non-verbal social cues are allowed to correlate. Finally, the one-factor model for Language & Communication was significantly improved by allowing the two items on expressive language to correlate, suggesting that there were variances specific to receptive and expressive language items, and these two sub-domains are better kept separate (model fit details for each subscale is summarized in Supplementary Table 1).

The descriptive results of the five domains of the TSI are shown in Table 5. The subscales were all at least moderately correlated, and Emotion Regulation and Cognitive Flexibility were highly correlated ($r = .71$). Descriptive results of all standard measures are summarized in Table 6.

Psychometric Properties

Reliability.—Overall, the internal consistency as measured by Cronbach's alpha was moderate to high for each domain, ranging from 0.84 to 0.91 (see Table 5). Average inter-item correlations ranged from 0.49 to 0.70, suggesting that items within each domain cohere well with the overall construct, with higher values observed for domains with fewer items (especially Emotion Regulation).

¹Modifications include: error correlations between Items 10 & 11 on expressive language; Items 24 & 25 on non-verbal social cues; error correlations between Item 20 & 21 on ability to handle deviations from expectations or routines error; added path between Item 9, Item 14 and Social Thinking skills; and added path between Item 7 and Cognitive Flexibility.

Convergent validity.

As shown in Table 7, the subscales of the TSI were highly correlated with established caregiver-report measures of the corresponding domain, providing evidence for the convergent validity of the TSI. Specific results are reported by skill domain: (1) For the Attention and Working Memory domain, the two established measures and the TSI were highly correlated with one another. The lowest correlation between the TSI and established measures of Attention/Working Memory domain was $r = 0.63$, which was higher than the strongest correlation between the TSI's Attention/Working Memory domain and measures of other domains, providing support for both convergent and divergent validity. (2) For the Language/Communication domain, the TSI correlated well with the communication skills measured by ABAS-II ($r = -.54$), and the TSI had weaker correlations with measures of other domains. (3) For the Emotion Regulation domain, the TSI was strongly correlated with the BRIEF Emotional Control subscale ($r = 0.74$). Interestingly, the Emotion Regulation domain of the TSI also highly correlated with the BRIEF Shift scale. Such patterns were mirrored by the correlation patterns between BRIEF Emotional Control scale and BRIEF Shift scale. (4) For the Cognitive Flexibility domain, the TSI demonstrated strong correlations with the Shift scale of the BRIEF, and weaker correlations with measures of the other domains. (5) For the Social Thinking Skills domain, the TSI demonstrated strong correlations with Social Problems scale of CBCL. Social skills tend to moderately correlate with communication skills, emotion regulation and cognitive flexibility, and this was true for social skills measured by both the TSI and CBCL, providing further support for the validity of the TSI.

Divergent validity.—The Somatic Complaints subscale of the CBCL was used as a negative control to assess the divergent validity of the TSI domains. Overall, we found low correlations between each domain of the TSI and Somatic Complaints subscale of the CBCL (ranging from 0.01 to 0.21), providing support for the divergent validity of the TSI domains, and suggesting that the observed cross-correlations between the TSI domains and corresponding constructs were not likely the result of shared method variance. Other selected measures also correlated weakly with Somatic Complaints (ranging from 0.08 to 0.26), except for a moderate correlation between Somatic Complaints and the Social Problems scale of CBCL ($r = 0.35$). This was unexpected, but may be a spurious finding or a function of shared variance of these particular constructs.

Skills by diagnostic categories.—We next tested whether particular skills were associated with particular diagnostic categories. While this relationship is not critical for the purpose of the current study (i.e., to screen for skill deficits that may be good targets for treating behavioral symptoms), it may provide additional information regarding the validity of the TSI subscales and support for its intended use as a cross-diagnostic measure.

The sample was diagnostically heterogeneous, with 287 youth assigned a diagnosis in at least one of the four major diagnostic categories (non-exclusive, see Table 1), including ASD, mood disorders (MOOD), ADHD, and anxiety disorders (ANX). Figure 2 illustrates the frequency of each of these diagnostic categories, alone and combined. We focused on the effects of these four diagnostic categories in predicting the five skill domains measured by

the TSI, and used Bonferroni corrections to adjust for multiple testing. Results of regression models with each skill domain of the TSI as the dependent variable and the diagnostic categories as predictors are summarized in Table 8 (see Supplementary Table 2 for detailed results). Given that all diagnostic categories were entered together in the regression models, associations between a particular diagnostic category and the TSI subscale scores were estimated while controlling for the impact of the remaining diagnostic categories.

Results indicate that deficits in emotion regulation skills are associated with having a diagnosis of autism spectrum, mood, or anxiety disorders, suggesting that emotion regulation may contribute to symptoms or behaviors observed in several disorders. Similarly, cognitive flexibility skills and social thinking skills are both associated with having a diagnosis of mood or autism spectrum disorders, while attention/working memory skills are associated only with having a diagnosis of ADHD. Communication skills were not associated with having any of the diagnoses in our sample; youth who were diagnosed with any of the four diagnostic categories did not have significantly different communication skills than youth who were subclinical, after statistical correction. Furthermore, the average score on all TSI subscales for subclinical youth (those assigned none of the four major psychiatric diagnoses) was below 3, suggesting that their average skills in the five domains were reported as ‘strengths.’ Youth with at least one diagnosis within the four major diagnostic categories scored significantly higher than those with no diagnosis, and specifically, their skills in the Attention & Working Memory, Emotion Regulation, and Cognitive Flexibility domains were reported as relative weaknesses (average scores above 3).

Criterion-related validity.—Despite some variation of skill deficits by disorder, most skill deficits occurred cross-diagnostically. Therefore, we focused on youth’s aggressive behavior independent of diagnosis, to explore the criterion-related validity of the TSI. Results of regression models with each cognitive skill domain entered separately as a predictor are summarized in Table 9. All five domains of cognitive skills assessed by the TSI, including the domains of Attention and Working Memory, Language and Communication, Emotion Regulation, Cognitive Flexibility, and Social Thinking, were significant predictors of aggressive behavior, after controlling for the effect of age and IQ.

To investigate the relative contribution of each lagging skill to aggressive behavior, all lagging skills were entered the regression model simultaneously (see Table 10). Given the high correlations between cognitive flexibility and emotion regulation skills, we were concerned about multi-collinearity. We tested four models with: (1) emotion regulation score representing both domains; (2) both emotion regulation and cognitive flexibility; (3) the average scores of these two; or (4) using the score of cognitive flexibility as the predictor; and found similar patterns of results. Given that emotion regulation was a stronger predictor of aggressive behavior, we kept the model that used emotion regulation to represent the two domains. As shown in Table 10, the general models accounted for over 40% of the variance in aggressive behavior, providing further support for the hypothesis that behavior symptoms are associated with lagging neurocognitive skills, and for the utility of the TSI. Skill domains that significantly predicted aggressive behavior were Emotion Regulation/Cognitive

Flexibility; Social Thinking Skills, and Language & Communication, with Emotion Regulation being the strongest predictor.

To control for the impact of common method variances of caregiver's report, we applied the partial correlation approach using Somatic Complaints subscale of CBCL/6–18 as the marker variable (Lindell, & Whitney, 2001; Podsakoff, MacKenzie, Lee, & Podsakoff 2003). After controlling for the bias, the variance in aggressive behavior accounted for by the TSI skills was 36.8%, similar to the explained variance without controlling for method variance. Furthermore, the same skill domains as above were significant predictors of aggressive behavior, with Emotion Regulation being the strongest predictor. These patterns suggest that skills measured by the TSI accounted for the variance in aggressive behavior, beyond shared method variance of caregiver-report.

Discussion

In this study, we evaluated the psychometric properties and criterion-related validity of a newly revised clinical instrument, the Thinking Skills Inventory (TSI). The TSI attempts to identify skill deficits relevant to youth mental health treatment, including attention and working memory, emotion regulation, social cognition, cognitive flexibility, and language/communication skills. The TSI was developed by licensed clinicians who are experts in treatment that specifically targets skill deficits underlying challenging behavior for remediation of behavioral concerns. Since challenging behavior can be present for youth exhibiting a variety of mental health concerns, and a variety of skill-focused therapies are becoming available, this instrument is relevant to child and adolescent therapeutic settings generally, regardless of specific intervention utilized. Thus, the TSI has potential relevance for screening in general youth clinical settings. To our knowledge, there are no brief, freely-available measures that assess all five of these therapeutically-relevant skills together.

Based on factor analyses outlined above, caregiver-reported TSI scores in a sample of clinically-referred youth reflected five internally reliable subscales: Attention and Working Memory Skills; Language and Communication Skills; Emotion Regulation Skills; Cognitive Flexibility Skills; and Social Thinking Skills. These five subscales demonstrated moderate to high internal consistency, ranging from 0.84 to 0.91. The measure also demonstrated convergent validity; scores were correlated with commonly-used and well-established caregiver-report measures of corresponding domains. In particular, the correlation patterns between three domains assessed by the TSI (attention and working memory, emotion regulation, and cognitive flexibility) and all other established measures were mirrored by the corresponding subscales of BRIEF (working memory, emotion control, and shift subscale, respectively); the TSI domain of language and communication was highly correlated with the communication scale of ABAS-II; and finally, patterns of correlations between the TSI domain of social thinking skills and all other measures closely resembled that of the CBCL-Social Problems scale.

Not surprisingly, we also found correlations among many of the skills, both measured by the TSI and by standard measures. The relationship between language/communication skills and social thinking skills is reasonable, because the ability to use and understand language is

often important in communications and social interactions with others. The moderate correlation between skills of emotion regulation (the Emotion Control subscale of the BRIEF or the Emotion Regulation domain of the TSI) and cognitive flexibility (the Shift subscale of the BRIEF or the Cognitive Flexibility domain of the TSI) is also expected, since both fall under the umbrella construct of executive functions and may reflect overlapping neural systems (Hendricks & Buchanan 2016). The moderate correlation between emotion regulation and social skills in this study is consistent with some previous studies (e.g., Berkovits & Baker, 2014; Eisenberg et al., 1993; Lopes, Salovey, Côté, & Beers, 2005), and suggests that emotion dysregulation, at least in a youth clinical population, might often occur in the context of poor social interactions. Given that both cognitive flexibility and social skills were correlated with emotion regulation, the moderate relationship between cognitive flexibility and social skills could be the result of their common correlate; it is also possible that poor cognitive flexibility may impact the quality of social interactions directly. This is a topic worthy of longitudinal study.

Finally, associations between skills and youths' diagnoses suggest that skill deficits are largely cross-diagnostic: youth with different diagnoses may have skills deficits in similar domains; and lagging skills in certain domains are not exclusively associated with one particular diagnosis (the only exception in this clinically referred sample was attention/working memory skills, which were mainly associated with ADHD). A deficit in four of the five cognitive skill domains (all except the language and communication skills in our sample) does provide information that the youth is more likely to have a diagnosable mental health concern. It is important to note, however, that although at a population level, diagnosis and skills are correlated, at the individual level, these skills will vary across individuals with the same diagnosis. Furthermore, a key finding is that TSI skills were associated with cross-diagnostic aggressive behavior in our sample, especially emotion regulation skills that explained the largest variance in aggressive behavior (for consistent findings that emotion regulation is a key skill related to challenging behavior, refer to Röhl, Koglin, & Petermann, 2012). Thus, it is the understanding of any one child's skill strengths and weaknesses that allows us to develop a personalized intervention. This adds support for a more dimensional and mechanism-focused framework for understanding and treating any individual's mental health concerns. As new therapeutic approaches are developed that address skills deficits underlying psychiatric diagnoses, research will emerge that indicates whether building skills leads to psychopathology remission and other treatment effects (e.g., for a recent example with emotion regulation skills, see Renna, Quintero, Fresco, & Mennin, 2017). Future research can also explore whether the TSI can be used to track skills growth as a result of treatment for specific skill deficits.

The domains of the TSI were derived conceptually, to identify strengths and weaknesses in skill domains that frequently occur in relation to child psychopathology and may indeed underlie symptoms of emotional and behavioral disorders. Although the original clinical tool on which the current TSI is based was designed prior to NIMH's RDoC framework, we note the convergence of TSI domains with several domains of RDoC. For example, the TSI's domains of attention/working memory, cognitive flexibility, and language overlap with RDoC's cognitive systems; the TSI domain of social thinking skills overlaps with RDoC's systems for social process; and the TSI domain of emotion regulation overlaps with RDoC's

concept of negative and positive valence (Meyers, DeSerisy, & Roy, 2017). This convergence with RDoC is not a coincidence but rather a reflection that such skills are viewed as fundamental to successful functioning across a range of contexts.

In sum, we have demonstrated that the TSI is a reliable and valid instrument. Nonetheless, we note areas that may benefit from further improvement. First, development of a normative base would enhance the ability of the TSI to more precisely assess the degree of deficit in different skills at different ages. Second, measurement properties may differ in pediatric or other settings in which there is less symptom variation at the more severe end of the distribution; thus, this research should be replicated in a sample of typically developing children. Third, while the TSI was intended to assist treatment planning for youth with behavioral concerns, there were only six youth in our sample (1.6%) who received a primary diagnosis of oppositional defiant disorder (ODD) or conduct disorder (CD). Rather, in the pediatric assessment clinic where this study was conducted, the majority of diagnoses of CD and ODD (16.4% and 7.6% respectively) were considered secondary manifestations of struggles in areas such as language, mood regulation, and executive functioning, consistent with the rationale behind the creation of the TSI. While the overall rate of ODD/CD in our participants is consistent with expectations for a clinically referred sample, we did not include the ODD/CD diagnostic groups in the analyses, as their diagnoses came from a different source than youth from the other diagnostic groups. Thus, conclusions specific to youth diagnosed with ODD/CD are limited.

Replication in a sample of youth referred specifically for behavior disorders would provide further insights on the criterion-related validity of the TSI.

Despite these limitations, we have established the psychometric properties of the TSI, a novel caregiver-report measure that aims to quickly identify cross-diagnostic skill deficits that may be contributing to youth psychopathology, behavior, and functioning. Furthermore, analyses revealed that TSI subscales predict cross-diagnostic aggressive behavior and that skill deficits are relevant across diagnostic groupings of patients. This freely available and concise measure should therefore be considered for use in youth clinical settings. Its comprehensive scope is unique in screening potential targets for treatment, without more time-consuming and expensive measurement across multiple domains using multiple instruments. The validated TSI (see Appendix B) represents a potentially valuable addition to a standard array of clinical assessment measures. Results from the TSI may be used for caregivers to have a better understanding of the factors leading to challenging behavior, and for clinicians to guide treatment or referrals for more in-depth evaluation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported in part by the Stanley Center for Psychiatric Research, and NIH R03MH106862 awarded to Dr. Alysia E. Doyle. Drs. Pollastri, Wang, Hill, and Ablon work for Think:Kids, a non-profit program in the Department of Psychiatry at Massachusetts General Hospital, from which staff train and consult on

Collaborative Problem Solving and use the Thinking Skills Inventory (TSI). The TSI is available at no cost to paid clients of Think:Kids as well as to the larger research and clinical communities.

Appendix A.: The Thinking Skills Inventory (TSI) as it was administered for this study (with item numbers added for reference).

Thinking Skills Inventory (TSI)

Instructions:

Below is a list of thinking skills required to solve problems, be flexible, and tolerate frustration. Many children with social, emotional and behavioral challenges will have deficits in some of these areas. The skills are organized into five categories.

Please rate the extent to which each skill is a strength or challenge to your child by **marking an "X"** in one column for each skill.

Executive Functioning Skills	<i>Consistent Strength</i>	<i>Sometimes a Strength</i>	<i>Depends</i>	<i>Sometimes Difficult</i>	<i>Consistently Difficult</i>
(1) Handles transitions, shifts easily from one task to another					
(2) Sticks with tasks requiring sustained attention					
(3) Does things in a logical sequence or set order					
(4) Keeps track of time; correctly assesses how much time a task will take					
(5) Reflects on multiple thoughts or ideas at the same time					
(6) Maintains focus during activities					
(7) Ignores irrelevant noises, people, or other stimuli; tunes things out when necessary					
(8) Thinks before responding; considers the likely outcomes or consequences of his/her actions					
(9) Considers a range of solutions to a problem					
Language Processing Skills	<i>Consistent Strength</i>	<i>Sometimes a Strength</i>	<i>Depends</i>	<i>Sometimes Difficult</i>	<i>Consistently Difficult</i>
(10) Expresses concerns, needs, or thoughts in words					
(11) Is able to tell someone what's bothering him or her					
(12) Understands spoken directions					
(13) Understands and follows conversations					
Emotion Regulation Skills	<i>Consistent Strength</i>	<i>Sometimes a Strength</i>	<i>Depends</i>	<i>Sometimes Difficult</i>	<i>Consistently Difficult</i>
(14) Thinks rationally, even when					

frustrated					
(15) Manages irritability in an age-appropriate way					
(16) Manages anxiety in an age-appropriate way					
(17) Manages disappointment in an age-appropriate way					
Cognitive Flexibility Skills	<i>Consistent Strength</i>	<i>Sometimes a Strength</i>	<i>Depends</i>	<i>Sometimes Difficult</i>	<i>Consistently Difficult</i>
(18) Is able to see “shades of gray” rather than thinking only in “black-and-white”					
(19) Thinks hypothetically, is able to envision different possibilities					
(20) Handles deviations from rules, routines, and original plans					
(21) Handles unpredictability, ambiguity, uncertainty, and novelty					
(22) Takes into account situational factors that may mean a change in plans (Example: “If it rains, we may need to cancel the trip.”)					
(23) Interprets information accurately/ <u>avoids</u> over-generalization or personalizing (Example: <u>Avoids</u> saying “Everyone’s out to get me,” “Nobody likes me,” “You always blame me,” “It’s not fair,” “I’m stupid,” “Things will never work out for me.”)					
Social Skills	<i>Consistent Strength</i>	<i>Sometimes a Strength</i>	<i>Depends</i>	<i>Sometimes Difficult</i>	<i>Consistently Difficult</i>
(24) Pays attention to verbal and nonverbal social cues					
(25) Accurately interprets nonverbal social cues (like facial expressions and tone of voice)					
(26) Starts conversations with peers, enters groups of peers appropriately					
(27) Seeks attention in appropriate ways					
(28) Understands how his or her behavior affects other people					
(29) Empathizes with others, appreciates others’ perspectives or points of view					
(30) Understands how he or she is coming across or being perceived by others					

Appendix B: The validated Thinking Skills Inventory (TSI) ,after EFA and CFA.

Thinking Skills Inventory (TSI)

Instructions:

Below is a list of thinking skills required to solve problems, be flexible, and tolerate frustration. Many children with social, emotional and behavioral challenges will have deficits in some of these areas. The skills are organized into five categories.

Please rate the extent to which each skill is a strength or challenge to your child by **marking an “X”** in one column for each skill.

Attention and Working Memory	<i>Consistent Strength</i>	<i>Sometimes a Strength</i>	<i>Depends</i>	<i>Sometimes Difficult</i>	<i>Consistently Difficult</i>
Sticks with tasks requiring sustained attention					
Does things in a logical sequence or set order					
Keeps track of time; correctly assesses how much time a task will take					
Reflects on multiple thoughts or ideas at the same time					
Maintains focus during activities					
Ignores irrelevant noises, people, or other stimuli; tunes things out when necessary					
Considers a range of solutions to a problem					
Language and Communication Skills	<i>Consistent Strength</i>	<i>Sometimes a Strength</i>	<i>Depends</i>	<i>Sometimes Difficult</i>	<i>Consistently Difficult</i>
Expresses concerns, needs, or thoughts in words					
Is able to tell someone what's bothering him or her					
Understands spoken directions					
Understands and follows conversations					
Emotion Regulation Skills	<i>Consistent Strength</i>	<i>Sometimes a Strength</i>	<i>Depends</i>	<i>Sometimes Difficult</i>	<i>Consistently Difficult</i>
Thinks rationally, even when frustrated					
Manages irritability in an age-appropriate way					
Manages anxiety in an age-appropriate way					
Manages disappointment in an age-appropriate way					
Cognitive Flexibility Skills	<i>Consistent Strength</i>	<i>Sometimes a Strength</i>	<i>Depends</i>	<i>Sometimes Difficult</i>	<i>Consistently Difficult</i>
Handles transitions, shifts easily from one task to another					
Handles deviations from rules, routines, and original plans					
Handles unpredictability, ambiguity, uncertainty, and novelty					
Takes into account situational factors that may mean a change in plans (Example: "If it rains, we may need to cancel the trip.")					

Social Thinking Skills	<i>Consistent Strength</i>	<i>Sometimes a Strength</i>	<i>Depends</i>	<i>Sometimes Difficult</i>	<i>Consistently Difficult</i>
Pays attention to verbal and nonverbal social cues					
Accurately interprets nonverbal social cues (like facial expressions and tone of voice)					
Starts conversations with peers, enters groups of peers appropriately					
Seeks attention in appropriate ways					
Understands how his or her behavior affects other people					
Empathizes with others, appreciates others' perspectives or points of view					
Understands how he or she is coming across or being perceived by others					

The most current version and any future updates will be available online at: [MASK 4]

References

References

- Allen MJ & Yen WM (2002). Introduction to measurement theory. Long Grove, IL: Waveland Press.
- Achenbach TM (2009). ASEBA: Development, findings, theory and applications. Burlington, VT: University of Vermont Research Center for Children, Youth and Families.
- Achenbach TM & Rescorla LA (2001). Manual for the ASEBA school-age forms and profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families ISBN 0-938565-73-7
- American Psychiatric Association (2000). Diagnostic and statistical manual of mental disorders (4th ed., Text Revision). Washington, DC: Author.
- American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC: Author.
- Bentler PM & Bonett DG (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588–606. 10.1037/0033-2909.88.3.588
- Berkovits LD & Baker BL (2014). Emotion dysregulation and social competence: stability, change and predictive power. *Journal of Intellectual Disability Research*, 58(8), 765–776. 10.1111/jir.12088 [PubMed: 23957719]
- Browne MW & Cudeck R (1993). Alternative ways of assessing model fit In Bollen KA & Long SL (Eds.), *Testing structural equation models* (pp. 136–162). London, UK: Sage Publications.
- Cohen J (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ: Routledge.
- Dodge KA, Lansford JE, Burks VS, Bates JE, Pettit GS, Fontaine R, & Price JM (2003). Peer rejection and social information-processing factors in the development of aggressive behavior problems in children. *Child Development*, 74(2), 374–393. [PubMed: 12705561]
- Dolan C (1994). Factor analysis of variables with 2, 3, 5, and 7 response categories: A comparison of categorical variable estimators using simulated data. *British Journal of Mathematical and Statistical Psychology*, 47(2), 309–326.
- Doyle AE, Vuijk PJ, Doty ND, McGrath LM, Willoughby BL, O'Donnell EH, ... Braaten EB (2017). Cross-disorder cognitive impairments in youth referred for neuropsychiatric evaluation. *Journal of the International Neuropsychological Society*, 1–13.

- Eisenberg N, Fabes RA, Bernzweig J, Karbon M, Poulin R, & Hanish L (1993). The relations of emotionality and regulation to preschoolers' social skills and sociometric status. *Child Development*, 64(5), 1418–1438. 10.2307/1131543 [PubMed: 8222881]
- Gioia GA (2000). BRIEF: Behavior rating inventory of executive function: Professional manual. Odessa: Psychological Assessment Resources.
- Gioia G, Isquith PK, Guy SC, & Kenworthy L; Reviewed by Baron IS (2000). Test review: Behavior rating inventory of executive function. *Child Neuropsychology* 6 (3): 235–238. [PubMed: 11419452]
- Gioia GA, Isquith PK, Guy SC, & Kenworthy L (2015). Behavior rating inventory of executive functions-2. Lutz: PAR.
- Greene RW (1998). *The Explosive Child: A New Approach for Understanding and Parenting Easily Frustrated, Chronically Inflexible Children*. New York: Harper Collins.
- Greene RW & Ablon JS (2005). *Treating explosive kids: The collaborative problem-solving approach*. New York, NY: Guilford Press.
- Harrison PL & Oakland T (2003). *Adaptive Behavior Assessment System - Second Edition*, San Antonio, TX: The Psychological Corporation.
- Hendricks MA & Buchanan TW (2016). Individual differences in cognitive control processes and their relationship to emotion regulation. *Cognition & Emotion*, 30(5), 912–924. 10.1080/02699931.2015.1032893 [PubMed: 25947896]
- Hill AL, Degnan KA, Calkins SD, & Keane SP (2006). Profiles of externalizing behavior problems for boys and girls across preschool: the roles of emotion regulation and inattention. *Developmental Psychology*, 42(5), 913. [PubMed: 16953696]
- Hooper D, Coughlan J, & Mullen M (2008). Structural equation modeling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), 53–60.
- Hu L & Bentler PM (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424.
- Hu L & Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55.
- Jamshidian M, Jalal S, & Jansen C (2014). MissMech: An R package for testing homoscedasticity, multivariate normality, and missing completely at random (MCAR). *Journal of Statistical Software*, 56(6), 1–31
- Kline P (2013). *Handbook of psychological testing*. London, England: Routledge.
- Korkmaz S, Goksuluk D, & Zararsiz G (2014). MVN: An R package for assessing multivariate normality. *The R Journal*, 6(2), 151–162.
- Kotov R, Ruggero CJ, Krueger RF, Watson D, Yuan Q, & Zimmerman M (2011). New dimensions in the quantitative classification of mental illness. *Archives of General Psychiatry*, 68, 1003–1011. [PubMed: 21969458]
- Lindell MK & Whitney DJ (2001). Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology*, 86(1), 114–121. 10.1037//0021-9010.86.1.114 [PubMed: 11302223]
- Lopes PN, Salovey P, Côté S, & Beers M (2005). Emotion regulation abilities and the quality of social interaction. *Emotion*, 5(1), 113–118. 10.1037/1528-3542.5.1.113 [PubMed: 15755224]
- MacCallum RC, Browne MW, & Sugawara HM (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), 130.
- Matthys W, Cuperus JM, & Engeland HV (1999). Deficient social problem-solving in boys with ODD/CD, with ADHD, and with both disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, 38(3), 311–321. [PubMed: 10087693]
- Meyers E, DeSerisy M, & Roy AK (2017). Disruptive mood dysregulation disorder (DMDD): An RDoC perspective. *Journal of Affective Disorders*, 216, 117–122. 10.1016/j.jad.2016.08.007 [PubMed: 27554606]
- Morris SE & Cuthbert BN (2012). Research domain criteria: cognitive systems, neural circuits, and dimensions of behavior. *Dialogues in Clinical Neuroscience*, 14(1), 29–37. [PubMed: 22577302]

- Nordhausen K and Oja H (2011). Multivariate L1 Methods: The Package MNM, *Journal of Statistical Software*, 43, 1–28.
- Nunnally JC (1978). *Psychometric theory*. New York, NY: McGraw-Hill.
- Podsakoff PM, MacKenzie SB, Lee J-Y, & Podsakoff NP (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. 10.1037/0021-9010.88.5.879 [PubMed: 14516251]
- Pollastri AR, Epstein LD, Heath GH, & Ablon JS (2013). The collaborative problem solving approach: Outcomes across settings. *Harvard Review of Psychiatry*, 21(4), 188–199. [PubMed: 24651507]
- Renna ME, Quintero JM, Fresco DM, & Mennin DS (2017). Emotion Regulation Therapy: A mechanism-targeted treatment for disorders of distress. *Frontiers in Psychology*, 8, 1–14. [PubMed: 28197108]
- Team RStudio (2016). RStudio: integrated development for R. Boston, MA: RStudio, Inc Available from <http://www.rstudio.com/>.
- Revelle W (2017) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, <https://CRAN.R-project.org/package=p psychVersion = 1.7.8>.
- Röll J, Koglin U, & Petermann F (2012). Emotion regulation and childhood aggression: Longitudinal associations. *Child Psychiatry & Human Development*, 43(6), 909–923. [PubMed: 22528031]
- Rosseel Yves. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36.
- Rushton J, Bruckman D, & Kelleher K (2002). Primary care referral of children with psychosocial problems. *Archives of pediatrics & adolescent medicine*, 156(6), 592–598. [PubMed: 12038893]
- Russell D (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in *Personality and Social Psychology Bulletin*. *Personality and Social Psychology Bulletin*, 28(12), 1629–46.
- Steiger JH (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences*, 42 (5), 893–98.
- Stieben J, Lewis MD, Granic I, Zelazo PD, Segalowitz S, & Pepler D (2007). Neurophysiological mechanisms of emotion regulation for subtypes of externalizing children. *Development and Psychopathology*, 19(2), 455–480. [PubMed: 17459179]
- Tabachnick BG & Fidell LS (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon/Pearson Education.
- Van Goozen SH, Cohen-Kettenis PT, Snoek H, Matthys W, Swaab-Barneveld H, & Van Engeland H (2004). Executive functioning in children: A comparison of hospitalized ODD and ODD/ADHD children and normal controls. *Journal of Child Psychology and Psychiatry*, 45(2), 284–292. [PubMed: 14982242]
- Zadeh ZY, Im-Bolter N, & Cohen NJ (2007). Social cognition and externalizing psychopathology: An investigation of the mediating role of language. *Journal of abnormal child psychology*, 35(2), 141–152. [PubMed: 17191148]

Non Graphical Solutions to Scree Test

Fig. 1.
Scree plot of exploratory factor analysis (EFA) results

Fig. 2.

Pie chart for number of youth assigned a diagnosis within any of the four diagnostic categories (alone or in combination), including autism spectrum (ASD), mood (MOOD), attention deficit/hyperactivity (ADHD), and anxiety disorders (ANX).

Table 1.

Diagnostic information and parent education level for the study sample

Diagnosis¹ (N = 384)	N	Percent (%)
Autism spectrum (Asperger's syndrome, Autistic disorder, Pervasive developmental disorder NOS)	54	14.1
Mood disorders (Bipolar disorder, Mood disorder NOS, Major depressive disorder, Dysthymia)	66	17.2
Attention-deficit/hyperactivity disorder (ADHD)	213	55.5
Anxiety disorders (Generalized anxiety disorder, Obsessive compulsive disorder, Panic disorder, Anxiety disorder NOS)	112	29.2
Nonaffective psychosis (Schizophrenia, Psychosis disorder NOS, Prodromal symptoms)	7	1.8
Oppositional Defiant Disorders or Conduct Disorders (Primary diagnosis of ODD or CD)	6	1.6
(Primary or secondary diagnoses of ODD or CD)	98	25.5
Parents' Highest Education²	Mother	Father
Did not complete High School (up to 11 years)	5	11
Completed High School	46	75
Some College	62	42
Completed College	129	95
More than College	115	93

¹: Diagnoses above are not mutually exclusive. Sum of Ns will exceed total samples size due to comorbidity.

²: Twenty-one participants were missing the highest education for both parents.

Table 2.

Caregiver-report measures used for validation & prediction analysis

Variables	Source Test	Subscale/Measurement
Aggressive Behavior	CBCL/6–18	Aggressive Behavior
Attention & Working Memory	BRIEF	Working Memory
Attention & Working Memory	CBCL/6–18	Attention Problem Scale
Language & Communication	ABAS-II	Communication
Emotion Regulation	BRIEF	Emotional Control
Cognitive Flexibility	BRIEF	Shift
Social Thinking	CBCL/6–18	Social Problem
Somatization	CBCL/6–18	Somatic Complaints

Table 3.

EFA factor loadings with oblique rotation, for five-and four-factor models respectively

FIVE FACTORS							FOUR FACTORS				
	ATT	LAN	ER	FLEX	SOC	Com-munality	ATT	LAN	EF	SOC	Com-munality
ATTENTION/WORKING MEMORY							ATTENTION/WORKING MEMORY				
TSI2	.837	-.099	.029	.004	-.004	.654	.817	-.102	.040	.017	.644
TSI3	.720	.254	-.042	.051	-.064	.709	.718	.261	.015	-.050	.710
TSI4	.714	.031	.021	.060	.053	.614	.707	.031	.080	.067	.613
TSI5	.421	.181	-.116	.215	.128	.483	.464	.204	.079	.123	.471
TSI6	.852	-.024	.061	-.154	.076	.673	.790	-.028	-.060	.103	.625
TSI7	.634	-.050	-.045	.182	-.033	.473	.660	-.040	.121	-.029	.474
TSI9	.416	.148	.167	.212	-.033	.474	.400	.134	.360	-.040	.475
LANGUAGE & COMMUNICATION							LANGUAGE & COMMUNICATION				
TSI10	-.060	.825	.179	-.045	.095	.815	-.116	.794	.146	.101	.790
TSI11	.040	.635	.264	-.010	.045	.626	-.026	.591	.257	.055	.591
TSI12	.146	.704	-.129	.081	-.037	.585	.163	.730	-.046	-.041	.588
TSI13	.074	.721	-.201	.114	.097	.676	.115	.752	-.089	.091	.662
EMOTION REGULATION							EMOTION REG/COGNITIVE FLEX				
TSI14	.048	.185	.483	.115	.202	.630	-.036	.135	.570	.194	.596
TSI15	.055	.057	.672	.107	.165	.755	-.061	-.010	.728	.165	.664
TSI16	.104	.038	.626	.240	.040	.727	.007	-.024	.813	.031	.680
TSI17	.058	.056	.643	.261	.058	.785	-.038	-.009	.850	.045	.741
COGNITIVE FLEXIBILITY											
TSI1	.173	.063	.037	.618	.108	.694	.247	.082	.566	.076	.635
TSI20	-.013	-.037	-.021	.892	.071	.813	.120	.009	.700	.040	.617
TSI21	-.019	.016	.058	.867	.040	.848	.097	.043	.763	.011	.697
TSI22	.013	.082	.248	.649	-.079	.645	.047	.073	.811	-.122	.628
SOCIAL THINKING SKILLS							SOCIAL THINKING SKILLS				
TSI24	.120	.103	-.120	.081	.733	.713	.162	.135	-.058	.722	.691
TSI25	-.027	.080	-.130	.047	.870	.776	.019	.115	-.101	.854	.746
TSI26	-.069	.194	.191	.066	.490	.547	-.096	.180	.236	.485	.548
TSI27	.101	-.011	.293	.065	.537	.636	.056	-.036	.333	.535	.629
TSI28	.117	.026	.223	.004	.643	.696	.080	.007	.206	.648	.691
TSI29	-.057	.214	.180	.003	.557	.598	-.088	.199	.167	.558	.596
TSI30	-.008	-.108	.066	.021	.878	.750	.000	-.108	.053	.887	.756

Note: Following the convention, items with loadings on more than one factor (≥ 0.35), or with no loadings on any factor (< 0.35) were dropped and the cells were left empty. ATT/WM = Attention & Working Memory; COMM = Language & Communication; ER = Emotion Regulation; FLEX = Cognitive Flexibility; SOC = Social Thinking; EF = Emotion Regulation/Cognitive Flexibility.

Table 4.

Confirmatory Factor Analysis and model comparisons for the TSI in its entirety

Models	df	χ^2	Relative χ^2/df	SRMSR	RMSEA	90% CI	CFI	AIC	BIC
1. Four factors	293	790.34	2.70	0.094	0.108	0.099-0.117	0.812	11441	11435
2. Five factors	289	704.35	2.44	0.095	0.099	0.090-0.109	0.845	11347	11340
3. Second-order Five factors	291	708.82	2.44	0.097	0.099	0.090-0.109	0.843	11349	11342
1b. Four factors, MI*	287	515.39	1.80	0.081	0.074	0.064-0.084	0.914	11145	11137
2b. Five factors, MI	283	468.75	1.66	0.073	0.067	0.056-0.078	0.930	11100	11092
3b. Second-order Five factors, MI	285	481.32	1.69	0.076	0.069	0.058-0.079	0.926	11109	11102

Note: df = degrees of freedom; χ^2 =Chi-square test (MLR); Relative $\chi^2 = \chi^2/\text{df}$; SRMSR = standardized root mean squared residual; RMSEA = root mean squared error of approximation (Robust); 90% CI = 90% confidence interval for RMSEA; CFI = comparative fit index (Robust).

*: The modifications informed by Modification Index were the same among all models.

Fit indices (Hooper, Coughlan, & Mullen, 2008) included: χ^2 test, relative χ^2 (χ^2/df , Tabachnick & Fidell, 2007); Comparative Fit Index (CFI, Hu & Bentler, 1999); Root Mean Square Error of Approximation (RMSEA, Hu & Bentler, 1999; Steiger, 2007; 90% confidence interval, Browne & Cudeck, 1993; MacCallum, Browne, & Sugawara, 1996); Standardized Root Mean Residual (SRMR, Hu & Bentler, 1998); differences in χ^2 (χ^2) were used for comparing nested models (Bentler & Bonet, 1980); and for non-nested models, Akaike Information Criterion (AIC) and sample-size-adjusted Bayesian Information Criterion (BIC) were used.

Table 5.

Five domains of the TSI: Reliability and Correlations

Domain	Subscale Mean	SD	Cronbach's alpha	Average inter-item correlation	COMM	ER	FLEX	SOC
ATT/WM	3.39	0.89	0.87	0.49	0.45	0.41	0.50	0.44
COMM	2.33	1.03	0.84	0.57		0.43	0.48	0.56
ER	3.20	1.15	0.90	0.70			0.71	0.67
FLEX	2.99	1.09	0.89	0.67				0.63
SOC	2.72	1.02	0.91	0.59				

Note: By convention, a Cronbach's alpha between .70 and .80 was considered acceptable, between .80 and .90 good, and higher than 0.90 excellent (Kline, 2013; Nunally, 1978). ATT/WM = Attention & Working Memory; COMM = Language & Communication; ER = Emotion Regulation; FLEX = Cognitive Flexibility; SOC = Social Thinking.

Table 6.

Descriptive results of the standard caregiver report measurements

Domain & Scale	Mean	SD	Range	% T scores ≥ 65
Challenging Behavior				
CBCL Aggressive Behavior	58.38	9.05	50–85	26.30
Attention & Working Memory				
CBCL Attention problems	64.19	10.20	50–97	44.27
BRIEF Working Memory	65.70	12.18	36–93	55.47
Language & Communication				
ABAS II Communication	8.46	3.41	1–15	–
Emotion Regulation				
BRIEF Emotion Control	57.88	13.58	35–91	33.59
Cognitive Flexibility				
BRIEF Shift	61.00	14.83	36–98	37.50
Social Thinking				
CBCL Social problems	59.33	8.26	50–88	26.82

Table 7.

Convergent and divergent validity of the TSI (correlations for convergent validity in gray)

	Attention & Work Mem		Language & Communication		Emotion Regulation		Cognitive Flexibility		Social Thinking	
	CBC L	TSI	ABA S	TSI	BRIF	TSI	BRIF	TSI	CBC L	TSI
BRIEF.WM	0.67	0.71	−0.35	0.35	0.28	0.23	0.33	0.30	0.32	0.33
CBCL.ATT		0.63	−0.43	0.33	0.35	0.38	0.33	0.36	0.51	0.50
ABAS.COM		−0.42		−0.54	−0.38	−0.51	−0.44	−0.47	−0.47	0.60
BRIEF.ER		0.30		0.27		0.74	0.72	0.59	0.51	0.53
BRIEF.FLX		0.35		0.35		0.59		0.69	0.56	0.54
CBCL.SOC		0.37		0.32		0.48		0.49		0.64

Note. correlations of 0.1 are interpreted as small, 0.3 as moderate, and 0.50 are considered large (Cohen, 1988). Cross-domain correlations can be read vertically, under each column for the TSI subscales. ATT/WM = Attention & Working Memory; COM = Language & Communication; ER = Emotion Regulation; FLEX = Cognitive Flexibility; SOC = Social Thinking.

* : On the ABAS-II, higher scores indicate better skills. For all other measures, higher scores indicate a greater deficit. Therefore, ABAS-II and other caregiver report measures were negatively correlated.

Table 8.TSI skills predicted by diagnostic categories^{*}

TSI SKILLS	Autism Spectrum	Mood Disorders	ADHD	Anxiety Disorders
	N = 54	N = 66	N = 213	N = 112
Attention & Working Memory	*	-		-
Communication	*	-	-	-
Emotion Regulation			-	
Cognitive Flexibility			-	*
Social Thinking			*	-

: $p < .0025$ (Bonferroni Correction, adjusted $\alpha = .05/20$);*
: $p < .05$

Each row represents a separate regression model, with youth's sex and IQ as covariates, and four main diagnostic categories as binary predictors.

*
: The diagnostic categories are not mutually exclusive. Sum of Ns will exceed the sample size due to comorbidity.

Table 9.

Results of linear regression models, with CBCL Aggressive Behavior subscale scores as the dependent variable and each TSI skill as predictor, entered separately.

Regression model		F-value	R ² (adjusted, %)	B	T
TSI (each lagging skill entered individually & separately)					
Block 1	Covariates: Age & IQ	5.56 [*]	2.38		
+ Block 2	Attention & Working Memory	14.59 ^{**}	7.45	2.92	5.63 ^{**}
+ Block 2	Communication	9.12 ^{**}	3.76	1.83	4.08 ^{**}
+ Block 2	Emotion Regulation	74.61 ^{**}	35.25	4.73	14.47 ^{**}
+ Block 2	Cognitive Flexibility	37.52 ^{**}	20.28	3.77	9.93 ^{**}
+ Block 2	Social Thinking	55.54 ^{**}	28.57	4.76	12.37 ^{**}

^{*}
: $p < .005$;

^{**}
: $p < .0001$;

Table 10.

Results of linear regression models, with CBCL Aggressive Behavior subscale scores as the dependent variable and all TSI skills as predictors, entered simultaneously.

Regression model		F-value	R ² (adjusted, %)	B	T
TSI (all lagging skills entered together)		45.77	42.46		
Block 1	Covariates: Age & IQ		2.38		*
+ Block 2	TSI skills		40.08		
	Attention & Working Memory			0.21	0.44 ^I
	Communication			-1.54	-3.47*
	Emotion Regulation ²			3.44	8.05**
	Social Thinking			2.92	5.56**
Partial correlation approach: Controlling for Common Method Variance		40.13	39.21		
Block 1	Covariates: Age & IQ		3.74		*
+ Block 2	Residuals of TSI skills & Aggressive Behavior (controlling the impact of Somatic Complaints)		36.83		
	Attention & Working Memory			0.06	0.13 ^I
	Communication			-1.41	-3.31*
	Emotion Regulation ²			3.12	7.51**
	Social Thinking			2.72	5.39**

^I: $p = \text{n.s.}$;

*: $p < .005$;

**: $p < .0001$

²: Cognitive Flexibility was not included, due to multi-collinearity (high correlations between cognitive flexibility & emotion regulation, $r = 0.72$ for TSI). In regression models with both domains, the effect of cognitive flexibility was not significant after controlling for emotion regulation; the changes of adjusted R² was less than 1% with or without cognitive flexibility.