Development and Validation of the Parenting Skill Use Diary (PSUD) in a Nationally Representative Sample

Oliver Lindhiem[1], **Rachel A. Vaughn-Coaxum**[1], **Janelle Higa**[2], **Jordan L. Harris**[2], **David J. Kolko**[1], **Paul A. Pilkonis**[1]

[1]University of Pittsburgh, Pittsburgh, PA

[2]University of Pittsburgh Medical Center, Pittsburgh, PA

Abstract

Objective—We describe the development and psychometric properties of an instrument designed to assess the use of effective parenting skills reported with a daily diary. The Parenting Skill Use Diary (PSUD) was developed iteratively relying on a "common elements" approach to quantify the use of evidence-based parenting techniques for responding to child misbehaviors and positive behaviors.

Method—The PSUD was administered online daily for seven days to parents/guardians of children aged 5–12. The nationally representative sample ($N = 1,570$) was selected to match the US population of such parents/guardians on key demographic variables.

Results—The instrument demonstrated the ability to capture significant between person variability in the appropriate use of parent management skills. A weekly summary score discriminated between parents/guardians whose children screened positive versus negative for Conduct Disorder (AUC = .72) and Oppositional Defiant Disorder (AUC = .70).

Conclusions—The results supported the reliability of validity of the diary as a research tool for examining mean differences.

Keywords

parenting; assessment; skill use; daily diary; disruptive behavior problems

---

Evidence-based treatments for childhood disruptive behavior disorders (DBDs) incorporate diverse techniques with an emphasis on teaching parent-management training (PMT) skills such as the use of rewards, praise, effective communication, and consequences for both positive and negative behaviors. These skills are conceptualized as "specific factors" or active agents of therapeutic change. To benefit from treatment, a parent/guardian must: 1) participate and engage in treatment, 2) acquire new skills or hone existing skills, and 3) incorporate these skills into daily practice. Theoretically, the day-to-day utilization of these skills leads to symptom reduction and eventual improvement in child behavior. From an "experimental therapeutic" framework (Gordon, 2017), the acquisition and utilization of

Corresponding Author: Oliver Lindhiem, PhD, Associate Professor of Psychiatry and Pediatrics, University of Pittsburgh School of Medicine, 3811 O'Hara St., Pittsburgh, PA 15213, Office: 537 Bellefield Towers, Phone: 412-246-5909, lindhiemoj@upmc.edu.

parenting skills are important treatment targets that bring about reductions in child disruptive behaviors. Measuring the engagement of these treatment targets is a critical step in refining treatments with the goal of improving outcomes. This approach is especially important for skills-based psychosocial treatments for DBDs. Skills such as problem-solving, praise, positive reinforcement, time-outs, support for emotion regulation, and clear communication can only be effective if they are practiced and utilized between treatment sessions.

Daily Diary Approach for the Assessment of Skill Use

Measuring the use of parenting skills at home is challenging and many current measures lack precision and ecological validity. For example, many parenting measures simply ask about parenting strategies over the "past 6 months" or otherwise rely on retrospective reports in general terms (e.g., compliance with homework assignments). To address this challenge, we have developed a daily internet diary to collect and aggregate data on the use of evidence-based parenting strategies. Modern mobile devices such as smartphones and internet access make daily diary methods a feasible approach to collecting data that has ecological validity and limits recall bias. Daily diaries have long been used successfully to assess child behaviors including problem behaviors (e.g., Chamberlain, Price, Reid, Landsverk, Fisher, & Stoolmiller, 2006; Chamberlain & Reid, 1987) and ADHD symptoms (e.g., Rosen & Factor, 2015; Whalen et al., 2006).

Skill Knowledge Versus Use

A daily diary approach to assessing skill use adopted in this study also recognizes the important distinction between knowledge and utilization. Skill knowledge involves a parent's or guardian's understanding of behavioral principles and ability to select appropriate responses to a variety of child behaviors and parenting scenarios. We have developed a measure of skill knowledge—the Knowledge of Effective Parenting Test (KEPT; Lindhiem et al., 2019)—but we also recognize the need to complement this measure with an assessment of the actual use of parenting skills. Skill use, in contrast, requires the execution of effective parenting strategies and behaviors in real time. This might include the practice of a skill assigned as "homework" by a therapist, but also refers more generally to skill use that is initiated spontaneously by the parent/guardian. Based on such a distinction, a parent/guardian might score high on a measure of skill knowledge but low on a measure of skill use. We expect that both constructs influence child behaviors (Lindhiem, Higa, Trentacosta, Herschell, & Kolko, 2014). Increasing both skill knowledge and skill use are implicit treatment targets common across evidence-based psychosocial treatments for childhood disruptive behavior disorders.

The Current Study

The current study involved the development and validation of the Parenting Skill Use Diary (PSUD). The PSUD is a daily diary measure in which parents/guardians report on the use of evidence-based parenting skills in their day-to-day lives. The development process was iterative, with modifications based on feedback from DBD experts, clinicians, and parents/guardians. The PSUD was designed to be consistent with the "common elements" approach

to the psychosocial treatment literature (Chorpita, Daleiden, & Weisz, 2005). Systematic studies of the evidence-based DBD treatment literature have identified several shared treatment elements including limit-setting, time-outs, praise, ignoring, communication skills, tangible rewards, and positive reinforcement (Chorpita et al., 2005; Garland, Hawley, Brookman-Frazee, & Hurlburt, 2008). These common elements have been incorporated into the PSUD. During the validation phase, the PSUD was administered daily for seven days to a nationally representative sample of parents/guardians of school-aged children. In addition to the PSUD, we collected data on parenting skill knowledge, discipline practices, child behavior, and parent psychopathology. These additional measures were used to assess the divergent and convergent validity of the PSUD. We also estimated norms and percentiles for the PSUD during this phase.

Method

Overview

The study was approved by the Institutional Review Board at the University of Pittsburgh. The Parenting Skill Use Diary (PSUD) was developed in three stages. In the first stage, the content of the diary was formulated using an iterative approach with feedback from DBD experts, clinicians, and parents. In the second stage, the PSUD was administered daily for one week to a national sample of parents/guardians of children aged 5–12, along with several other measures of related parenting constructs and child and parent psychopathology with which to establish convergent and divergent validity. In the third stage, we conducted analyses to examine the reliability and validity of the PSUD. During this final stage, we also estimated national norms, percentiles, and optimal cutoff scores that differentiated parents/guardians of children who screened positive versus negative for Oppositional Defiant Disorder (ODD) and Conduct Disorder (CD) on a standardized rating scale.

Content Development

Diary content and format—The Parenting Skill Use Diary (PSUD) was designed to assess the daily use of parenting skills in response to both positive (e.g., sharing, helping) and negative (e.g., hitting, fighting) child behaviors. Refer to the online Supplemental Material for the full measure. For each diary entry, parents/guardians are presented with a checklist of behaviors to report on for the past day. For each behavior they select, they are next asked to identify which skills (e.g., praise, reward, time-out, loss-of-privilege) they used in responding to the behaviors. Subsequent diary questions are individualized for each parent/guardian based on what behaviors/situations occurred during each reporting period through a series of if-then algorithms programmed into the online diary. To ensure that parents/guardians understand each of the terms, definitions appear in a pop-up box when the cursor is moved over a word.

Data Collection

Procedure—The PSUD and other study measures were administered by the survey company YouGov (www.yougov.com) to a national sample of parents/guardians of children aged 5–12 that was representative of the US population on key demographic variables. Participants on YouGov panels have been recruited through web advertising, email

campaigns, partner contacts, random digit dialing, and mail (based on voter registration). Participants reached via these recruitment strategies have opted-in and agreed to be contacted to complete surveys for which they are relevant. Through these recruitment strategies, YouGov has built a panel of 1.2 million U.S. residents. Recent studies have found that 84% of Americans use the internet (Perrin & Duggan, 2015), that 79% have regular access to internet in the home (File & Ryan, 2014), and that internet samples reasonably represent the overall population (Hays, Liu, & Kapteyn, 2015). Links to the surveys were emailed to YouGov panelists. The diaries could be completed on any device with a web browser (e.g., phone, tablet, laptop, or desktop machine). Each day, respondents were sent one email invitation and up to one reminder. (The full text of the emails can be found in the online Supplemental Materials.) 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 seven-day set of diaries.

Participants—Data were collected from an initial sample of 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 1.

Measures—Additional measures of related parenting constructs and child and parent psychopathology were administered to examine convergent and divergent validity of the PSUD. Constructs included child behavior problems, child depression, child anxiety, child trauma, parent discipline strategies, parent psychopathology (depression and anxiety), and parenting knowledge.

Knowledge of Effective Parenting Test (KEPT): The KEPT is a 21-item multiple choice test of parenting knowledge for parents/guardians of children aged 5–12 (Lindhiem et al., 2019). The measure has strong internal consistency (Cronbach's alpha = .84 in the study sample) and substantial evidence of construct validity. Children whose parents/guardians score below the 25th percentile are approximately 8 to 10 times as likely to screen positive for CD compared to children of parents/guardians who score above this percentile (Lindhiem, Vaughn-Coaxum, Higa, Harris, Kolko, & Pilkonis, 2018).

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 measure has well-established construct validity (Shelton, Frick, & Wooten, 1996). Internal consistency in the study sample was high, with Cronbach's alphas for the subscales ranging from .79 to .85.

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). Internal consistency in the study sample was acceptable, with a Cronbach's alpha of .60.

Vanderbilt Assessment Scale-Parent Report (VAS-P): The VAS-P is a 55-item parent-report screen for ADHD (inattentive, hyperactive, and combined type), ODD, and 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 strong internal consistency (Cronbach's alphas ranging from .90 to .96 in the study sample) 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 measure has strong convergent and divergent validity (Birmaher et al., 1997). Internal consistency in the study sample was high, with a Cronbach's alpha of .95.

Patient Heath Questionnaire depression scale (PHQ-8): The PHQ-8 was used to measure parent/guardian symptoms of depression. Items are rated using a 4-point scale from "not at all" to "nearly every day." Total scores range from 0 to 24. The measure has strong internal consistency (Cronbach's alpha of .91 in the study sample) and strong evidence of construct validity (Pressler et al., 2011).

Generalized Anxiety Disorder Screener (GAD-7): The GAD-7 was used to measure parent/guardian symptoms 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 = .93 in the study sample), good test-retest reliability (intraclass correlation = .83), and strong convergent validity with other measures of anxiety (Spitzer, Kroenke, Williams, & Löwe, 2006).

Coding and Scoring Procedures

Descriptive count data—Positive child behaviors ("Did something helpful"; "Followed directions"; "Said something nice"; "Shared"; "Was kind") and misbehaviors ("Annoyed Others"; "Argued"; "Bullied or fought"; "Destroyed property"; "Hit, kicked, or bit"; "Lied"; "Stole"; "Swore"; "Talked back"; "Threw a temper tantrum"; "Was defiant or oppositional") selected in Question 1 were summed for each day and then combined for a total weekly score. Next, strategies for responding to positive child behaviors ("Reward"; "Praise"; "Point System") and strategies for responding to misbehaviors ("Ignored"; "Time-out"' "Natural consequence"; "Loss-of-privilege"; "Logical consequence") selected in Question 2 were summed for each day and then combined for a total weekly score.

Overall proportion score—An overall proportion score representing the "effective" application of evidence-based parenting strategies was calculated for each day and then combined for a total weekly score. Step 1: each instance of skill use was coded as "within the range of effective responses" (1) or "outside the range of effective responses" (0), based on established PMT principles in the extant empirical literature. For example, ignoring non-compliance would be coded "within the range of effective responses," whereas ignoring a

tantrum would be coded "outside the range of effective responses." The coding scheme has more than one "effective" answer for most behaviors, recognizing a range of reasonable responses. See Table 2 for the complete coding scheme. Step 2: the denominator of the proportion score was calculated as the sum of the total "strategies" reported. This was calculated by summing all parenting strategies used across all behaviors for that day (not including "Other strategy" or "None of these"). This method of coding captures responses that have multiple parenting strategies used for one behavior and keeps the denominator greater than or equal to the numerator. Step 3: the numerator of the proportion score was calculated as the number of appropriate or "effective" strategies that were utilized. This is calculated by summing all "effective" parenting strategy responses for each child behavior reported. Scores with fewer than three events in the denominator were considered "missing" based on our analysis on the minimum number of events needed to estimate reliable proportion scores. Step 4: an overall proportion of effective skill use was calculated for the assessment period by dividing the weekly total number "effective" strategies (the numerator estimated in step 3) by the weekly total number of strategies used (the denominator estimated in step 2). The overall proportion score is interpreted as the proportion of times the parent/guardian used an appropriate skill in responding to discrete child behaviors during the week. The use of a proportion score controls for the opportunity to use the skills which is dependent on the frequency of behaviors that differ across children (see Lindhiem, Shaffer, & Kolko, 2014; Shaffer, Lindhiem, & Kolko, 2016). The proportion score is therefore an index that captures the percentage of time parents used "effective" strategies of all the strategies they used.

Analyses

All analyses were conducted using SPSS version 26 and R version 3.5.2. Labels for the magnitudes of effect sizes are based on Cohen's definitions of "small" ($r = .10$), "medium" ($r = .30$), and "large" ($r = .50$; Cohen, 1988).

Variability within and between persons—To optimize use of the daily repeated assessments we used mixed effects models to examine within and between person variability in skill-use and child behaviors while accounting for the nested structure of the data. All models were fit in R (version 3.6.1). Daily PSUD scores (Level 1) were nested within individuals (Level 2). Model specifications were based on prior studies of daily diary assessments (Heiy & Cheavens, 2014; Reynolds, Robels, & Repetti, 2016). Four separate linear mixed effects models were fit for the daily PSUD score of each component of the proportion score (i.e., skill use, child behaviors). Each score was treated as a repeated-measures dependent variable (lmer function of the lme4 R package). An intercept-only model with no independent variable and a random effect for the intercept was fit to each PSUD score to allow for the calculation of an intraclass correlation coefficient (ICC). The ICC indicates the proportion of variability in each PSUD score that can be attributed to between-person versus within-person differences (icc function of the sjstats R package). To evaluate within and between person variability of the proportion score, a logistic mixed effects regression was used to account for the binomial nature of the score following published recommendations (Chen, Cheng, Berkout, & Lindhiem, 2017). The logistic model configuration matched the linear regressions: intercept-only model, random effect for the

intercept, and an ICC calculated from the sjstats R package. Although the PSUD was designed to measure between-person variability is skill use, we also used the standard deviation of each participants weekly PSUD score as a proxy for within-personal variability.

Concurrent and incremental validity—We tested the concurrent validity of the PSUD proportion score using hierarchical linear regression, with separate models for testing associations with ODD symptoms and CD symptoms. We first entered the overall PSUD proportion score (i.e., skill use) into a linear regression equation as the only independent variable. Next, we tested the relative importance of skill use versus skill knowledge by controlling for skill knowledge (i.e., KEPT score). We also tested the incremental validity of the PSUD in accounting for variability in ODD and CD symptoms over and above APQ scores. Finally, we used receiver operating characteristic (ROC) analyses to determine the optimal cutoff scores for distinguishing between positive and negative screens for ODD and CD based on the VAS-P. The cutoff scores were defined as the scores that maximized sensitivity and specificity based on ROC analyses.

Convergent and divergent validity—We also examined patterns of convergent and divergent validity by computing correlations with related constructs. We expected higher correlations with more closely related constructs (e.g., other parenting constructs) than with more distinct constructs (e.g., parent psychopathology). Difference between correlations were tested for statistical significance using Steiger's method for comparing correlations (Steiger, 1980).

National norms and percentiles—Finally, we estimated national norms (*M*s, *SD*s, and percentiles) for the PSUD proportion score. Norms were also estimated for the "clinical" populations of parents/guardians of children who screened positive for ODD and CD on the Vanderbilt Assessment Scale-Parent Version (VAS-P).

Results

Preliminary Analyses

Families completed an average of six diaries out of seven ($M = 6.10$; $SD = 1.56$) for a total compliance rate of 87.2%. Typical completion time for a diary ranged from 1.6 minutes (10th percentile) to 11.7 minutes (90th percentile) with a median of 3.8 minutes. Table 3 summarizes the percentage of days parents/guardians reported specific child behaviors and parenting strategies. Parents/guardians reported on an average of ten positive child behaviors ($M = 9.71$; $SD = 6.28$) and four child misbehaviors ($M = 3.66$; $SD = 4.53$) over the seven days of diaries. They reported on an average of eight strategies ($M = 8.26$; $SD = 7.13$) for responding to positive behaviors and three strategies ($M = 3.10$; $SD = 4.50$) for responding to misbehaviors. The proportion of "effective" use of strategies over the seven days of diaries ranged from .00 to 1.00 ($M = .83$; $SD = .20$). Metrics of normality indicated a negative skew (skewness = −1.52; $SE = .07$) caused by a ceiling effect. We addressed this analytically by dropping data points at the ceiling from analyses for all inferential (not descriptive) statistics, as recommended by Austin & Brunner (2002). This approach is

recommended for large samples as it prevents Type-1 error inflation and produces unbiased results.

Variability within and between persons

The full sample was used for these analyses. Results from the linear mixed effects models revealed ICC values ranging from 0.33 to 0.38 across the four primary PSUD components. The ICC for strategies used to respond to misbehaviors was 0.34. In other words, 34% of the variability in parents' skill use for misbehaviors was attributable to differences between responders. The remainder of the variability in skill use was attributable to within-person variability (i.e. day-to-day differences in skill use) plus measurement error. The proportion of between-person variability in strategies used to respond to positive behaviors was 38% (ICC=0.38). For daily reports of child behaviors, 33% of the variability in positive behaviors was accounted for by between-person variability (ICC=0.33) and 35% of the variability in negative behaviors was attributable to between-person differences (ICC=0.35). For the overall PSUD proportion score, results from the logistic mixed effects model produced an ICC of 0.21, with 21% of the variability in scores attributable to between-person differences.

Concurrent and Incremental Validity

CD symptoms—As the single independent variable in a linear regression model, PSUD proportion scores were significantly associated with CD symptoms ($\beta = -.29$, $t = -9.17$, $p < .001$) and accounted for 9% of the variance (adjusted $R^2 = .09$). The PSUD proportion score remained significantly associated with CD symptoms ($\beta = -.25$, $t = -8.36$, $p < .001$) after the KEPT score (i.e., skill knowledge) was added to the model. The KEPT score was significantly associated with CD symptoms ($\beta = -.33$, $t = -10.68$, $p < .001$) with the PSUD proportion score already in the model and accounted for additional variance, R^2 Change $= .10$, F Change(1, 886) = 113.98, $p < .001$.

ODD symptoms—As the single independent variable in a linear regression model, PSUD proportion scores were significantly associated with ODD symptoms ($\beta = -.25$, $t = -7.62$, $p < .001$) and accounted for 6% of the variance (adjusted $R^2 = .06$). The PSUD proportion score remained significantly associated with ODD symptoms ($\beta = -.25$, $t = -7.53$, $p < .001$) after the KEPT score (i.e., skill knowledge) was added to the model. The KEPT score was not significantly associated with ODD symptoms ($\beta = -.01$, $t = -.22$, $p = .82$) with the PSUD proportion score already in the model and did not account for additional variance, R^2 Change = .00, F Change(1, 886) = 0.05, $p > .05$.

Incremental validity over the APQ—Using linear regression models, the PSUD proportion score significantly associated with CD symptoms ($\beta = -.11$, $t = -3.95$, $p < .001$) and ODD symptoms ($\beta = -.14$, $t = -4.28$, $p < .001$) the five APQ subscale scores already in the model. The additional variance explained by the PSUD proportion score was modest (1–2%) but significant for CD symptoms, R^2 Change = .01, F Change(1, 882) = 15.62, $p < .001$, and ODD symptoms, R^2 Change = .02, F Change(1, 882) = 18.31, $p < .001$.

ROC analyses—ROC analyses were used to determine whether the PSUD can reliably differentiate between positive and negative screens for CD and ODD based on the VAS-P.

Although the PSUD proportion score is not a diagnostic test for CD or ODD, the AUC, sensitivity, and specificity provide additional tests of the concurrent validity of the PSUD proportion score as a measure of an established risk factor and a target for treatment. The PSUD proportion score reliably differentiated between positive and negative screens for both CD (AUC = .72) and ODD (AUC = .70). See Figures 1 and 2 for ROC curves. Children whose parents/guardians scored below .78 (31st %ile) were almost four times as likely (risk ratio = 3.86) to screen positive for CD compared to children whose parents/guardians scored .78 or above (Sensitivity = .71; Specificity = .61). Children whose parents/guardians scored below .78 (31st %ile) were almost three times as likely (risk ratio = 2.93) to screen positive for ODD compared to children whose parents/guardians scored .78 or above (Sensitivity = .71; Specificity = .59).

Convergent and Divergent Validity

The PSUD proportion score correlated with other parenting constructs, child behaviors and symptoms, and parent/guardian psychopathology in expected ways (see Table 3). PSUD proportion scores were positively correlated with all measures of child disruptive behavior and internalizing symptoms, but the strongest associations were with symptoms of Conduct Disorder (CD) and Oppositional Defiant Disorder (ODD). These correlations were stronger than the next strongest correlations with ADHD symptoms and internalizing symptoms ($zs >$ 1.96; $p < .05$). Within the domain of parenting constructs, the PSUD was positively associated with the "positive parenting" and "involvement" subscales from the APQ and negatively associated with the "corporal punishment," "poor monitoring/supervision," and "inconsistent discipline" subscales. Notably, and in support of the distinction between skill knowledge and skill use, the correlation between PSUD proportion scores (i.e., skill use) and KEPT scores (i.e., skill knowledge) was small ($r = .12$, $p < .01$). Finally, there were small but significant associations between measures of parent psychopathology and PSUD scores. Specifically, parents/guardians who reported more symptoms of depression (PHQ-8) and anxiety (GAD-7) scored lower on the PSUD. Overall, the pattern of converging and diverging correlations support the validity of the PSUD as a measure of parenting skill use.

National Norms and Percentiles

Table 5 summarizes national norms for the PSUD proportion score. 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, PSUD proportion scores were higher for the general population than for parents/guardians of children who screened positive for ODD and CD. Parents/guardians of children who screened positive for ODD scored about three-fourths of a standard deviation lower compared to those who screened negative, $F(1, 887) = 37.56$, $p < .001$. This represents a large effect size (Cohen's $d = .77$). Differences were even larger for CD, for which the mean score on the PSUD was a full standard deviation lower for parents/guardians of children who screened positive for CD compared to those who screened negative, $F(1, 887) = 66.03$, $p < .001$. This represents a very large effect size (Cohen's $d = 1.10$). Percentiles for the PSUD are summarized in Table 6.

Discussion

The PSUD was developed to assess a critical component of parenting, namely the use of evidence-based parenting techniques for responding to child misbehaviors and positive behaviors. The techniques assessed are known from the empirical literature on common elements to be associated with childhood externalizing behaviors (e.g., Kaminski, Valle, Filene, & Boyle, 2008). The PSUD has numerous strengths including the assessment of evidence-based parenting techniques, norms and percentiles based on a large nationally representative sample, and evidence of strong reliability and validity. The scoring system was also designed to account for the dyadic nature of parenting skills and is based on the congruity between the context (i.e. the child's behavior or the presenting problem) and the parenting strategy.

The current study supported the reliability and validity of the PSUD. Analyses indicated that over 20% of variability in PSUD proportion scores could be attributed to between-person differences. The PSUD also correlated with related parenting constructs in expected ways. Specifically, PSUD proportion scores were positively correlated with the positive parenting and involvement subscales from the APQ, and negatively correlated with the corporal punishment, poor monitoring, and inconsistent discipline subscales from the APQ. Although the additional variance in CD and ODD symptoms explained by the PSUD proportion score over and above the five APQ subscales was modest, this was a conservative test of incremental validity insofar as the APQ and assessment of CD/ODD symptoms are both retrospective survey reports that were completed currently. Furthermore, several of the subscales of the APQ include items that assess child behavior and even overlap with CD and ODD symptoms. For example, the APQ item "your child is out after dark without an adult" overlaps with the CD symptom "often stays out at night despite parental prohibition." The high correlations between the APQ and symptoms of CD and ODD are likely due, at least in part, to this content overlap, making the APQ a conservative benchmark against which to test incremental validity. Therefore, these results likely serve as a lower-bound estimate for the incremental validity of the PSUD.

The strongest associations the PSUD had with a measures of child psychopathology were with CD and ODD symptoms. Specifically, higher PSUD proportion scores were associated with fewer symptoms of CD and ODD. The results also support the important distinction between skill knowledge and skill use. PSUD scores (i.e., skill use) were weakly correlated with KEPT scores (i.e., skill knowledge). This supports the importance of the distinction between skill knowledge and skill use. Overall, skill use was more important than skill knowledge in accounting for variance in ODD and CD symptoms. When entered as simultaneous independent variables in regression models, only skill use was associated with ODD symptoms whereas both skill use and skill knowledge were associated with CD symptoms. Together, parenting skill knowledge and skill use accounted for almost 20% of the variability in CD symptoms. ROC analyses further supported the validity of the PSUD as measuring an important risk factor for both CD and ODD, with risk ratios of 3.86 and 2.93 respectively.

It is worth noting that only a little over 20% of the variability in the PSUD proportion score could be attributed to between-person differences. Possible reasons for this lack of stability are worth considering. One likely cause might be that the consistent use of effective parenting strategies is less impacted by stable traits than by daily conditions, including the spillover of work stress to the home environment (e.g., Bolger, DeLongis, Kessler, & Wethington, 1989; Grzywacz, Almeida, & McDonald, 2002). This could be tested in future research. The low ICC might also suggest that the PSUD could be useful for detecting generalization of treatment gains in real life settings from one day to the next.

The negative skew of the PSUD also resulted in a ceiling effect such that roughly one-third of the sample had a "perfect" PSUD proportion score of 1.0. The PSUD is therefore only useful for assessing deficits in the use of parenting skills for those in the bottom two-thirds of the population and will not be useful in examining the full range of skill use in its current form. This, however, is consistent with the intended purpose of PSUD which is to identify levels of skill use that are associated with disruptive child behaviors. Children were identified to be at risk for disruptive behavior problems when parents/guardians score in the bottom one-third (31st %ile) of the distribution. The PSUD is therefore only appropriate as a research tool for clinical populations. The negative skew of the PSUD proportion score also raises an issue for data analyses that assume a normal distribution. When PSUD scores are analyzed as outcomes, we recommend following recent "best practice" guidelines for analyzing proportion score data (Chen et al., 2017). These recommendations include transforming the data [e.g., log(-log(1-p))] prior to analyses and beta regression.

Study Strengths and Limitations

The study had several notable strengths including the selection of content based on the common elements literature and the collection of data from a nationally representative sample of parents/guardians of children aged 5–12. Limitations included the reliance on parents/guardians as the only informants for child behavior. This limitation is somewhat mitigated by 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). Further mitigating this limitation, the specificity in the pattern of converging and diverging correlations cannot be accounted for by shared method variance.

Summary and Conclusions

The PSUD was developed as a measure of skill use for use in treatment outcome studies. It is intended to complement the KEPT (Lindhiem et al., 2019) which measures skill knowledge. Skill knowledge and skill utilization are important treatment targets for DBD treatments, and careful measurement of these constructs in the context of clinical trials has the potential to help refine existing treatments by identifying both successful and failed engagement of treatment targets (i.e., the daily use of evidence-based PMT skills).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

Austin PC, & Brunner LJ (2002). Type I error inflation in the presence of a ceiling effect. Statistical Practice, 57(2), 97–104. doi: 10.1198/0003130031450

Bolger N, DeLongis A, Kessler RC, & Wethington E (1989). The contagion of stress across multiple roles. Journal of Marriage and the Family, 175–183. doi:10.2307/352378

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]

Chamberlain P, Price JM, Reid JB, Landsverk J, Fisher PA, & Stoolmiller M (2006). Who disrupts from placement in foster and kinship care?. Child abuse & neglect, 30(4), 409–424. doi:10.1016/j.chiabu.2005.11.004 [PubMed: 16600372]

Chamberlain P, & Reid JB (1987). Parent observation and report of child symptoms. Behavioral Assessment, 9(1), 97–109.

Chen K, Cheng Y, Berkout O, & Lindhiem O (2017). Analyzing proportion scores as outcomes for prevention trials: A statistical primer. Prevention Science, 18(3), 312–321. doi:10.1007/s11121-016-0643-6 [PubMed: 26960687]

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. doi:10.1007/s11020-005-1962-6 [PubMed: 15832690]

Cohen J (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.nd

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]

File T, & Ryan C (2014). Computer and internet use in the United States: 2013. Retrieved from Washington, DC.: United States Census Bureau.

Furtkamp E, Giffort 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(5), 505–514. doi:10.1097/CHI.0b013e31816765c2

Gordon JA (2017). An experimental therapeutic approach to psychosocial interventions. In NIMH Director's Message, 3 20, 2017. Retrieved from https://www.nimh.nih.gov/about/director/messages/2017/an-experimental-therapeutic-approach-to-psychosocial-interventions.shtml

Grzywacz JG, Almeida DM, & McDonald DA (2002). Work–family spillover and daily reports of work and family stress in the adult labor force. Family relations, 51(1), 28–36. doi:10.1111/j.1741-3729.2002.00028.x

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]

Heiy JE, & Cheavens JS (2014). Back to basics: A naturalistic assessment of the experience and regulation of emotion. Emotion, 14(5), 878–891. doi: 10.1037/a0037231 [PubMed: 24999913]

Kaminski JW, Valle LA, Filene JH, & Boyle CL (2008). A meta-analytic review of components associated with parent training program effectiveness. Journal of Abnormal Child Psychology, 36, 567–589. doi:10.1007/s10802-007-9201-9 [PubMed: 18205039]

Lindhiem O, Higa J, Trentacosta C, Herschell A, & 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, 41–66. doi: 10.1007/s10567-013-0136-0 [PubMed: 23649324]

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]

Lindhiem O, Vaughn-Coaxum R, Higa J, Harris JL, Kolko DJ, & Pilkonis PA (2019). Development and validation of the Knowledge of Effective Parenting Test (KEPT) in a nationally representative sample. Psychological Assessment, 31(6), 781–792. doi: 10.1037/pas0000699 [PubMed: 30742461]

Loeber R, Green SM, Lahey BB, & Stouthamer-Loeber M (1989). Optimal informants on childhood disruptive behaviors. Development and Psychopathology, 1(4), 317–337. doi:10.1017/S095457940000050X

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. doi:10.4037/ajcc2010931 [PubMed: 20378777]

Reis HT, & Gable SL (2000). Event sampling and other methods for studying everyday experience. In Reis HT & Judd CM (Eds.), Handbook of research methods in social and personality psychology (pp. 190–222). New York, NY, US: Cambridge University Press.

Reynolds BM, Robles TF, & Repetti RL (2016). Measurement reactivity and fatigue effects in daily diary research with families. Developmental Psychology, 52(3), 442–456. doi: 10.1037/dev0000081 [PubMed: 26689757]

Rosen PJ, & Factor PI (2015). Emotional impulsivity and emotional and behavioral difficulties among children with ADHD: An ecological momentary assessment study. Journal of Attention Disorders, 19(9), 779–793. doi:10.1177/1087054712463064 [PubMed: 23172248]

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(3), 317–329. doi:10.1207/s15374424jccp2503_8

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. doi:10.1001/archinte.166.10.1092 [PubMed: 16717171]

Steiger JH (1980). Tests for comparing elements of a correlation matrix. Psychological Bulletin, 87, 245–251.

Sturmey 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(3), 223–227. doi:10.1016/0005-7916(87)90004-8 [PubMed: 3667950]

Whalen CK, Henker B, Jamner LD, Ishikawa SS, Floro JN, Swindle R,… & Johnston JA (2006). Toward mapping daily challenges of living with ADHD: Maternal and child perspectives using electronic diaries. Journal of Abnormal Child Psychology, 34(1), 115–130. doi:10.1007/s10802-005-9008-5 [PubMed: 16463071]

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 Pediactric Psychology, 28(8), 559–568. doi:10.1093/jpepsy/jsg046
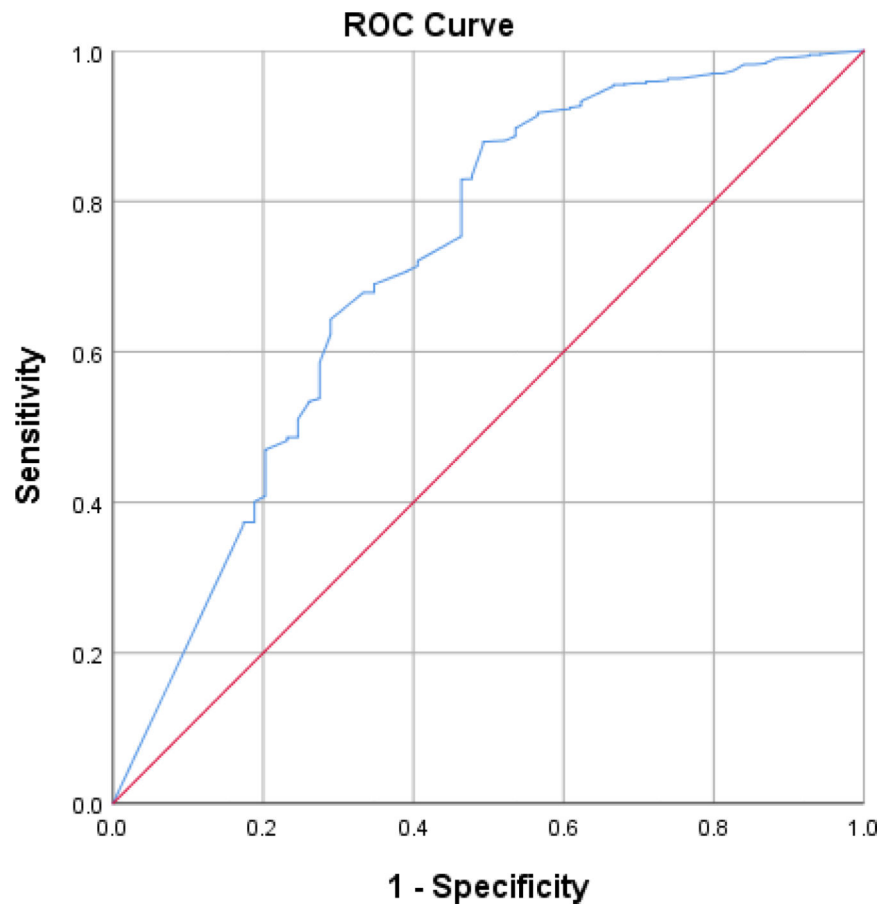
Figure 1.
ROC curve for distinguishing between positive and negative screens for CD from the PSUD proportion score (AUC = .72).
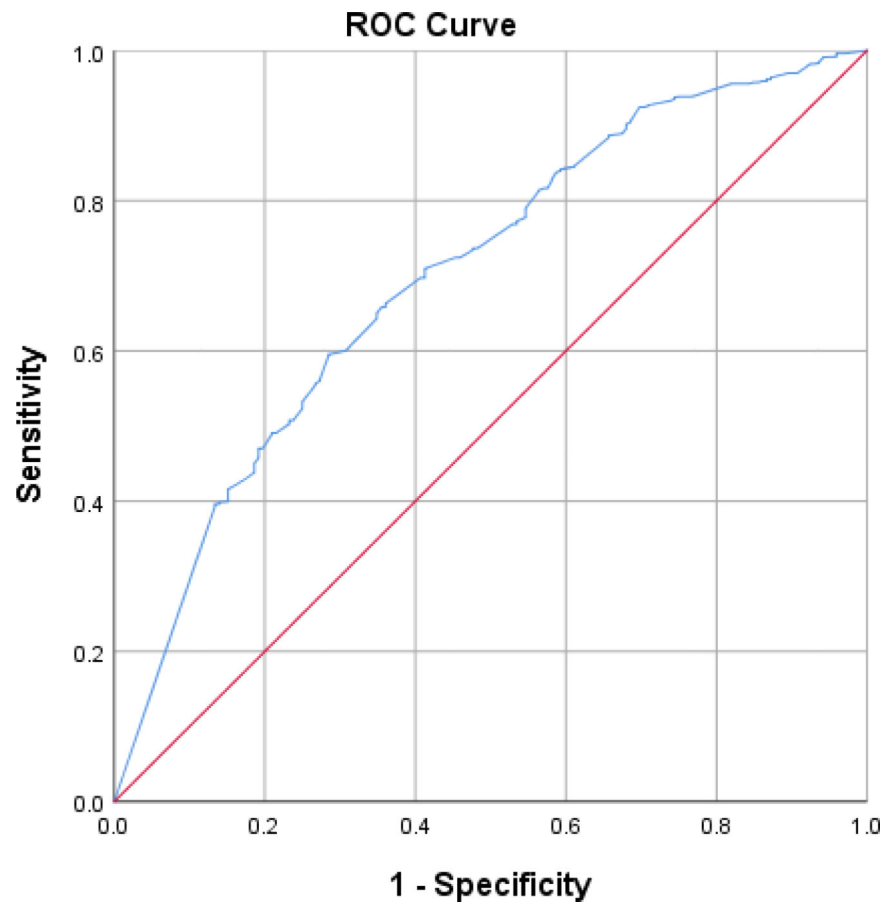
Figure 2.
ROC curve for distinguishing between positive and negative screens for ODD from the PSUD proportion score (AUC = .70).

Table 1

Participant Demographics

	Full Sample (N = 1,570)	Subsample (N = 889)
Parent/Guardian Age	*M* = 41.5; *SD* = 10.6	*M* = 41.0; *SD* = 10.2
Parent/Guardian Gender		
Female	63.7%	64.5%
Male	36.3%	35.5%
Child Age	*M* = 8.6; *SD* = 2.3	*M* = 8.3; *SD* = 2.4
Child Gender		
Female	46.8%	46.6%
Male	53.2%	53.4%
Parent/Guardian Race		
White	78.7%	82.2%
Black or African American	13.9%	13.0%
Asian	3.4%	3.4%
American Indian or Alaska Native	2.6%	2.0%
Native Hawaiian or Pacific Islander	.6%	.3%
Some other race	5.5%	3.6%
Parent/Guardian Ethnicity		
Hispanic, Latino, or Spanish Origin	16.2%	13.9%
Parent/Guardian Born in the US	89.7%	89.7%
Parent/Guardian Education		
High school or lower	25.6%	20.9%
Some college or higher	74.4%	79.1%
Marital Status (% married)	68.2%	69.1%
Family Income		
Less than $10,000	6.0%	4.7%
$10,000 – $29,999	20.0%	20.0%
$30,000 – $49,999	20.3%	19.4%
$50,000 – $69,999	16.6%	17.4%
$70,000 and above	37.2%	38.5%
US Region		
Northeast	18.3%	18.8%
Midwest	24.3%	23.7%
South	37.0%	36.4%
West	20.4%	21.0%

Table 2

Coding Key for Calculating the PSUD Proportion Score

Behavior	Type of Behavior	Range of Effective Parenting Strategies
Annoyed others	Misbehavior	Loss-of-privilege, Logical consequence
Argued	Misbehavior	Loss-of-privilege, Logical consequence
Bullied or fought	Misbehavior	Time-out, Loss-of-privilege, Logical consequence
Destroyed property	Misbehavior	Time-out, Loss-of-privilege. Logical consequence
Did something helpful	Positive behavior	Reward, Praise, Point system
Followed Directions	Positive behavior	Reward, Praise, Point system
Hit, kicked, or bit	Misbehavior	Time-out, Loss-of-privilege, Logical consequence
Lied	Misbehavior	Loss-of-privilege, Logical consequence
Said something nice	Positive behavior	Reward, Praise, Point system
Shared	Positive behavior	Reward, Praise, Point system
Stole	Misbehavior	Loss-of-privilege, Logical consequence
Swore	Misbehavior	Loss-of-privilege, Logical consequence
Talked back	Misbehavior	Loss-of-privilege, Logical consequence
Threw a temper tantrum	Misbehavior	Ignoring
Was defiant or oppositional	Misbehavior	Loss-of-privilege, Logical consequence
Was kind	Positive behavior	Reward, Praise, Point system

Table 3

Percentage of Days of Parents/Guardians Reported Specific Child Behaviors and Parenting Strategies

Positive Child Behaviors		Negative Child Behaviors	
Followed directions	41.8%	Talked back	14.9%
Did something helpful	35.6%	Argued	14.4%
Was kind	36.1%	Annoyed others	10.1%
Said something nice	28.5%	Was defiant or oppositional	8.7%
Shared	20.9%	Threw a temper tantrum	6.2%
		Lied	3.2%
		Hit, kicked, or bit	1.9%
		Bullied or fought	1.5%
		Swore	1.5%
		Destroyed property	1.1%
		Stole	0.1%
Parenting Strategy for Positive Child Behaviors		Parenting Strategy for Negative Child Behavior	
Praise	52.3%	Ignore	15.6%
Reward	24.5%	Loss-of-privilege	13.6%
Point System	9.0%	Time-out	12.0%
		Natural consequence	6.6%
		Logical consequence	5.3%

Table 4

Correlations Between the PSUD Proportion Score, Measures of Child Behavior, Parenting, and Parent and Child Psychopathology

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Parenting															
1. PSUD Proportion Score[a]	.12**	-.22**	-.22**	-.20**	.31**	.28**	-.06	-.29**	-.25**	-.19**	-.19**	-.19**	-.17**	-.14**	-.17**
2. KEPT-Full (21 items)	-	-.42**	-.44**	-.19**	-.02	.02	.44**	-.31**	-.02	-.05*	.00	.04	-.18**	.00	.00
3. Corporal Punishment (APQ subscale)		-	.59**	.44**	-.17**	-.14**	-.09**	.58**	.31**	.35**	.34**	.29**	.39**	.20**	.21**
4. Poor Monitoring (APQ subscale)			-	.54**	-.19**	-.05	-.08**	.58**	.28**	.26**	.31**	.31**	.36**	.18**	.18**
5. Inconsistent Discipline (APQ subscale)				-	-.21**	-.17**	-.02	.45**	.43**	.37**	.40**	.37**	.37**	.28**	.27**
6. Positive Parenting (APQ subscale)					-	.66**	-.14**	-.23**	-.19**	-.13**	-.21**	-.18**	-.11**	-.08**	-.11**
7. Involvement (APQ subscale)						-	-.08**	-.16**	-.24**	-.16**	-.25**	-.15**	-.10**	-.16**	-.15**
8. Knowledge of Behavior Principles							-	.00	.12**	.11**	.10**	.16**	.03	.10**	.12**
Child Behavior / Psychopathology															
9. CD Symptoms (VADPRS)								-	.59**	.49**	.48**	.61**	.61**	.40**	.41**
10. ODD Symptoms (VADPRS)									-	.72**	.69**	.60**	.51**	.44**	.47**
11. Hyperactivity / Impulsivity (VADPRS)										-	.74**	.48**	.45**	.40**	.43**
12. Inattention (VADPRS)											-	.53**	.46**	.43**	.43**
13. Internalizing Symptoms (VADPRS)												-	.72**	.49**	.49**
14. Anxiety (SCARED)													-	.52**	.55**
Parent Psychopathology															
15. Depression (PHQ8)														-	.79**
16. Anxiety (GAD7)															-

[a] n = 889.

* p < .05.

** p < .01.

Table 5

Weekly Means and Standard Deviations for the PSUD

	Overall (*N* = 1,570)	ODD positive screen (*N* = 192)	CD positive screen (*N* = 82)
Child Positive Behaviors	9.71 (6.28)	7.81 (5.67)	5.42 (2.42)
Child Misbehaviors	3.66 (4.53)	7.28 (8.21)	6.80 (10.05)
Responses to Positive Behaviors	8.26 (7.13)	7.58 (6.89)	6.15 (7.48)
Responses to Misbehaviors	3.10 (4.50)	6.49 (7.84)	6.56 (9.90)
PSUD Proportion Score	.83 (.20)	.70 (.25)	.63 (.30)

Table 6

Percentiles for the PSUD Proportion Score

Score	%ile
.30	2
.35	3
.40	4
.45	5
.50	6
.55	9
.60	10
.65	14
.70	20
.75	26
.80	34
.85	44
.90	53
.95	62
1.00	64