



Published in final edited form as:

Psychol Assess. 2019 June ; 31(6): 781–792. doi:10.1037/pas0000699.

Development and Validation of the Knowledge of Effective Parenting Test (KEPT) in a Nationally Representative Sample

Oliver Lindhiem, Rachel A. Vaughn-Coaxum, Janelle Higa, Jordan L. Harris, David J. Kolko, and Paul A. Pilkonis

Department of Psychiatry, University of Pittsburgh School of Medicine

Abstract

We report on the development and psychometric properties of an instrument for the assessment of knowledge of effective parenting skills specific to conduct problems using an item response theory (IRT) framework. The initial item pool (36 items) for the Knowledge of Effective Parenting Test (KEPT) was administered online to a national sample ($N = 1,570$) selected to match the U.S. population on key demographic variables. Items with strong psychometric properties and without significant differential item functioning (DIF) by race/ethnicity were retained, resulting in a 21-item version of the KEPT with excellent reliability and validity. We also created a brief 10-item version of the KEPT to reduce respondent burden and to enhance its utility for repeated measurement in longitudinal and intervention research. We report norms and percentiles for both the 21-item version (KEPT-Full) and the 10-item version (KEPT-Brief).

Keywords

parenting; assessment; item response theory; disruptive behavior disorders

Effective parenting is a critical component of healthy child development and effective approaches are associated with fewer behavioral problems, decreased levels of aggression, and improved social skills (Gryczowski, Jordan, & Mercer, 2010; Kaiser, McBurnett, & Piffner, 2011). Effective parenting strategies are also associated with academic achievement and enhanced school performance (Swanson, Valiente, Lemery-Chalfant, & O'Brien, 2011). Parenting is also associated with long-term adjustment, affecting children throughout adolescence and into adulthood (Aquilino & Supple, 2001; Morris, Silk, Steinberg, Myers, & Robinson, 2007). For example, children raised with consistent and effective parental practices experience higher levels of positive adjustment (fewer depressive symptoms, decreased risk for substance abuse, and higher self-efficacy) in young adulthood. Conversely, children raised with ineffective parenting strategies experience lower levels of self-esteem and self-efficacy, decreased life satisfaction, heightened irritability, and an increased risk for substance abuse problems into young adulthood (Aquilino & Supple, 2001; Ben-Ami & Baker, 2012; Morris et al., 2007).

Parenting Skills as Target Mechanisms for Treating Disruptive Behavior Disorders

Parent management training (PMT) skills are specific treatment targets common across evidence-based psychosocial treatments for childhood disruptive behavior disorders (Lindhiem, Higa, Trentacosta, Herschell, & Kolko, 2014). Such treatments typically include (a) teaching caregivers skills including the use of praise, rewards, consequences (e.g., removing privileges), and time-outs, and (b) teaching children skills including problem-solving, communication, and emotion labeling (Eyberg, Nelson, & Boggs, 2008). Substantial evidence points to skill deficits in these areas as pathogenic and maintenance factors for disruptive behavior disorders. As a result, skills in these domains are conceptualized as active agents of therapeutic change and are specifically targeted in treatments. Most research on treatment adherence has focused on how well clinicians follow treatment protocols, and less research has focused on the extent to which patients and families acquire knowledge of relevant skills and use them outside of treatment sessions (Huey, Henggeler, Brondino, & Pickrel, 2000; Nock & Ferriter, 2005; Waltz, Addis, Koerner, & Jacobson, 1993). Despite the conceptual importance of skills as target mechanisms of treatment, skill acquisition in these domains is challenging to assess.

The Challenges of Measuring Parenting

Parenting involves complex and multifaceted patterns of behavior that are difficult to measure (e.g., Eddy, 2017; Lindhiem & Shaffer, 2017). Quantifying effective parenting is challenging for many reasons, including contextual factors, cultural considerations, and limits on observable behavior. Because of the complexity involved in measuring and quantifying the various dimensions of parenting, it is not surprising that existing measures are limited in several important ways, including recall biases inherent in retrospective self-reports, the contrived nature of laboratory observations, and the reactivity of home observations creating social desirability effects. Existing measures of parenting skills also tend to lack adequate psychometric properties or established normative data (Hurley, Huscroft-D'Angelo, Trout, Griffith, & Epstein, 2014). There have also been numerous calls for measures of parenting that are culturally sensitive (Beidas et al., 2015; Jensen-Doss & Hawley, 2010; Locke & Prinz, 2002). Thus, measures of parenting are still needed that reduce social desirability bias, are culturally sensitive, and are sensitive to change (Locke & Prinz, 2002; Smith, 2011). Finally, if evidence-based assessments are to be adopted widely, they must be available at low cost (ideally free), brief (to minimize respondent burden and to enhance compliance), and validated across relevant populations including ethnic minority groups and low SES populations (Beidas et al., 2015; Jensen-Doss & Hawley, 2010).

The Current Study

To address these challenges and needs, we developed a brief online measure of parent management training (PMT) skills that has strong psychometric properties, including national norms, and is available at no cost. The Knowledge of Effective Parenting Test (KEPT) is a measure of parenting knowledge specific to conduct problems for parents/guardians of children aged 5–12 years, intended for researchers conducting trials of

parenting interventions. Respondents answer multiple-choice questions with four options in response to video vignettes of both common and challenging parenting scenarios. The assessment covers content including age-appropriate expectations, effective communication, planned-ignoring, rewards, consequences, time-outs, attending to positive behavior, and praise. These domains were selected because of their empirical association with childhood externalizing behavior and their representation in the 16 psychosocial evidence-based treatments (EBTs) currently available for children and adolescents with disruptive behaviors (Eyberg et al., 2008). By developing a measure that uses a common elements approach (Chorpita, Daleiden, & Weisz, 2005; Garland, Hawley, Brookman-Frazee, & Hurlburt, 2008) and that is also “free, brief, and validated” (Beidas et al., 2015), our goal was to maximize the likelihood of successful dissemination and adoption by developmental researchers and developers of early interventions. Finally, the use of an item response theory (IRT) framework has allowed us to select items from our initial item pool that are the most informative with regard to knowledge of effective parenting and that are free from significant differential item functioning (DIF). The lack of DIF ensures a measure of knowledge of parenting skills appropriate for racially and ethnically diverse samples, without racial bias in the meaning of item scores. The availability of measures that function equivalently across diverse groups is increasingly essential, as the fastest growing population of racial and ethnic minorities in the United States is children under the age of 18. The U.S. Census Bureau projects that in the next two-to-three years slightly less than 50% of children will identify as non-Hispanic White (Vespa, Armstrong, & Medina, 2018).

Method

Overview

The project was approved by the Institutional Review Board at the University of Pittsburgh. The Knowledge of Effective Parenting Test (KEPT) was developed in several stages. In the first stage, an item pool of 36 multiple-choice questions was developed to measure skills typically taught during parent management training (PMT). In the second stage, all 36 items were administered online to a national sample ($N = 1,570$) of parents/guardians of children aged 5–12 years, along with several other measures used to establish convergent and divergent validity. In the third stage, IRT analyses were performed to select items with strong psychometric properties and without significant DIF by race/ethnicity for the final version of the KEPT. In this stage, we also created a brief 10-item version of the KEPT to enhance its utility as a research tool for both longitudinal and intervention research. In the fourth and final stage, we examined the reliability and validity of the full 21-item version (KEPT-Full) and brief 10-item version (KEPT-Brief). Measures of related parenting constructs as well as child and parent psychopathology were used to examine convergent and divergent validity. We also estimated national norms and percentiles.

Item Development (Stage 1)

Items were developed through an iterative process that began by reviewing the “common elements” literature (Chorpita et al., 2005; Garland et al., 2008) to identify parenting domains that are consistently represented in empirically based treatments for disruptive behavior problems. We started by reviewing the major treatments for disruptive behavior

disorders including Parent Management Training – Oregon Model (PMTO; Bullard et al., 2010), Parenting Through Change (Martinez & Forgatch, 2001), Fast Track (Conduct Problems Prevention Research Group, 2004), Parent Child Interaction Therapy (PCIT; Funderburk & Eyberg, 2011), Incredible Years (Webster-Stratton & Reid, 2003), Problem-Solving Skills Training (Kazdin, 2003), Positive Parenting Program (Triple P; Bor, Sanders, & Markie-Dadds, 2002), and Keeping Foster Parents Trained and Supported (KEEP; Leathers, Spielfogel, McMeel, & Atkins, 2011). Using this approach, six core domains of parenting skills (praise, rewards/point system, attending and ignoring, commands/expectations, consequences, and time-outs) were identified (see Table 1).

Next, multiple-choice items were drafted to assess each of these six domains. A subset of the questions (items 15–36) was based on video vignettes depicting common parenting scenarios (e.g., arguing, interrupting, non-compliance). Eight of the vignettes were videotaped to create audio-visual clips that are imbedded in the online measure. Professional actors of ethnically diverse backgrounds were hired from the standardized patient program at the University of Pittsburgh Medical Center (UPMC) to perform the vignettes. Five child actors (2 boys; 3 girls) ranged in age from 7 to 12.

To ensure that the items had strong face validity and were relevant across the major treatments for disruptive behavior disorders (DBD), we used a modified Delphi technique to solicit feedback on the item pool from key DBD treatment experts. We identified a group of 20 researchers involved in the development of one or more of nine evidence-based protocols for the treatment or prevention of childhood behavior problems (PMTO, Parenting Through Change, Fast Track, PCIT, Incredible Years, Problem-Solving Skills Training, Triple P, and KEEP). These 20 experts were contacted via email to provide feedback on a subset of 14 items. Items were presented in an online survey format and survey links were emailed to the experts, who were asked to indicate whether they agreed or disagreed on the “correct or best” answer indicated for each item. If they disagreed, they were given the option to suggest how to improve the item. The experts also provided general feedback on the entire assessment. We sent four email requests at weekly intervals to each of the 20 experts, generating a response rate of 65% (13/20). The respondents included at least one expert from each of the nine evidence-based protocols for the treatment or prevention of childhood behavior problems. Only two items had less than 80% expert agreement on the “correct or best” answer, indicating good consensus. See Figure 1 for a sample screen shot and item from the KEPT.

Data Collection (Stage 2)

Procedure.—We contracted with the survey company YouGov to administer all 36 items (see Supplemental Material), along with additional measures of parenting, child behavior, and parent and child psychopathology, to a national sample of parents/guardians of children aged 5–12 years that was representative of the U.S. population on key demographic variables. YouGov is a large survey company that uses a panel of over 1.2 million U.S. residents to conduct online surveys and assessments. Participants on the YouGov panels have been recruited through web advertising, permission-based email campaigns, partner contacts, random digit dialing, and mail (based on voter registration). YouGov has also

obtained a large amount of demographic information from the panelists in previous studies. Links to the survey were emailed to YouGov panelists. Recent studies have found that 84% of Americans use the internet (Perrin & Duggan, 2015), 79% have regular access to internet in the home (File & Ryan, 2014), and internet samples reasonably represent the overall population (Hays, Liu, & Kapteyn, 2015). Participants reached via these recruitment strategies have opted-in and agreed to be contacted to complete surveys for which they are relevant. The response rate to the emailed survey link was 52.8%. Of those who initiated the survey, 46.2% met eligibility criteria. Participants received \$25 in compensation for completing the full set of questions.

Participants.—The final sample included 1,570 parents/guardians of children aged 5–12 from all 50 states and the District of Columbia. Key demographic variables are summarized in Table 2.

Measures.—Several additional measures of related parenting constructs and child and parent psychopathology were administered to examine convergent and divergent validity of the KEPT-Full and KEPT-Brief. Constructs included child behavior problems, child depression, child anxiety, child trauma, parent discipline strategies, parent psychopathology (depression and anxiety), and parent knowledge of behavioral principles.

Alabama Parenting Questionnaire (APQ): The APQ is a 42-item measure that assesses five dimensions of parenting: (1) positive involvement, (2) supervision and monitoring, (3) use of positive discipline techniques, (4) consistency in the use of such discipline, and (5) use of corporal punishment, using a 5-point scale ranging from 1 (never) to 5 (always). The internal consistency of the scale has been reported to be acceptable with alphas for the five domains ranging from .46 to .80. The measure has well-established construct validity (Shelton, Frick, & Wooten, 1996).

Knowledge of Behavior Principles (KOBP-10): The KOBP-10 is an abbreviated, 10-item version of the Knowledge of Behavior Principles measure. Questions are presented in a multiple-choice format. Short forms of the KOBP have Cronbach's alphas ranging from .66 to .89 and strong content validity (Furtkamp, Giffort, & Schiers, 1982; Sturmey, Newton, Milne, & Burdett, 1987).

Vanderbilt Assessment Scale-Parent Report (VAS-P): The VAS-P is a 55-item parent-report screen for Attention-Deficit/Hyperactivity Disorder (ADHD), Oppositional Defiant Disorder (ODD), and Conduct Disorder (CD). It also includes seven items on internalizing symptoms and eight items on school performance and social functioning. Symptom items are rated using a 4-point scale (“never” to “very often”) and the performance items are rated on a 5-point scale from “problematic” to “above average.” The measure has Cronbach's alphas ranging from .79 to .95 and strong evidence of construct validity (Wolraich et al., 2003).

Screen for Child Anxiety Related Disorders (SCARED): The SCARED is a 41-item measure of general anxiety, separation anxiety, social phobia, and school phobia. Items are rated using a 3-point Likert scale. The internal consistency of the SCARED is good with

Cronbach's alphas ranging from .74 to .93. The measure has strong convergent and divergent validity (Birmaher et al., 1997).

Patient Health Questionnaire depression scale (PHQ-8). The PHQ-8 measures symptoms of depression using a 4-point scale from “not at all” to “nearly every day.” Total scores range from 0 to 24. The measure has a reported Cronbach's alpha of .82 and evidence of construct validity (Pressler et al., 2011).

Generalized Anxiety Disorder Screener (GAD-7). The GAD-7 is a 7-item measure of anxiety. Items are rated on a 4-point scale from “not at all” to “nearly every day.” The measure includes an item to assess the duration of anxiety symptoms. The measure has excellent internal consistency (Cronbach's alpha = 0.92), good test-retest reliability (intraclass correlation = 0.83), and strong convergent validity with other measures of anxiety (Spitzer, Kroenke, Williams, & Löwe, 2006).

Item Selection / IRT Analyses (Stage 3)

We began by computing item-total correlations for each item. Items with item-total correlations above .30 were analyzed further to ensure unidimensionality of items before initiating the IRT analyses. Unidimensionality was tested by examining the ratio of first to second eigenvalues from an exploratory factor analysis (Hawes et al., 2014; Morizot, Ainsworth, & Reise, 2007). One common threshold used to support unidimensionality is a ratio of ≤ 3 . Two additional goodness of fit indices were evaluated. The Very Simple Structure criterion (VSS) has been shown to identify more consistently the optimal number of factors to extract than maximum likelihood tests and Kaiser's eigenvalue cut-off of 1.0. The VSS evaluation is a degradation of the factor solution, testing how well the factor matrix fits the correlation matrix. The maximum value achieved represents the ideal number of factors to extract. The Velicer Minimum Average Partial (MAP) test characterizes the squared, average partial correlations among the scale items after removing the effect of the factors.

IRT analyses were conducted using IRTPRO (Cai, Toit, & Thissen, 2011) and the R statistical environment (R Core Team, 2018). Local independence of items was tested by examining the correlations of residuals (these should be close to zero), after accounting for the latent trait score. We generated item characteristic curves (ICCs), item parameters, and indices of differential item functioning (DIF) for each item. We used a three-parameter logistic (3PL) IRT model to estimate the discrimination (α), difficulty (β), and guessing (γ) parameters for each item. The discrimination (α) parameter is the slope of the item characteristic curve (ICC) at β and measures how well the item discriminates between participants whose latent trait ability level (θ) is above the item's β parameter and those whose latent trait ability level (θ) is below the item's β parameter. Higher α parameters indicate better discrimination. The difficulty parameter (β) is defined as the latent trait level at which a participant has a .50 chance of answering the item correctly. The guessing (γ) parameter estimates the lower asymptote of the ICC and is appropriate when participants have a reasonable chance of guessing the correct answer. Because most of the items consist

of multiple-choice questions with four options (and only one correct answer), we expected each γ parameter to be approximately .25.

Differential item functioning (DIF).—We conducted DIF analyses to determine whether any of the items had significantly different ICCs and model parameters across our largest racial/ethnic subgroups. DIF occurs when an item performs differently for individuals from distinct groups at equivalent levels of the latent trait (when group membership is not related to measurement conditions). Tests for DIF that produce null results suggest the item is invariant across groups. DIF was examined across racial/ethnic groups: Hispanic or Latino, Black (non-Hispanic), and White (non-Hispanic), the three largest racial/ethnic groups of children in the United States currently (Vespa et al., 2018). Given the lack of consensus on the optimal form of DIF detection, two procedures were used and compared to identify whether items displayed DIF with both methods. Logistic Regression methods of DIF detection are based on fitting logistic regression models with a criterion such as the total scale score (Model 1), group membership in comparison to a reference group (Model 2), and the interaction between scale score and group membership (Model 3). Then, χ^2 statistics for each model are compared to detect DIF and determine the size of the DIF effect. A logistic regression approach to DIF detection was implemented in the R environment using the Likelihood Ratio test (Thissen, Steinberg, & Wainer, 1993). Pseudo R^2 values were calculated to determine effect size, representing the gain in log-likelihood from the explanatory variables in the model. Another well-established method of DIF detection is the Mantel-Haenszel (MH) χ^2 (Holland & Thayer, 1988; Mantel & Haenszel, 1959). The MH method involves matching respondents on total scale score and using a series of contingency tables for each item to make pair-wise comparisons between group-membership categories based on item performance. The MH method was also employed in the R software package. Items that demonstrated DIF across the two methods or that displayed a large DIF effect size were eliminated from the final measure.

Item selection.—Results from the IRT analyses were used to select items for the final version of the KEPT as well as a brief (10-item) version of the KEPT. Criteria for selecting the 10 “best” items for the brief version were: 1) items with high discrimination (α) parameters, and 2) items without significant DIF by racial/ethnic categories. See Figure 2 for a flowchart depicting the item selection process.

Psychometric Characteristics and Norms (Stage 4)

Reliability and validity.—We computed internal consistency (Cronbach’s α) for the KEPT-Full (21 items) and KEPT-Brief (10 items). We also examined convergent and divergent validity by computing correlations with related constructs. We expected higher correlations with more closely related constructs (e.g., knowledge of behavior principles and discipline practices) than with more distinct constructs (e.g., parent psychopathology).

National norms and percentiles.—We estimated national norms (M s, SD s, and percentiles) for the KEPT-Full (21 items) and KEPT-Brief (10 items) based on the full sample ($N = 1,570$). Norms were also estimated for the “clinical” population described

above (parents/guardians of children who screened positive for ODD and CD on the Vanderbilt Assessment Scale-Parent Version).

Results

Preliminary Analyses

Item-total correlations for the initial 36 items ranged from .09 to .59 ($M = .39$; $SD = .11$). Seven items with correlations below .30 were removed prior to subsequent analyses. The dimensionality of the remaining 29 items was examined using exploratory factor analyses (EFA) with maximum-likelihood (ML) estimation. The ratio of the first eigenvalue (8.64) to the second eigenvalue (.87) was 9.92, far exceeding the recommended cut-off of 3.0 for documenting unidimensionality. A unidimensional structure was also supported by additional goodness-of-fit indices. The MAP minimum value of .01 was reached with the one-factor solution, and the VSS maximum value of .62 was also reached with a one-factor solution. Finally, the correlations of the residuals for each item were all less than .10, indicating sufficient local item independence.

IRT Analyses

Model fit.—Table 3 presents the item-fit statistics for the 3PL models. Item-fit statistics indicated that the 3PL models fit the data well, with large Comparative Fit Index (CFI) and Tucker Lewis Index (TLI) coefficients. Separate models were fit for respondents identifying with each of the three racial/ethnic groups, White non-Hispanic, Black non-Hispanic, and Hispanic, prior to the DIF tests, and these separate models also had excellent fit.

Differential item functioning (DIF).—Six items were removed for evidence of DIF across racial and/or ethnic categories. Items were removed if they showed significant DIF using multiple methods, large effect sizes for DIF, or both. All six items that were removed demonstrated DIF for non-Hispanic Black respondents compared to non-Hispanic White respondents. Four of the six items also demonstrated DIF between non-Hispanic Black respondents and Hispanic respondents.

Item selection.—After the DIF analyses, two additional items were removed for having discrimination parameters below 1.0 in the full sample. The remaining 21 items comprised the final KEPT measure. Ten items with the strongest psychometric properties (highest discrimination parameters) were selected for the KEPT-Brief. Table 4 presents factor loadings for the full and brief versions of the KEPT. Table 5 summarizes the item parameters for each of the items in the KEPT-Full and KEPT-Brief. All discrimination (α) parameters were above 1.0, ranging from 1.05 to 2.20 ($M = 1.58$; $SD = .39$). Difficulty (β) parameters ranged from -1.92 to 1.51 ($M = -.30$; $SD = .69$). Guessing (γ) parameters ranged from .09 to .29 ($M = .17$; $SD = .06$). See Table 1 for the item content for the KEPT-Full and KEPT-Brief. Figures 3 and 4 show the test information curves for the KEPT-Full and KEPT-Brief along with the standard errors of measurement. As would be expected, the KEPT-Full affords greater precision of measurement with standard errors ranging from .33 to .52 for typical levels of theta (-1.5 to 1.5) compared to .40 to .74 for the KEPT-Brief.

Reliability and Validity

Cronbach's alphas were .84 for the KEPT-Full and .76 for KEPT-Brief. Table 6 summarizes correlations between the KEPT-Full (21 items) and KEPT-Brief (10 items) and other measures of child behavior and child and parent psychopathology. As expected, the largest correlations were with other measures of parenting. Correlations were largest between the KEPT-Full and corporal punishment ($r = -.42$), poor monitoring/supervision ($r = -.44$), and knowledge of behavioral principles ($r = .44$). In contrast, the correlations with the other two APQ scales (involvement and positive parenting) were low and non-significant. Within the domain of child behavior/psychopathology, the KEPT-Full had the highest correlation with conduct problems ($r = -.31$). The correlation with oppositional and defiant behaviors was small ($r = -.02$) and nonsignificant. Scores on the KEPT-Full were not correlated with measures of parent depression ($r = .00$) or parent anxiety ($r = .00$). These converging and diverging patterns of correlations support the validity of the KEPT as a measure of knowledge of parenting skills with particular relevance to conduct problems. For the KEPT-Full, children whose parents/guardians scored below the 25th percentile were twice as likely (risk ratio = 2.0) to screen positive for ODD and more than eight times as likely (risk ratio = 8.2) to screen positive for CD compared to children whose parents/guardians scored above the 25th percentile. For the KEPT-Brief, these risk ratios were 2.0 for ODD and 10.7 for CD.

National Norms and Percentiles

Table 7 summarizes national norms for the KEPT-Full and KEPT-Brief. Scores did not differ by gender of the parent/guardian for either the KEPT-Full, $F(1,1568) = .01, p = .92$, or the KEPT-Brief, $F(1,1568) = .17, p = .68$. Means and standard deviations are also reported for parents/guardians of children who screened positive for ODD ($N = 192$; 12.2% of the full sample) and CD ($N = 82$; 5.2% of the full sample). As expected, mean scores were higher for the general population than for parents/guardians of children who screened positive for ODD and CD. Differences were most pronounced for CD, for which the mean score on the KEPT-Full was more than a full standard deviation lower for parents/guardians of children who screened positive for CD compared to those who screened negative, $F(1, 1568) = 108.31, p < .00001$. This difference represents a very large effect, Cohen's $d = 1.14$. For ODD, the difference was smaller but still significant. Parents/guardians of children who screened positive for ODD scored about one-third standard deviation lower ($d = .36$) on the KEPT-Full compared to those who screened negative, $F(1, 1568) = 21.65, p < .00001$. Percentiles for the KEPT-Full and KEPT-Brief are summarized in Table 8.

Discussion

The KEPT is unique among parenting measures because it was designed specifically as an assessment of a potential treatment target for evidence-based psychosocial treatments from a common elements perspective. Content was selected by reviewing the common elements literature on disruptive behavior disorders (DBDs) treatments and soliciting feedback from domain experts on pilot test items. The final item pool was tested rigorously using factor analyses and models from item response theory (IRT) to ensure that the final measure assessed a single construct and only contained items with strong psychometric properties. Each item was tested for differential item functioning (DIF) to ensure that the measure was

free from racial or ethnic biases for the major U.S. census categories (White, African American or Black, Hispanic). The result is a measure of knowledge of parenting skills (treatment targets common to evidence-based psychosocial treatments for DBDs) with excellent reliability and validity. Most importantly, scores on the KEPT differentiate between parents/guardians of children who screen positive versus negative for ODD and CD. The difference in scores is particularly large for CD. The mean score on the KEPT-Full was more than a full standard deviation lower for parents/guardians of children who screened positive for CD compared to those who screened negative, a very large effect (Cohen's $d = 1.14$). The KEPT was also designed to reduce social desirability bias. Because the measure is a test of knowledge of parenting skills with empirically derived correct and incorrect answers, it is difficult for respondents to inflate their score.

We also developed a brief version of the KEPT (10 items) to enhance its clinical utility. Assessment measures are more likely to be adopted if they are free, brief, and validated (Beidas et al., 2015). In developing the KEPT, we ensured that the measure met these three key criteria. The KEPT is freely available from the developers by request, can be administered in just a few minutes, and has strong psychometric properties. We acknowledge that the KEPT-Full has stronger indices of reliability and validity than the KEPT-Brief, and we recommend using the full version if time allows.

The KEPT as a Measure of Target Engagement in Treatment Trials

The KEPT is well suited to measure knowledge of parenting skills relevant to conduct problems as a treatment target in the context of clinical trials because it was designed to eliminate the problem of direction of causality that confounds so many measures of parenting. Because many measures of parenting ask respondents about their own use of discipline strategies and parenting behaviors, it is challenging to establish whether changes in parenting behaviors are, in fact, contributing to the reduction of behavior problems or vice versa (Lindhiem, Shaffer, & Kolko, 2014; Shaffer, Lindhiem, & Kolko, 2016). It is equally plausible that a reduction in behavior problems is contributing to changes in parenting behaviors. Because the KEPT measures parenting knowledge, however, it is far less plausible that reductions in child behavior problems could boost a caregiver's score on the KEPT.

Strengths and Limitations

This study has numerous strengths including the systematic development of an item pool from a common elements approach, the solicitation of feedback on test items from a large cross-section of content experts, administration of test items to a large nationally representative sample, rigorous analyses of item characteristics and DIF of every item using models from IRT, the development of both a long form and a short form, the creation of national norms and percentiles, and evidence of strong reliability and validity. Limitations include a small number of items for each domain and the reliance on parents/guardians as the only informants for child behavior. This limitation is somewhat mitigated by substantial evidence that parents/guardians are accurate reporters of children's externalizing behaviors (e.g., De Los Reyes & Kazdin, 2005; Loeber, Green, Lahey, & Stouthamer-Loeber, 1989). If the correlations between the KEPT and other measures of parenting and child behavior were

due largely to shared method variance, then all correlations would be expected to be positive and convergent. However, this was not the case, as correlations with parent psychopathology, for example, were zero. Rather, the patterns of converging and diverging correlations between the KEPT and the other measures support the validity of the KEPT as a distinctive measure of knowledge of parenting skills. Another limitation is that feasibility considerations constrained our ability to collect observational data on parenting behaviors. This is somewhat less problematic than it would otherwise be, given that the KEPT is a measure of parenting knowledge versus a self-report measure of actual parenting behaviors. Finally, we were surprised at the non-significant correlation between the KEPT and the ODD subscale on the Vanderbilt Assessment Scale. This might suggest that deficits in PMT knowledge are more of a risk-factor for CD than for ODD, and certainly requires further investigation.

Future Directions

The KEPT is intended as a measure of parenting knowledge specific to conduct problems for use in treatment outcome studies. Before it can be adopted, however, the KEPT needs additional research evaluating its construct validity, test-retest reliability, and sensitivity to intervention effects. A reliable and valid measure of PMT knowledge will allow researchers to test whether instances of treatment failure are due to failure of target engagement (i.e., failure to understand and acquire PMT skills). This is a critical question because many families fail to respond to treatment, typically ranging from 20–60% depending on sample characteristics (age, diagnosis, severity, comorbid conditions, etc.) and definitions of nonresponse. The ability to test whether instances of treatment failure are due to failure of target engagement will help to refine existing treatments and lead to better outcomes.

It is important to note that the KEPT is not intended as a standalone measure but as a supplement to measures of actual parenting behaviors. Skill knowledge and skill use are conceptually related and understanding the variables that influence the translation of knowledge into practice is an important next step. There is meta-analytic evidence, for example, that parenting-based interventions are more effective in teaching parents new skills (i.e., knowledge acquisition) than in changing actual parenting behaviors (Lindhiem, et al., 2014). Measures of actual skill use with strong psychometric properties are needed to investigate the distinct antecedents and correlates of behavior change in parenting as well as techniques to enhance the utilization of effective parenting strategies.

Finally, another next step will be to develop a computerized adaptive testing (CAT) version of the KEPT using the 21 items from the KEPT-Full as an item bank. This could enhance the utility of the KEPT for repeated measurements in the context of clinical trials and longitudinal studies. A CAT version of the KEPT would administer items until a predetermined standard error of measurement was reached. In general, CAT achieves adequate precision in measurement with fewer items than static tests (often with as few as 4 to 6 items). A CAT option would mean an even briefer measure than the static 10-item KEPT-Brief with comparable measurement precision.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This study was supported by grants to the first author from the National Institute of Mental Health (K01MH093508) and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R21HD090145). Janelle Higa is now at S. W. King Intermediate School in Kaneohe, HI. The authors would like to thank Karen Bierman, Stephen Boggs, David DeGarmo, Sheila Eyberg, Marion Forgatch, John Lochman, Robert McMahon, Cheryl McNeil, and Matthew Sanders for providing feedback on preliminary test items.

References

- Aquilino WS, & Supple AJ (2001). Long-term effects of parenting practices during adolescence on well-being: Outcomes in young adulthood. *Journal of Family Issues*, 22(3), 289–308. doi: 10.1177/019251301022003002
- Beidas RS, Stewart RE, Walsh L, Lucas S, Downey MM, Jackson K, ... Mandell DS (2015). Free, brief, and validated: standardized instruments for low-resource mental health settings. *Cognitive and Behavioral Practice*, 22(1), 5–19. doi:10.1016/j.cbpra.2014.02.002 [PubMed: 25642130]
- Ben-Ami N, & Baker AJ (2012). The long-term correlates of childhood exposure to parental alienation on adult self-sufficiency and well-being. *The American Journal of Family Therapy*, 40(2), 169–183. doi:10.1080/01926187.2011.601206
- Birmaher B, Khetarpal S, Brent D, Cully M, Balach L, Kaufman J, & Neer SM (1997). The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale construction and psychometric characteristics. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36(4), 545–553. doi:10.1097/00004583-199704000-00018 [PubMed: 9100430]
- Bor W, Sanders MR, & Markie-Dadds C (2002). The effects of the Triple P-positive Parenting Program on preschool children with co-occurring disruptive behavior and attentional/hyperactive difficulties. *Journal of Abnormal Child Psychology*, 30(6), 571–587. [PubMed: 12481972]
- Bullard L, Wachlarowicz M, DeLeeuw J, Snyder J, Low S, Forgatch MS, & DeGarmo DS (2010). Effects of the Oregon model of Parent Management Training (PMTO) on marital adjustment in new stepfamilies: A randomized trial. *Journal of Family Psychology*, 24(2), 485–496. [PubMed: 20731495]
- Cai L, Toit. d. CSH, & Thissen D (2011). IRTPRO: Flexible, multidimensional, multiple categorical IRT modeling. Chicago, IL: Scientific Software International.
- Chorpita BF, Daleiden EL, & Weisz JR (2005). Identifying and selecting the common elements of evidence based interventions: A distillation and matching model. *Mental health services research*, 7(1), 5–20. [PubMed: 15832690]
- Conduct Problems Prevention Research Group. (2004). The effects of the Fast Track Program on serious problem outcomes at the end of elementary school. *Journal of Clinical Child and Adolescent Psychology*, 33(4), 650–661. [PubMed: 15498733]
- De Los Reyes A, & Kazdin AE (2005). Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, 131(4), 483–509. doi: 10.1037/0033-2909.131.4.483 [PubMed: 16060799]
- Eddy JM (2017). Facing a Fundamental Problem in Prevention Science: the Measurement of a Key Construct. *Prevention Science*, 18(3), 322–325. [PubMed: 28091964]
- Eyberg SM, Nelson MM, & Boggs SR (2008). Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. *Journal of Clinical Child & Adolescent Psychology*, 37(1), 215–237. [PubMed: 18444059]
- File T, & Ryan C (2014). Computer and internet use in the United States: 2013. Washington, DC: United States Census Bureau.

- Funderburk B, & Eyberg SM (2011). Parent-child interaction therapy In Norcross JC, VandenBos GR, & Freedheim DK (Eds.), *History of psychotherapy: Continuity and change* (2nd ed., pp. 415–420). Washington, DC, US: American Psychological Association.
- Furtkamp E, Gifford D, & Schiers W (1982). In-class evaluation of behavior modification knowledge: Parallel tests for use in applied settings. *Journal of Behavior Therapy & Experimental Psychiatry*, 13(2), 131–134. [PubMed: 7130408]
- Garland AF, Hawley KM, Brookman-Frazee L, & Hurlburt M (2008). Identifying common elements of evidence-based psychosocial treatments for children's disruptive behavior problems. *Journal of the American Academy for Child and Adolescent Psychiatry*, 47, 505–514.
- Gryczowski MR, Jordan SS, & Mercer SH (2010). Differential relations between mothers' and fathers' parenting practices and child externalizing behavior. *Journal of Child and Family Studies*, 19(5), 539–546. doi:10.1007/s10826-009-9326-2
- Hawes SW, Byrd AL, Henderson CE, Gazda RL, Burke JD, Loeber R, & Pardini DA (2014). Refining the parent-reported Inventory of Callous-Unemotional Traits in boys with conduct problems. *Psychological Assessment*, 26, 256–266. doi:10.1037/a0034718 [PubMed: 24188153]
- Hays RD, Liu H, & Kapteyn A (2015). Use of internet panels to conduct surveys. *Behavior Research Methods*, 47(3), 685–690. doi:10.3758/s13428-015-0617-9 [PubMed: 26170052]
- Holland PW, & Thayer DT (1988). Differential item performance and the Mantel-Haenszel procedure In Wainer H & Braun HI (Eds.), *Test validity* (pp. 129–145). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Huey SJ, Henggeler SW, Brondino MJ, & Pickrel SG (2000). Mechanisms of change in multisystemic therapy: Reducing delinquent behavior through therapist adherence and improved family and peer functioning. *Journal of Consulting & Clinical Psychology*, 68(3), 451–467. [PubMed: 10883562]
- Hurley KD, Huscroft-D'Angelo J, Trout A, Griffith A, & Epstein M (2014). Assessing parenting skills and attitudes: A review of the psychometrics of parenting measures. *Journal of Child and Family Studies*, 23, 812–823. doi:10.1007/s10826-013-9733-2
- Jensen-Doss A, & Hawley KM (2010). Understanding barriers to evidence-based assessment: Clinician attitudes toward standardized assessment tools. *Journal of Clinical Child & Adolescent Psychology*, 39(6), 885–896. doi:10.1080/15374416.2010.517169 [PubMed: 21058134]
- Kaiser NM, McBurnett K, & Piffner LJ (2011). Child ADHD severity and positive and negative parenting as predictors of child social functioning: Evaluation of three theoretical models. *Journal of attention disorders*, 15(3), 193–203. doi:10.1177/1087054709356171 [PubMed: 20424006]
- Kazdin AE (2003). Problem-solving skills training and parent management training for conduct disorder In Kazdin AE & Weisz JR (Eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 241–262). New York: Guilford.
- Leathers SJ, Spielfogel JE, McMeel LS, & Atkins MS (2011). Use of a parent management training intervention with urban foster parents: A pilot study. *Children and Youth Services Review*, 33(7), 1270–1279. doi:10.1016/j.childyouth.2011.02.022 [PubMed: 21686093]
- Lindhiem O, Higa J, Trentacosta CJ, Herschell AD, & Kolko DJ (2014). Skill acquisition and utilization during evidence-based psychosocial treatments for childhood disruptive behavior problems: A review and meta-analysis. *Clinical Child and Family Psychology Review*, 17(1), 41–66. doi:10.1007/s10567-013-0136-0 [PubMed: 23649324]
- Lindhiem O, & Shaffer A (2017). Introduction to the Special Series: Current Directions for Measuring Parenting Constructs to Inform Prevention Science. *Prevention Science*, 18(3), 253–256. doi: 10.1007/s11121-016-0724-6 [PubMed: 27834033]
- Lindhiem O, Shaffer A, & Kolko DJ (2014). Quantifying discipline practices using absolute vs. relative frequencies: Clinical and research implications for child welfare. *Journal of Interpersonal Violence*, 29(1), 66–81. doi:10.1177/0886260513504650 [PubMed: 24106146]
- Locke LM, & Prinz RJ (2002). Measurement of parental discipline and nurturance. *Clinical Psychology Review*, 22(6), 895–929. doi:10.1016/S0272-7358(02)00133-2 [PubMed: 12214330]
- Loeber R, Green SM, Lahey BB, & Stouthamer-Loeber M (1989). Optimal informants on childhood disruptive behaviors. *Development and Psychopathology*, 1(4), 317–337.
- Mantel N, & Haenszel W (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the national cancer institute*, 22(4), 719–748. [PubMed: 13655060]

- Martinez CRJ, & Forgatch MS (2001). Preventing problems with boys' noncompliance: Effects of a parent training intervention for divorcing mothers. *Journal of Consulting and Clinical Psychology*, 69, 416–428. [PubMed: 11495171]
- Morizot J, Ainsworth AT, & Reise SP (2007). Toward modern psychometrics: Application of item response theory models in personality research In Robins RW, Fraley RC, & Krueger RF (Eds.), *Handbook of Research Methods in Personality Psychology* (pp. 407–421). New York, NY: Guilford Press.
- Morris AS, Silk JS, Steinberg L, Myers SS, & Robinson LR (2007). The role of the family context in the development of emotion regulation. *Social Development*, 16(2), 361–388. doi:10.1111/j.1467-9507.2007.00389.x [PubMed: 19756175]
- Nock MK, & Ferriter C (2005). Parent Management of attendance and adherence in child and adolescent therapy: A conceptual and empirical review. *Clinical Child and Family Psychology Review*, 8(2), 149–166. doi:10.1007/s10567-005-4753-0 [PubMed: 15984084]
- Perrin A, & Duggan M (2015). *Americans' internet access: 2000–2015*. Washington, D.C.: Pew Research Center.
- Pressler SJ, Subramanian U, Perkins SM, Gradus-Pizlo I, Kareken D, Kim J, ... Sloan R (2011). Measuring depressive symptoms in heart failure: Validity and reliability of the Patient Health Questionnaire–8. *American Journal of Critical Care*, 20(2), 146–152. [PubMed: 20378777]
- R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing Retrieved from <http://www.r-project.org/>
- Shaffer A, Lindhiem O, & Kolko D (2016). Treatment Effects of a Primary Care Intervention on Parenting Behaviors: Sometimes It's Relative. *Prevention Science*, 18(3), 305–311. doi:10.1007/s11121-016-0689-5
- Shelton KK, Frick PJ, & Wooten J (1996). Assessment of parenting practices in families of elementary school-age children. *Journal of Clinical Child Psychology*, 25, 317–329.
- Smith M (2011). Measures for assessing parenting in research and practice. *Child and Adolescent Mental Health*, 16(3), 158–166. doi:10.1111/j.1475-3588.2010.00585.x
- Spitzer RL, Kroenke K, Williams JB, & Löwe B (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, 166(10), 1092–1097. [PubMed: 16717171]
- Sturme P, Newton T, Milne D, & Burdett C (1987). Parallel forms of the knowledge of behavioral principles as applied to children questionnaire: An independent, multi-centered, British replication. *Journal of Behavior Therapy & Experimental Psychiatry*, 18, 223–227. [PubMed: 3667950]
- Swanson J, Valiente C, Lemery-Chalfant K, & Caitlin O'Brien T (2011). Predicting early adolescents' academic achievement, social competence, and physical health from parenting, ego resilience, and engagement coping. *The Journal of Early Adolescence*, 31(4), 548–576. doi: 10.1177/0272431610366249
- Thissen D, Steinberg L, & Wainer H (1993). Detection of differential item functioning using the parameters of item response models In Holland PW & Wainer H (Eds.), *Differential Item Functioning* (pp. 67–113). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Vespa J, Armstrong DM, & Medina L (2018). *Demographic turning points for the United States: Population projections for 2020 to 2060*. Washington, DC: Curr Popul Rep. US Census Bureau.
- Waltz J, Addis ME, Koerner K, & Jacobson NS (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology*, 61, 620–630. [PubMed: 8370857]
- Webster-Stratton C, & Reid MJ (2003). The Incredible Years parents, teachers, and children training series: A multifaceted treatment approach for young children with conduct problems In Kazdin AE & Weisz JR (Eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 224–240). New York: Guilford.
- Wolraich ML, Lambert W, Doffing MA, Bickman L, Simmons T, & Worley KB (2003). Psychometric properties of the Vanderbilt ADHD diagnostic parent rating scale in a referred population. *Journal of Pediatric Psychology*, 28(8), 559–567.

Public Significance Statement:

The study describes a new online research tool that will allow for the remote assessment of parenting knowledge relevant to conduct problems. Results indicate strong psychometric properties with implications for parenting-based intervention trials and developmental research.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Please watch the clip below and answer the following questions.



29. What effective praise technique does the mother use at the beginning of the scene?

- A. Her praise is specific/labeled
- B. She speaks slowly
- C. She adds some constructive criticism
- D. She makes good eye contact when she praises

Figure 1.
Sample screen shot and item from the Knowledge of Effective Parenting Test (KEPT).

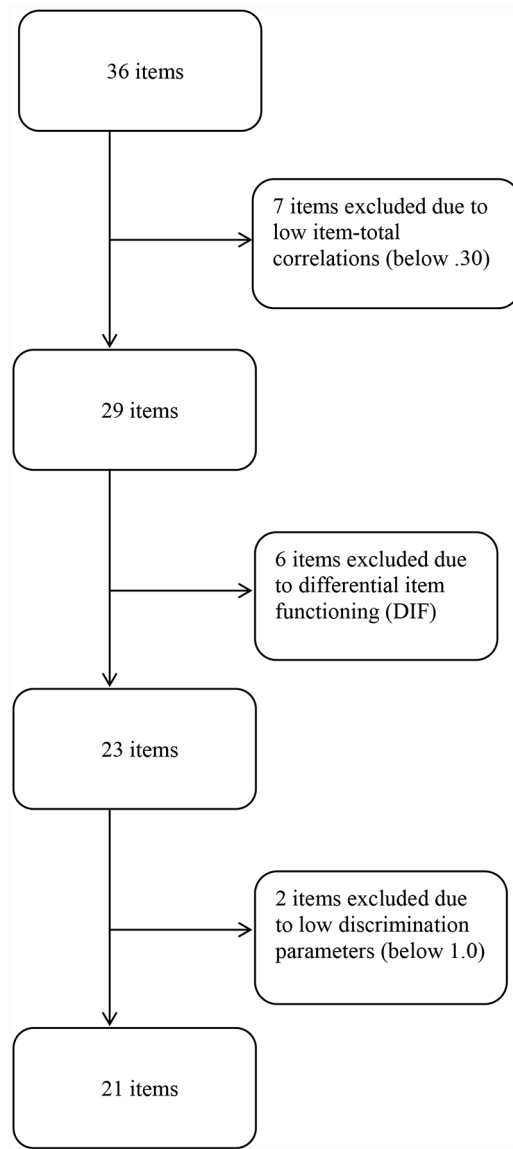


Figure 2.
Item selection process for the KEPT.

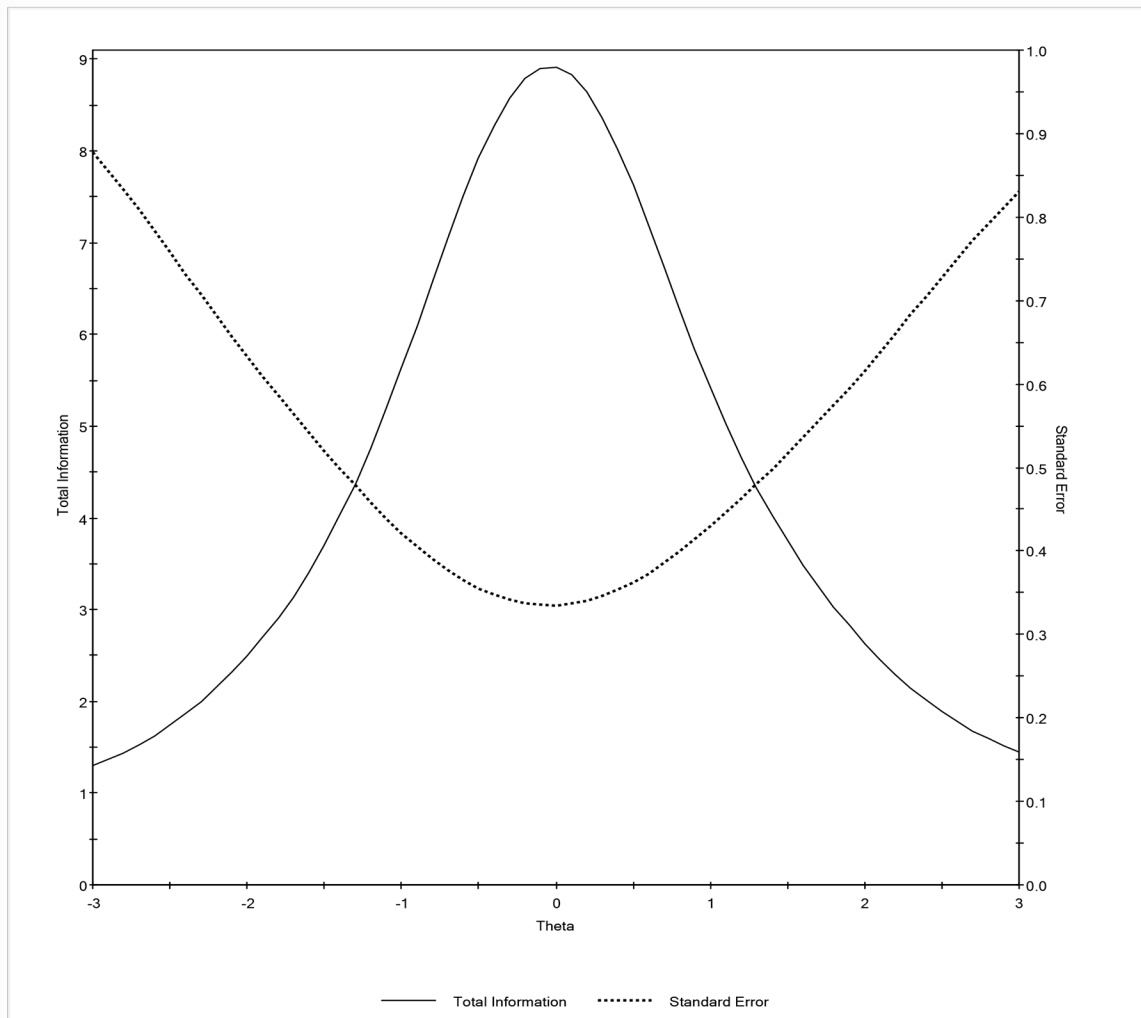


Figure 3.
Total test information curve for the KEPT-Full (21 items).

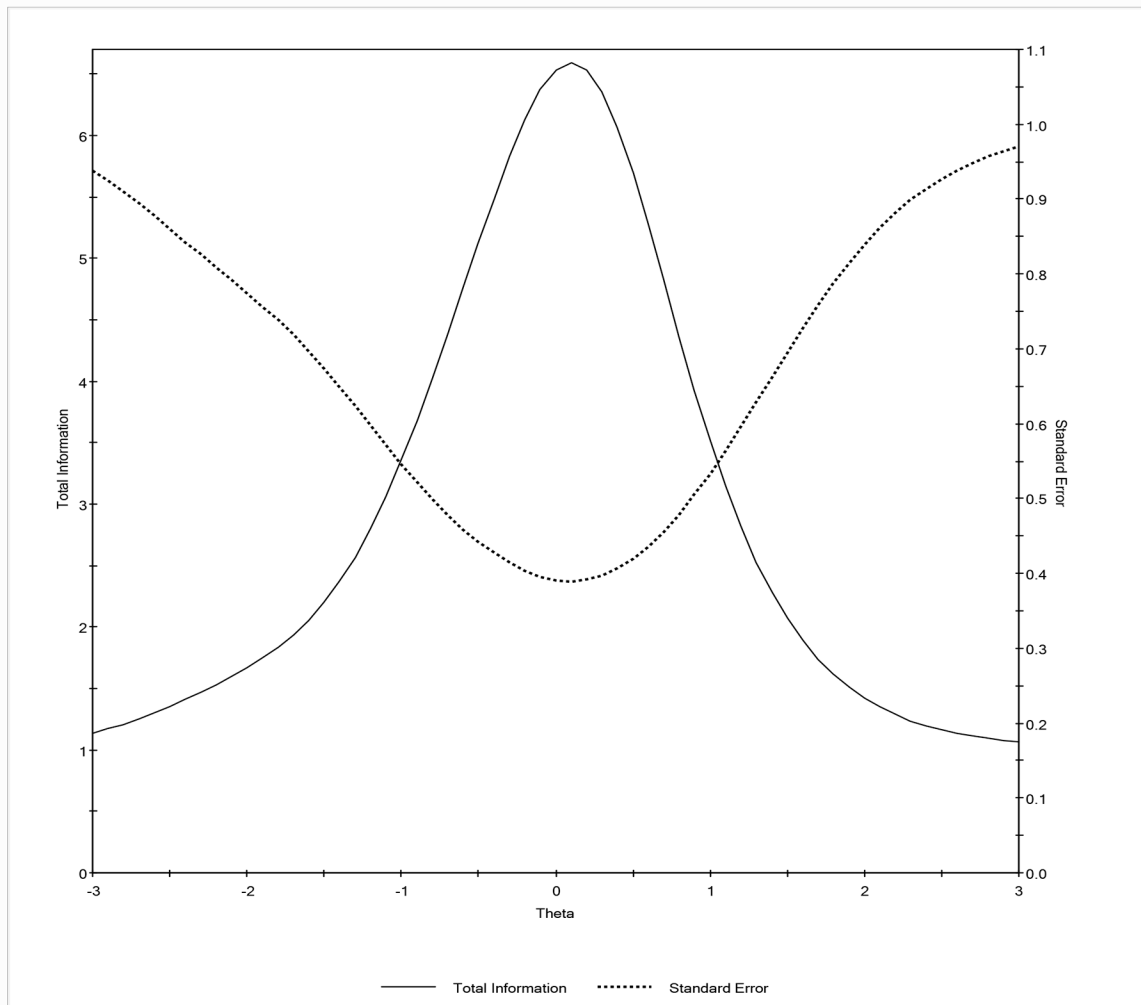


Figure 4.
Total test information curve for the KEPT-Brief (10 items).

Table 1

Content for the Item Pool (36 items), KEPT-Full (21 items), and KEPT-Brief (10 items)

	Item Pool (36 items)	KEPT-Full (21 items)	KEPT-Brief (10 items)
Praise	3 items	2 items	1 item
Rewards / Point System	7 items	6 items	2 items
Attending and Ignoring	4 items	4 items	1 item
Commands / Expectations	6 items	3 items	3 items
Consequences	5 items	3 items	2 items
Time-outs	3 items	-	-
Two or more skills	8 items	3 items	1 item

Table 2

Participant Demographic Characteristics (N = 1,570)

Parent/Guardian Age	$M = 41.5; SD = 10.6$
Parent/Guardian Gender	
Female	63.7%
Male	36.3%
Child Age	$M = 8.6; SD = 2.3$
Child Gender	
Female	46.8%
Male	53.2%
Parent/Guardian Race	
White	78.7%
Black or African American	13.9%
Asian	3.4%
American Indian or Alaska Native	2.6%
Native Hawaiian or Pacific Islander	.6%
Some other race	5.5%
Parent/Guardian Ethnicity	
Hispanic, Latino, or Spanish Origin	16.2%
Parent/Guardian Born in the United States	89.7%
Parent/Guardian Education	
High school or lower	25.6%
Some college or higher	74.4%
Marital Status (% married)	68.2%
Family Income	
Less than \$10,000	6.0%
\$10,000 – \$29,999	20.0%
\$30,000 – \$49,999	20.3%
\$50,000 – \$69,999	16.6%
\$70,000 and above	37.2%
U.S. Region	
Northeast	18.3%
Midwest	24.3%
South	37.0%
West	20.4%

Table 3

3PL Model Fit Indices

	3PL model		
	RMSEA	TLI	CFI
<u>KEPT-Full (21 items)</u>			
Total sample	.03	.97	.98
<i>Race/ethnicity</i>			
White (non-Hispanic)	.04	.96	.97
Black (non-Hispanic)	.04	.91	.93
Hispanic	.02	.99	.99
<u>KEPT-Brief (10 items)</u>			
Total sample	.03	.99	.99
<i>Race/ethnicity</i>			
White (non-Hispanic)	.03	.99	.99
Black (non-Hispanic)	.00	1.00	1.00
Hispanic	.00	1.00	1.00

Note. 3PL = 3 parameter logistic model; RMSEA = Root mean square error of approximation; CFI = Comparative fit index; TLI = Tucker Lewis Index.

Table 4

Factor Loadings Resulting from EFA

Item number	<i>KEPT-full</i>	<i>KEPT-brief</i>
2	0.50	--
6	0.63	0.60
7	0.64	--
8	0.59	0.60
9	0.62	--
10	0.37	--
12	0.71	0.70
13	0.65	--
14	0.45	--
16	0.63	0.63
17	0.52	--
18	0.57	0.60
20	0.57	0.59
25	0.60	0.59
28	0.66	0.65
29	0.66	0.66
30	0.51	--
31	0.49	--
32	0.48	--
33	0.72	0.69
34	0.48	--
Internal consistency (α)	$\alpha = .84$	$\alpha = .76$

Note. EFA=Exploratory factor analysis based on tetrachoric correlations. The seven initial items in the pool with low inter-item correlations were removed prior to computing the EFA.

Table 5

3PL Item Parameters for the KEPT-Full (21 items) and KEPT-Brief (10 items)

Item	Discrimination Parameter		Difficulty Parameter		Guessing Parameter	
	α	<i>s.e.</i>	β	<i>s.e.</i>	γ	<i>s.e.</i>
2	1.06	0.13	-0.65	0.24	0.16	0.09
6*	1.76	0.19	-1.92	0.18	0.18	0.09
7	1.57	0.17	-0.60	0.14	0.14	0.06
8*	1.61	0.22	-0.14	0.14	0.21	0.06
9	1.47	0.17	0.04	0.11	0.09	0.05
10	1.81	0.35	1.51	0.11	0.14	0.02
12*	2.10	0.23	0.25	0.06	0.09	0.03
13	1.67	0.17	-0.87	0.15	0.16	0.07
14	1.35	0.25	0.39	0.16	0.25	0.06
16*	1.61	0.17	-0.60	0.14	0.15	0.06
17	1.08	0.12	-0.68	0.21	0.13	0.08
18*	2.01	0.29	0.04	0.10	0.28	0.04
20*	1.99	0.30	0.18	0.10	0.25	0.04
25*	2.11	0.33	-0.06	0.11	0.29	0.05
28*	1.82	0.23	-0.60	0.15	0.20	0.07
29*	1.73	0.19	-0.29	0.11	0.13	0.05
30	1.06	0.12	-1.15	0.24	0.15	0.09
31	1.05	0.14	0.31	0.15	0.11	0.05
32	1.11	0.15	-0.25	0.20	0.19	0.07
33*	2.20	0.24	-0.42	0.08	0.14	0.05
34	1.05	0.13	-0.86	0.25	0.20	0.09

Note. The ten items marked with an asterisk are included in the KEPT-Brief.

Table 6
Correlations Between the KEPT-Full, KEPT-Brief, and Measures of Child Behavior, Parenting, and Parent and Child Psychopathology

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
<u>Parenting</u>																
1. KEPT-Full (21 items)	.84	.94**	-.42**	-.44**	-.19**	-.02	.02	.44**	-.31**	-.02	-.05*	.00	.04	-.18**	.00	.00
2. KEPT-Brief (10 items)	.76	-	-.39**	-.41**	-.17**	-.03	.00	.43**	-.29**	-.03	-.06*	-.01	.03	-.18**	.00	.00
3. Corporal Punishment (APQ subscale)	.80	.80	-	.59**	.44**	-.17**	-.14**	-.09**	.58**	.31**	.35**	.34**	.29**	.39**	.20**	.21**
4. Poor Monitoring (APQ subscale)	.85	.85	-	.54**	-.19**	-.05	-.05	-.08**	.58**	.28**	.26**	.31**	.31**	.36**	.18**	.18**
5. Inconsistent Discipline (APQ subscale)	.79	.79	-	-.21**	-.17**	-.02	-.02	-.02	.45**	.43**	.37**	.40**	.37**	.37**	.28**	.27**
6. Positive Parenting (APQ subscale)	.81	.81	-	.66**	-.14**	-.19**	-.13**	-.14**	-.23**	-.19**	-.13**	-.21**	-.18**	-.11**	-.08**	-.11**
7. Involvement (APQ subscale)	.79	.79	-	-.08**	-.16**	-.24**	-.16**	-.16**	-.16**	-.24**	-.16**	-.25**	-.15**	-.10**	-.16**	-.15**
8. Knowledge of Behavior Principles (KOBP)	.60	.60	-	.12**	.11**	.10**	.10**	.12**	.11**	.12**	.11**	.10**	.16**	.03	.10**	.12**
<u>Child Behavior / Psychopathology</u>																
9. CD Symptoms (VADPRS)	.96	.96	-	.59**	.49**	.48**	.61**	.61**	.49**	.59**	.49**	.48**	.61**	.61**	.40**	.41**
10. ODD Symptoms (VADPRS)	.91	.91	-	.72**	.69**	.69**	.60**	.60**	.72**	.69**	.69**	.69**	.60**	.51**	.44**	.47**
11. Hyperactivity / Impulsivity (VADPRS)	.91	.91	-	.74**	.48**	.48**	.48**	.48**	.74**	.48**	.48**	.48**	.48**	.45**	.40**	.43**
12. Inattention (VADPRS)	.92	.92	-	.53**	.43**	.43**	.43**	.43**	.53**	.43**	.43**	.43**	.43**	.46**	.43**	.43**
13. Internalizing Symptoms (VADPRS)	.90	.90	-	.72**	.49**	.49**	.49**	.49**	.72**	.49**	.49**	.49**	.49**	.72**	.49**	.49**
14. Anxiety (SCARED)	.95	.95	-	.52**	.55**	.55**	.55**	.55**	.52**	.55**	.55**	.55**	.55**	.52**	.55**	.55**
<u>Parent Psychopathology</u>																
15. Depression (PHQ8)	.91	.91	-	.79**	.79**	.79**	.79**	.79**	.79**	.79**	.79**	.79**	.79**	.79**	.79**	.79**
16. Anxiety (GAD7)	.93	.93	-	.79**	.79**	.79**	.79**	.79**	.79**	.79**	.79**	.79**	.79**	.79**	.79**	.79**

* $p < .05$.** $p < .01$. α = Cronbach's alpha

Table 7

Means and Standard Deviations for the KEPT-Full and KEPT-Brief

	Full Sample (N = 1,570)	ODD positive screen (N = 192)	CD positive screen (N = 82)
KEPT-Full (21 items)	13.56 (4.67)	12.10 (5.38)	8.51 (4.47)
KEPT-Brief (10 items)	6.69 (2.57)	5.84 (2.82)	4.02 (2.28)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 8

Percentiles for the KEPT-Full and KEPT-Brief

KEPT-Full		KEPT-Brief	
Score	%ile	Score	%ile
0	0	0	0
1	0	1	0
2	0	2	2
3	1	3	7
4	1	4	14
5	3	5	23
6	6	6	33
7	9	7	45
8	13	8	56
9	17	9	68
10	22	10	83
11	27		
12	33		
13	39		
14	46		
15	52		
16	59		
17	68		
18	75		
19	84		
20	91		
21	97		