Article

Psychometric Evaluation of a New Digitally Animated Child Self-Report Assessment Instrument: The Interactive Child Distress Screener

Assessment
2023, Vol. 30(3) 907–922

DOI: 10.1177/10731911211072907
journals.sagepub.com/home/asm
SAGE

Kirsty Zieschank[1], Michael J. Ireland[1], Jamin Day[2], and Sonja March[1]

Abstract

The present study aimed to provide the first psychometric evaluation of the newly developed, digitally animated assessment instrument: the *Interactive Child Distress Screener* (*ICDS*). The latent factor structure of the novel ICDS was first established using exploratory factor analysis (EFA) on 15 pairs of animated items using a community sample ($N = 266$) of child–parent dyads. EFA results support a two-factor structure representing two broad domains of internalizing and externalizing difficulties ($r = .52$) and comprised 12 items. The reliability of the factors was strong with ordinal alpha and omega coefficients above .84 and .87, respectively, for each of the subscales. Convergent validity for the overall sample was supported with established child and parent-reported measures of internalizing and externalizing problems; however, the ICDS factors demonstrated convergence greater in magnitude with other child-reported measures such as the Me and My School Survey. Satisfaction and utility ratings by children indicated that the digital format was highly acceptable.

Keywords

ICDS, digital assessment, exploratory factor analysis, child self-report scale, psychometrics

Mental health problems throughout childhood can have detrimental effects on psychosocial well-being, academic development, and future achievement (Guzman et al., 2011; Merikangas et al., 2010; Ogundele, 2018). Prevalence rates are high with one in six American children, one in seven Australian children, and two in five British children meeting the criteria for a clinical mental health diagnosis (Deighton et al., 2019; Lawrence et al., 2016; Whitney & Peterson, 2019). The most commonly diagnosed mental health problems in school-aged children (5–11 years) are disruptive behavioral problems (conduct disorders and attention-deficit hyperactivity disorder), anxiety, and mood disorders (i.e., depression; Centers for Disease Control and Prevention, 2020; Lawrence et al., 2016). The available evidence suggests that the symptoms of such problems can occur for 2 to 4 years without detection, and almost half of children suffering will never receive treatment (National Research Council and Institute of Medicine., 2009). With such delayed recognition, valuable opportunities to intervene are missed before problems advance to clinical levels. Accordingly, universal mental health screening (hereafter referred to as screening) for emotional and behavioral difficulties in accessible locations such as primary health care settings and schools has long been promoted to aid detection and prevention (Carter et al., 2004; Dowdy et al., 2010; Humphrey & Wigelsworth, 2016; Mihalopoulos et al., 2012; The Royal Australian and New Zealand College of Psychiatrists, 2010).

Ideally, a screening instrument should be brief, cost-effective, psychometrically sound, and successfully discriminate between children who require further evaluation and those who do not (Goodman-Scott et al., 2019; Ivey, 2020; Newlove-Delgado & Ford, 2020) and to increase accessibility, require no training to procure, administer, and score. A recent review by McCrae and Brown (2018) describes three suitably broad screening instruments for use with school-aged children: the *Behavior Assessment System for Children, Second Edition Behavioral and Emotional Screening System* (BASC-2-BESS; Kamphaus & Reynolds,

[1]University of Southern Queensland, Springfield Campus, Springfield Central, Australia
[2]The University of Newcastle, Callaghan, New South Wales, Australia

Corresponding Author:
Kirsty Zieschank, School of Psychology and Counselling & Centre for Health Research, University of Southern Queensland, Springfield Campus, P.O. Box 4196, Springfield Central, Queensland 4300, Australia.
Email: kirsty.zieschank@usq.edu.au

2007), the *Pediatric Symptom Checklist* (PSC-17; Gardner et al., 1999), and the *Strengths and Difficulties Questionnaire* (SDQ; Goodman, 1997). The *Brief Problem Monitor* (BPM; Achenbach et al., 2011) and the *Me and My School Questionnaire* (M&MS; Deighton et al., 2013) are other common instruments. Unfortunately, there are many barriers to implementing universal screening due to limited accessibility to these tools. Such barriers include awareness of, and restricted access of appropriate instruments to specific professionals, the extensive parent and professional investment necessary to conduct most assessments, and financial costs (Ivey, 2020; Wood & McDaniel, 2020). For example, the BPM and BESS are expensive (AUD250 for 50 administrations and AUD450 for 25 administrations respectively) and like the SDQ, they require professional access and training for administration, scoring, and interpretation. Such factors ultimately limit the potential of such instruments for universal application.

Deighton et al. (2014) argue that child-reported measures may be less burdensome to administer. Yet, instruments assessing mental health in school-aged children (i.e., $<$11 years old) almost exclusively use adult caregivers (e.g., parents and teachers) as proxy informants. Out of the five measures previously listed, only the BESS and M&MS collect self-reported information from children under 11 years of age. Evidence shows that both parent and child perspectives are important and that with suitable measures, young children can not only provide unique and clinically useful information but more accurate accounts of their internalizing symptoms than adult informants (Arseneault et al., 2005; Dowdy et al., 2010; Jeffrey et al., 2020; Moffa et al., 2019). There is also growing impetus to foster client-centered practices and a shift toward patient-reported outcomes that encourage children to have a greater voice in their own health choices and care.

A ubiquitous feature of existing measures, and standard psychological screening measures generally, is their presentation and response method. That is, they comprise written statements or questions and utilize Likert-style response scales with temporal sequencing (e.g., ". . . over the past 5 days": "over the last six months"). Cognitive testing with children has demonstrated that word recognition problems, misunderstanding content, response option incongruence, and misapplying response options to content are issues that can occur when children are required to interpret assessment items (Bowen, 2008). Furthermore, Piagetian theory suggests that young children engage in dichotomous thinking and are more likely to focus exclusively on the two extremes of Likert-type scales (Chambers & Johnston, 2002). With young children's more limited capacity for reading comprehension, content appraisal, and scaling of responses, we believe that the prevailing presentation style degrades the reliability and validity of their answers and that a novel presentation approach is required. In a recent systematic literature review investigating children's ability to self-report, its authors concluded that measures for children should accommodate developmental variations by minimizing assessment demands through refined wording or including auditory and pictographic response items (Bevans et al., 2020).

There have been some attempts to develop such instruments including a computer-administered version of the SDQ with added static color graphics that was pilot-tested with children as early as 2001. It demonstrated clinical sensitivity in children 11 years and older and was able to discriminate between community and clinical populations based on receiver operating characteristic (ROC) curve analysis (ROC = 0.76, 95% confidence interval [CI] [0.68, 0.85]) (Truman et al., 2003). The computer version of the SDQ also demonstrated higher user satisfaction ratings and improved engagement compared with the standard pencil-and-paper version (Truman et al., 2003). Another example is the *Mood Assessment via Animated Characters* (MAAC) instrument for assessing anxiety. Measuring 16 emotions, the MAAC was able to discriminate between anxious and non-anxious children utilizing static and animated images (Manassis et al., 2009). Although it is not designed as a screening instrument, the *Berkeley Puppet Interview* (BPI) represents another child-focused method for engaging children in a structured discussion about their emotions and behaviors via hand puppets. Ringoot et al. (2017) demonstrated that child reports collected via the BPI predicted treatment referral up to 2 years later and consistently correlated with parent ratings on the *Child Behavior Checklist* (Piper et al., 2014). Although none of the aforementioned instruments are easily accessible (as described earlier), scalable, or intended directly for the purpose of broadly screening for emotional and behavioral difficulties, such results partly realize the potential of integrating visual components and improving self-reported assessment with young children. With the proliferation of the internet, and the ease with which information can now be accessed on digital devices (e.g., smartphones, tablets), there exists great scope to facilitate access to universal screening in school environments, community agencies, open-access platforms, and primary health care settings.

To advance self-reported screening for young children by overcoming the aforementioned issues, the current research sought to psychometrically evaluate a recently developed digitally animated assessment tool: the Interactive Child Distress Screener (ICDS). Given the pervasive reach of the internet, digital assessments have the potential to facilitate broad access to those with more limited resources or capacity to attend professional services. They can also be less resource-intensive than paper-based instruments because they can be accessed instantaneously without ordering or printing materials and can be developed to allow automated immediate scoring and reporting of results minimizing the need for specialist input. This also reduces financial costs for families, which may in turn

Table 1. Included Child Participants as a Function of Age and Gender (N = 266).

	Age in years, *n* (%)							
Gender	5	6	7	8	9	10	11	(*N*)
Male	18 (13.24)	24 (17.65)	27 (19.85)	24 (17.65)	13 (9.56)	12 (8.82)	18 (13.24)	136
Female	15 (11.54)	22 (16.92)	21 (16.15)	21 (16.15)	23 (17.69)	10 (7.69)	18 (13.85)	130
Total	33 (12.41)	46 (17.29)	48 (18.05)	45 (16.92)	36 (13.53)	22 (8.27)	36 (13.53)	266

Note. One parent participated with each child. Parent participants $N = 266$.

facilitate administration in routine care settings and greater scalability. This study aimed to psychometrically evaluate the ICDS, which is an important step toward identifying its utility as a self-report screening instrument for children.

The ICDS was co-developed with more than 100 children (aged 4–12 years) and has been described in previous feasibility and development studies (March et al., 2018; Zieschank, Day et al., 2021; Zieschank, Machin et al., 2021). In the first of these development studies, child participants discussed, defined, and modeled audiovisual and behavioral exemplars for each of the contrasting emotional and behavioral construct pairs as the first step of the item content co-design process. The resulting shared interpretations formed the narrative framework for and subsequent creation of 30 prototype animated ICDS assessment items (Zieschank, Machin et al., 2021). In the second of these studies, child participants' understanding of the prototypes was evaluated throughout iterative co-design cycles of animation testing, analysis, and refinement (Zieschank, Day et al., 2021). The content validity of the animated items was supported when participants as young as 5 years could accurately identify the intended emotional and behavioral constructs depicted in each animation.

Aims

The ICDS assessment items were broadly developed under the domains of behavioral and emotional difficulties (March et al., 2018) and form a brief, digitally animated, child self-report assessment instrument accessible via web-enabled digital devices. The current study aims to psychometrically evaluate the novel ICDS in a community sample of primary school-aged children (5–11 years) as the first step in establishing its utility as a screening tool. The current study examines its structural validity, internal consistency, and convergent validity. If psychometric evaluation of the ICDS is adequate, it will confirm that the instrument has potential for further investigation as a self-reported screening tool within health care and educational settings to facilitate early intervention.

A further aim was to examine whether the ICDS was acceptable to users and functional as expected as a brief easy-to-use assessment instrument as demonstrated through high satisfaction and utility ratings. We had no a priori hypotheses regarding the factor structure of the items and adopted an exploratory approach to determining this. Based on overlapping theoretical constructs, we hypothesized moderate to strong positive correlations between the ICDS and other child-reported measures of behavioral and emotional difficulties (i.e., M&MS and the Brief Problem Monitor–Youth Form). Due to the expectation that parent and child reports of the same construct would share less variance than converging child reports, we hypothesized small to moderate positive correlations between the ICDS and parent-reported measures of behavioral and emotional difficulties (i.e., the Strengths and Difficulties Questionnaire–Parent Form and the Brief Problem Monitor–Parent Form).

Method

Participants

Two hundred sixty-six parents with children (51% male, $M_{age} = 7.81$ years, $SD = 1.91$) were recruited in dyads for this study from two community sources: a local state primary (elementary) school in Brisbane ($N = 74$, 54% male, $M_{age} = 7.43$ years, $SD = 1.90$) Australia and an online sample via national social media advertising ($N = 192$, 50% male, $M_{age} = 7.95$ years, $SD = 1.89$). Inclusion criteria for each dyad required the child participant to complete the ICDS and M&MS and the corresponding parent to complete at least one parent-rated measure (i.e., either the BPM–Parent form [BPM-P] or SDQ–Parent form [SDQ-P]) to examine convergent validity. The proportion of child participants across ages and gender is presented in Table 1. For brevity, the term "parents" is used throughout this article and includes mothers, fathers, stepparents, or legal guardians of a child participant.

Procedure

The current study was observational utilizing a cross-sectional survey approach with parent–child dyads recruited from both school and online community sources. Ethical

approval was obtained from the University of Southern Queensland and the Queensland Government Department of Education.

Recruitment

School-Based Recruitment and Procedure. Participants were recruited from one Brisbane school so that we could directly observe a subsample of children completing the ICDS measure to examine its practical utility. The school was provided with research study invitation packs to distribute to students. Each pack contained individualized parent and child information and consent forms, one copy each of the BPM-P and SDQ-P for the parent to complete, and an envelope to return them. Seventy-four parents (10.6% response rate) returned completed parent-rated measures and consented to their child's involvement. This seemingly low response rate was likely due to the multiple steps required for children to participate. For example, the children needed to receive the invitation packs from their class teacher, give it to their parent, have their parent read information and complete consent forms and measures, return the completed measures back to their class teacher within 2 weeks, and then participate themselves in-person at their school on a specific day.

The first author met with each child participant individually during their school day to complete all child-report measures in person. The children were asked to write their name under their parents on the consent form if they agreed to participate and were advised that even if their parent had consented, they could withdraw at any time. Each child participant then completed the ICDS on a touchscreen tablet, and all other measures (M&MS, BPM, and ICDS Satisfaction Survey) were completed as recommended with pen and paper. Demographic data (age and gender) were collected from each child within the ICDS application. All children were observed by the first author (K.Z.) while they completed the ICDS to conduct the utility assessment. Incentives were not offered to participants sourced from the school.

Online Recruitment and Procedure. Advertisements on social networks (i.e., Facebook and Instagram) were utilized to recruit 192 parent–child dyads. Identical information and consent materials provided to school participants were also presented online using the survey website. Parents provided online consent and completed a brief demographic questionnaire about their child (child age, gender, school year level, and parent email address) followed by parent-report measures (BPM-P and SDQ-P). Parents were then instructed to have their child complete the M&MS, BPM, ICDS, and ICDS Satisfaction Survey. An AUD15 gift card was offered to compensate participating families for their time, with 64 (33%) providing contact details to receive this.

General Procedure Details

All participant data collected in this study were scored and evaluated so that we could inform participants of any potential elevated levels of distress, in line with ethical guidelines. That is, all parents were informed if any of their child's scores indicated an elevated level of distress according to the norms of the completed measures. Along with the notification, participants were provided with recommendations to obtain further assessment and referral information. Regardless of recruitment source, each member of the dyad completed the measures in the same sequence. For parents, this included the BPM-P and the SDQ-P, and for children, this included the ICDS, M&MS, BPM-Y, and ICDS satisfaction scale. The self-reported M&MS and BPM-Y were not developed for completion by children younger than 8 and 11 years, respectively; however, the M&MS was co-developed with children and the BPM Manual states that younger children ($<$ 11 years) can act as an informant "if they are able" (Achenbach et al., 2011). Therefore, we included children outside the intended age range for these measures to provide comparative child-reported data for the purpose of evaluating the utility of the self-reported ICDS within these age ranges. For the school-based sample, the measures were read to the younger participants if required by the researcher. For the online sample, parents were asked to assist their child by reading items only if required. The instruction stressed they were to read the item verbatim but allow their child to answer the question themselves. The M&MS was used with permission from the author and the BPM-P, BPM-Y, and SDQ measures were used under license.

Measures

ICDS. The ICDS is a 15-item, animated, digital assessment instrument that was designed to detect self-reported emotional and behavioral difficulties among 5- to 11-year-old children. The item constructs were generated in a prior feasibility study by an expert panel of child psychologists and psychometricians (March et al., 2018). Each of the 15 assessment items comprised two animations that depict children experiencing contrasting emotional or behavioral states. Negatively valenced states (e.g., sadness) are categorized as "Target Items" and are scored with a "1." Contrasting animations are positively valenced (e.g., happiness) and considered as control or comparison items, which are scored as "0."

The ICDS is delivered via a web-based application (app) and is optimized for use on computers, mobile devices (smartphones), and tablets (iOS and Android). The app opens with an administration page where the test administrator and parent details (email address) are captured, and an anonymous participant code is automatically generated.

A *Welcome* page follows to guide children on independently completing the rest of the measure by clicking or tapping on images, accompanied by a voiceover. Children initially choose a colorful "*Buddy*" that is situated at the bottom of the screen throughout the measure and the child to be aged 5 through 11 years. *Buddy* is a spoken voiceover assistant that is activated by the user via a tap or click that provides instructions automatically and on-demand throughout the measure. Requested demographic information includes age (shown as numerals from four to 13) and gender (shown as girl and boy images and written words). An "*other*" choice is provided for those who prefer not to identify as one of these two genders. After submitting these responses, each pair of animated items is then presented sequentially. As each animation appears on screen the user is initially directed verbally to tap or click a "*play*" button (represented by a stereotypical triangle) centered on the animation. The first animation seen by the user in each pair (i.e., target or contrasting animation) is randomized. Tapping on the triangle activates the video that enlarges to fill the screen.

After participants have viewed both animations in a pair, they are asked to respond to the audible and written question "Which one is most like you?" by clicking or tapping on their chosen animation. Developmental research shows that in middle childhood, executive function capabilities, including memory, self-monitoring, and cognitive processing, are still developing, and making retrospective temporal judgments is more difficult (Droit-Volet & Coull, 2015; McCormack & Hoerl, 2017). As a result, children tend to rely on the information that is most readily available when responding to questions. It is for this reason that questionnaires often ask children to comment on how they are feeling "today." With respect to our broad screening questions, which spanned home, school, and social environments, we were not wanting to focus on one day or a specific time period or event (e.g., "during the last week"), as this would likely elicit inexact responses. Instead, we expected that without time context their response was more likely to reflect recent symptomatology rather than trait-like behaviors, as recent negative symptoms would likely be most salient to the child. The sequence of ICDS webpages as seen by the user, including the animation viewing and response sequence, is shown in Figure 1.

ICDS Utility Measure. To determine how functional the delivery format of the ICDS application was for child users, the number of instances that each child requested assistance in completing the ICDS was tallied. Utility observations were made regarding children's capacity to use the ICDS and successfully complete the demographic collection page, navigate through the application by following audible and written directions, appropriately use the "play" buttons and "next" arrows, select responses for each item, and submit their data. In addition, any verbal requests for assistance and use of the in-app helper assistant "Buddy" were recorded. The utility assessment was omitted for the online sample of participants.

M&MS Questionnaire. The M&MS is a brief, 16-item self-reported mental health measure for children as young as eight years old (Deighton et al., 2013). It yields composite scores for emotional difficulties (10 items) and behavioral difficulties (6 items), as well as a total score (range 0–32). Higher scores reflect greater difficulties. The items consist of short written statements such as "Nobody likes me" and "I lose my temper" and utilizes three response options (0: *never*; 1: *sometimes*; 2: *always*) regarding how the participant feels at the time of assessment. The M&MS is reported to demonstrate good internal consistency across ages 8–11 years (behavioral difficulties: $\alpha = .68–.80$; emotional difficulties: $\alpha = .72–.77$). Construct validity has been examined with 11-year-old children demonstrating moderate to strong correlations between the M&MS subscales and corresponding child-reported SDQ behavioral ($r = .56–.67$, $p < .001$) and emotional ($r = .70–.85$, $p < .001$) subscales (Deighton et al., 2013; Patalay et al., 2014). Internal consistency for the current sample is reported in Table 4.

ICDS Satisfaction Survey. The final ICDS webpage invites both child and parent participants to leave written feedback about the ICDS and provides an option for children to complete a nine-item author-developed measure of user satisfaction. The first four questions of the survey asked child participants to compare the digital format of the ICDS to the written format of the M&MS (i.e., as an example of a pen and paper style survey) across several factors and choose which they preferred. Questions asked which format they: (a) liked more, (b) thought was easier to understand, (c) would want to do again, and (d) would recommend to other children. The survey was presented as a simple 9-point Likert-type scale represented as a line, which had a picture of the M&MS scale placed at the far-left of the line (coded as a score of "1"), a zero placed at the mid-point (coded as a score of "5"), and a picture of the ICDS logo placed at the far-right (coded as a score of "9"). For each question, the participants were asked to indicate their response by making a mark on the line closest to their preferred measure. If they had no preference, they were advised to make a mark toward the middle of the line. Marks reflecting scores between 1 and 3 were coded as a distinct preference for the M&MS (i.e., a paper-based survey), scores between 4 and 6 were deemed reflective of a participant having no preference, and scores of 7 to 9 were rated as an explicit preference for the ICDS. This was calculated for each of the first four questions.

The final five questions rated participant satisfaction with the ICDS specifically, utilizing a yes/no scale represented by sad and happy face emoticons placed at the

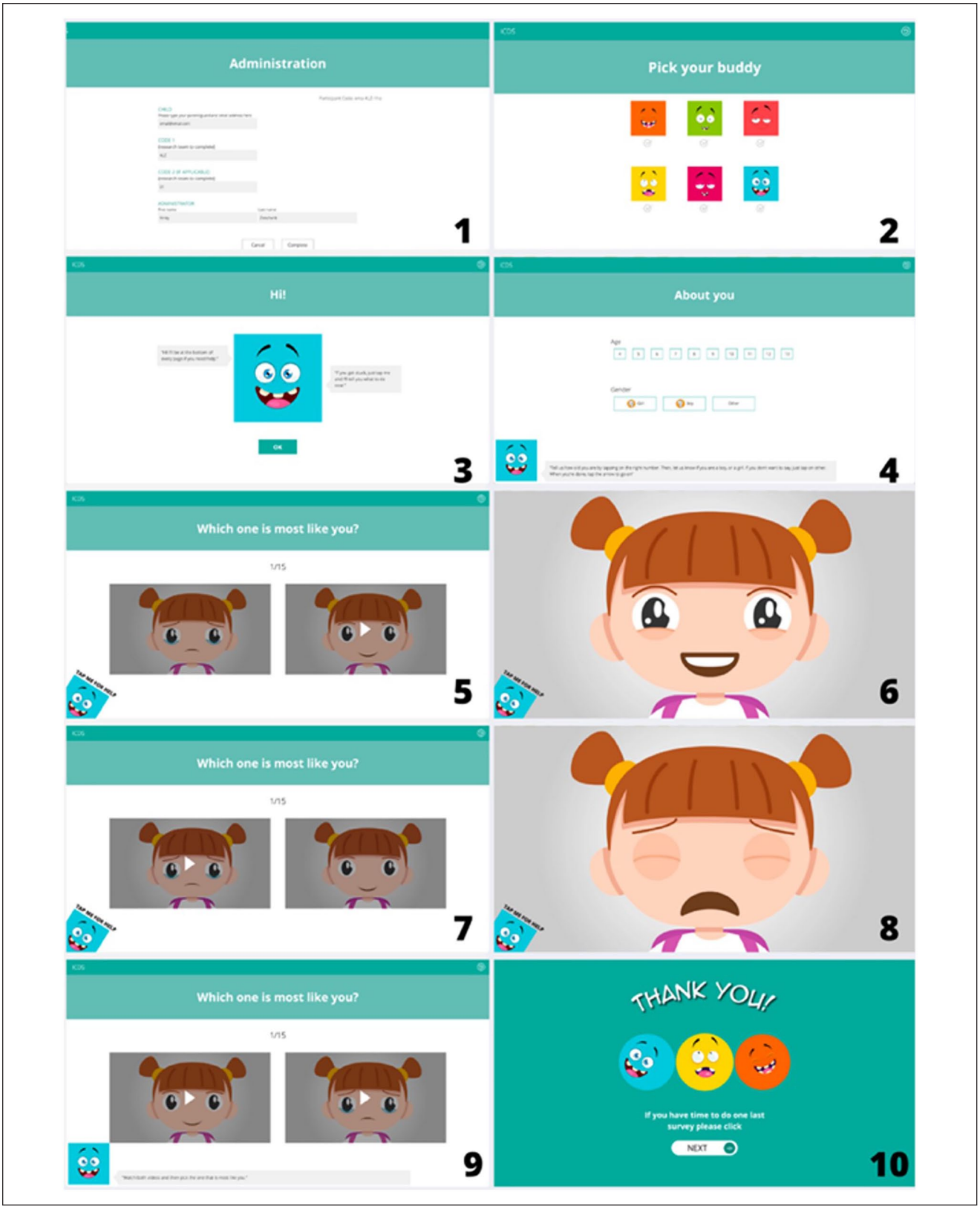

Figure 1. ICDS Web Application Screenshots.
Note. 1 = Administration page; 2 = Buddy helper assistant selection; 3 = Buddy assistant instructions; 4 =Demographic collection; 5–9 = Animated item, play, view, and response selection pages (repeats for each item); 10 = End and submission page.

extreme ends of the line. Participants were asked to rate whether they thought the ICDS (a) was fun to do, (b) had easy instructions, (c) had a response option that was easy to understand, (d) took too long to complete, and (e) had likable characters. Participants responded by choosing either a sad or happy emoticon image or placing a mark anywhere between the two if they could not make a clear choice. A sad-face response was scored as "0" and indicated dissatisfaction, and a smiley-face was scored as "1" and indicated satisfaction with the ICDS. Any response between the two emoticons was rated as undecided.

SDQ–Parent Form. The SDQ-P by Goodman (1997) is a 25-item parent-rated measure for children aged 4 to 17 years. It comprises five, 5-item subscales: emotional symptoms, peer problems, conduct problems, hyperactivity inattention, and prosocial behavior. A total difficulties score (range 0–40) excludes the prosocial scale. Broader dimensions may be examined by calculating externalizing (sum of conduct and hyperactivity items) and internalizing (sum of emotional and peer problem items) subscale scores. Higher scores reflect greater difficulties on each subscale and total score. The SDQ-P utilizes three response options (0: *not true*; 1: *somewhat true*; 2: *certainly true*) regarding the young person's behavior over the last 6 months or school year. Written statements include "Nervous in new situations, easily loses confidence" and "Often unhappy, depressed or tearful." Equivalent statements are used in the SDQ–Youth form (i.e., "I am nervous in new situations, I easily lose confidence").

The SDQ has been evaluated in multiple countries with some variability in reporting and results. An Australian study with a large community sample of young children (N = 1,359) aged 4 to 9 years reported moderate to strong internal reliability (α = .59–.80) across all five subscales (Hawes & Dadds, 2004) and stability at 12 months (α = .61–.77) when compared with diagnostic interviews. This is consistent with the original findings by Goodman (2001), which demonstrated similar coefficients for each subscale (α = .57–.82) and stability at 6 months (α = .57–.72). Internal consistency for the current sample is reported in Table 4.

Brief Problem Monitor. The BPM by Achenbach and colleagues (2011) is a 19-item measure with separate parent-report (BPM-P) and self-report options for youth aged 11 to 18 years (BPM-Y). Both result in internalizing, externalizing, and attention subscale scores and a total score (range 0–38). The BPM-P and BPM-Y consist of similarly written statements with slight changes for each informant. For example: "Feels worthless or inferior" and "Disobedient at home" for parents become "I feel worthless or inferior" and "I disobey my parents" for youth informants. Both versions utilize three response options (0: *not true*; 1: *somewhat true*; 2: *certainly true*). The BPM is an abbreviated version of the well-validated and widely utilized 113-item Child Behavior Checklist (Piper et al., 2014). Prior studies have reported test–retest reliability correlations and internal consistencies for the BPM-P total score as r = .85, p < .001, α = .92, BPM-Y total score as r = .89, p < .001, α = .86, and mean Cronbach's α coefficients across the subscales as .84 (BPM-P) and .76 (BPM-Y; Achenbach et al., 2011; Piper et al., 2014). Internal consistency for the current sample is reported in Table 4.

Psychometric Analyses and Results

Factor Structure

Exploratory factor analysis (EFA) was conducted on the 15 items of the ICDS. With a sample size of 266 participants, our item to participant ratio was 1:17. This indicates the sample was sufficient for EFA (Comrey, 1988; Hoe, 2008). Violations of univariate and multivariate normality were apparent. Skew and kurtosis exceeded conventional cutoffs ($>\pm1.5$) on six items (1, 5, 8, 9, 10, and 14), indicating violations to univariate normality (Field et al., 2012). Multivariate tests for skewness and kurtosis proposed by Mardia (1970) revealed that while skew was nonsignificant (m_{skewness} 109.46, $p > .999$), there was evidence of excessive multivariate kurtosis (m_{kurtosis} 407.56, $p < .001$). Given the high multivariate kurtosis and the dichotomous response scale, we ran factor analyses using a tetrachoric correlation matrix and applied a maximum likelihood estimator using Mplus. The Mplus MLR estimator was specifically employed as it provides robust standard errors and overcomes such distribution asymmetries (Muthén & Muthén, 2012). Regarding the factorability of the correlation matrix, the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy was high (.81), and Bartlett's Test of Sphericity was significant 702.30 (df = 105; $p < .001$), indicating that the data were suitable for factor analyses and for yielding distinct and reliable factors (Field et al., 2012).

Taken in isolation, the scree test notoriously suffers from "subjectivity and ambiguity" (Hayton et al., 2004). However, we conducted parallel analysis and plotted the scree of the obtained eigenvalues against those yielded by a reduced correlation matrix for simulated variables with population correlations of 0 (i.e., no common factors). The results for the full 15 items are displayed in Figure 2 and suggest a two-factor solution (m = 2).

A similar interpretation is supported by the parallel analysis in conjunction with the scree plot for the reduced 12-item set (omitting Items 8, 13, and 14) as shown in Figure 2. These obtained factors represent the final factor solution that was reported.

Kaiser's criterion suggested up to four factors to extract (eigenvalues >1.00); however, it is notably lenient (Floyd & Widaman, 1995). Therefore, model fit indices for one- to

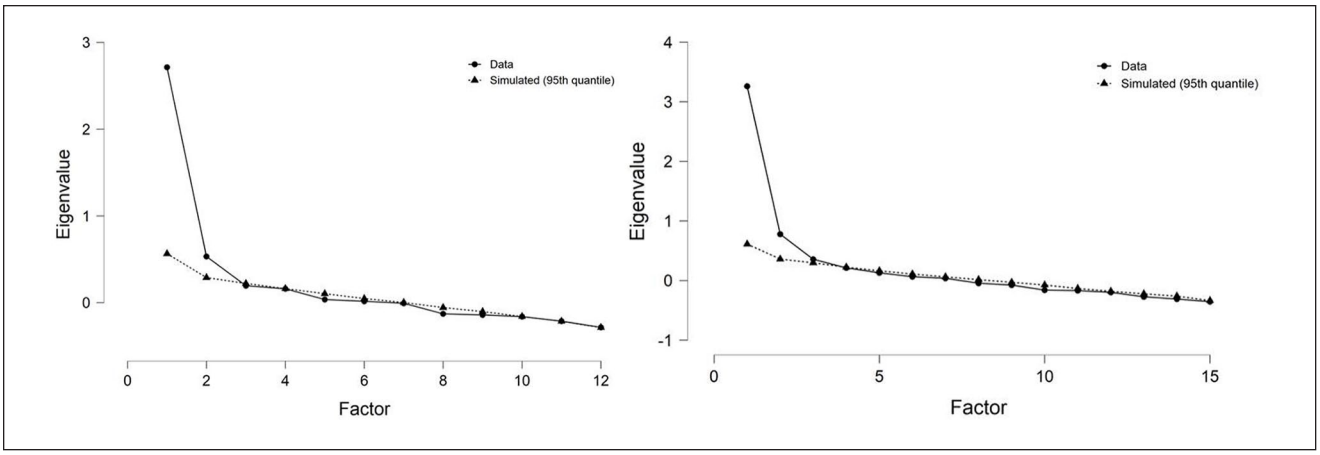

Figure 2. ICDS 15-Item and 12-Item Scree Plots.
Note. ICDS = *Interactive Child Distress Screener.*

Table 2. Model Fit for 1- to 4-Factor Solutions.

Solution	χ^2	$\Delta\chi^2$	RMSEA	CFI	SRMR
1-factor	165.65*	—	.06	.80	.07
2-factor	98.53	57.72*	.03	.94	.05
3-factor	66.90	27.67*	.02	.99	.04
4-factor	63.52	9.33	.03	.97	.03

Note. RMSEA = root mean square error of approximation; CFI = comparative fit index; SRMR = standardized root mean square residual.
*$p < .001$.

four-factor solutions were appraised. These indices are reported in Table 2 and as indicated by the results; the three-factor solution yields the best model fit. A rotation approach was selected on both theoretical and empirical grounds.

Theoretically, any underlying factors are expected to overlap consistent with established accounts of mental distress factors. Empirically, an inspection of the correlation between factors revealed they were moderately correlated ($r > .30$). Both of these considerations led to the decision to use an oblique geomin rotation and a three-factor solution was estimated. An inspection of the loading matrix showed a clear structure with a couple of exceptions. Item 14 (assessing physical aggression) exhibited weak cross-loadings across the three factors (.16, .11, and .21), and Factor 3 comprised only two items (8—hyperactive behavior; 13—distracted—inattentive behavior). Therefore, the decision was made to remove all three items and reserve them for further item development in the future.

Following the removal of Items 14 (physical aggression), 8 (hyperactive behavior), and 13 (distracted—inattentive behavior), a final two-factor model was extracted using the remaining 12 items. The final solution exhibited excellent model fit, $\chi^2(df = 43), 37.68, p = .701$; root mean square error of approximation (RMSEA) $< .01$, $p = .998$; 90% CI [<.00, .03]; comparative fit index (CFI) = 1.00; Tucker–Lewis index (TLI) = 1.02; and standardized root mean square residual (SRMR)= .03. Geomin rotated loadings for the final 12 items are reported in Table 3. The two remaining factors were strongly correlated, $r = .52$. Two alternative statistical methods were also used to inform the ideal number of factors to extract: Schwartz's Bayesian Information Criterion (BIC) dimensionality test (Schwarz, 1978) and the Hull method for selecting the number of common factors (Lorenzo-Seva et al., 2011). These analyses further confirmed that a two-factor solution best fit the data.

The final interpretation of the factors is consistent with conventional facets of internalizing and externalizing difficulties with Factor 1 exhibiting high loadings for items representing emotional distress and Factor 2 demonstrating high loadings for items representing behavioral concerns. The implied measurement model is depicted in Figure 3.

Internal Consistency

To estimate the internal consistency of the factors, we calculated Cronbach's alpha as well as ordinal omega coefficients using the tetrachoric correlation matrix. This method provides a more accurate estimate of reliability for dichotomous variables (Gadermann et al., 2012). Reliability estimates demonstrated excellent internal consistency for

Table 3. Geomin Rotated Factor Loadings on Final 12-Item EFA.

	Factor		
Items	1	2	Dimensions
ICDS_1 Sadness	**.512***	.122	Emotional
ICDS_2 Worry	**.534***	.017	
ICDS_3 Sleep difficulties	**.329***	.119	
ICDS_6 Shyness	**.491***	.016	
ICDS_9 Loneliness	**.451***	.006	
ICDS_10 Bullied/Excluded	**.775***	−.266	
ICDS_11 Fearful	**.539***	−.166	
ICDS_15 Physical symptoms	**.492***	.101	
ICDS_4 Anger	.072	**.510***	Behavioral
ICDS_5 Disobedience (School)	−.008	**.575***	
ICDS_7 Argumentativeness	.238	**.407***	
ICDS_12 Disobedience (Home)	.079	**.428***	

Note. Matrix: Tetrachoric correlations, Extraction: MLR, Rotation: Geomin. Loadings larger than .30 are in bold.
*Significant at 5% level.

Table 4. Reliability Statistics for Comparative Subscales on all Measures (Cronbach's α) (N = 266).

Scale	ICDS	M&MS	BPM-Y	BPM-P	SDQ-P
Behavioral	.84[a] .87[b]	.79	.73	.85	.86
Emotional	.88[a] .91[b]	.85	.74	.85	.80
Total Score	.90[a] .93[b]	.89	.84	.90	.88

Note. ICDS = Interactive Child Distress Screener; M&MS = Me and My School Questionnaire; BPM-Y = Brief Problem Monitor–Youth; BPM-P = Brief Problem Monitor–Parent; SDQ-P = Strengths and Difficulties Questionnaire–Parent.
[a]Indicates Cronbach's alpha statistic. [b] Indicates ordinal omega statistic.

Factor 1 (emotional distress items; α = .88 and ω = .91) and Factor 2 (behavioral distress items; α = .84 and ω = .87). The reliability estimates for each measure used in the study are shown in Table 4.

Construct-Related Validity

To assess the convergent validity of the ICDS, we correlated the ICDS subscale scores with the corresponding scales of the child-reported M&MS and BPM-Y, and parent-reported BPM-P and SDQ-P. These results are presented in Table 5. For the whole sample, positive correlations were found at the .001 alpha level for the behavioral subscale (r = .36–.50) and emotional subscale scores (r = .31–.58). Associations were larger in magnitude between the ICDS scores and youth-reported measures than between the ICDS scores and parent-reported measures across all subscales and age groups with one exception (BPM-Y < BMP-P behavioral subscale in the youngest age group). Given that some of the measures were used with participants outside their intended age range, correlations were also explored by age subgroups. We were also interested in the patterns of correlations between the ICDS and parent-reported measures across different age groups.

For the 5- to 7-year-old sub-sample of children, moderately strong, positive correlations were found across all youth completed subscales. Correlations with the ICDS were weaker for parent-completed subscales in this age group, particularly for the emotion-focused subscales. Correlations with the M&MS were significant at the .001 level for both ICDS subscales, while the emotion-focused subscale was not significantly correlated with the SDQ-P for this group. Moderately strong to strong significant, positive correlations were found for the 8- to 10-year-old subsample on all measures. For the 11-year-old subsample of children, strong correlations were found for both subscales on youth completed measures, while emotion-focused subscales showed moderately strong correlations with parent completed measures. Correlations between the behavioral subscale and the parent-reported measures were weaker and not significant for the BPM-P measure. The complete correlation table for the whole sample is shown in Table 6.

User Acceptability and Satisfaction

The acceptability of the ICDS was determined by examining preference survey responses and written feedback from 136 children (53.7% male, M_{age} = 7.65 years, SD = 1.93).

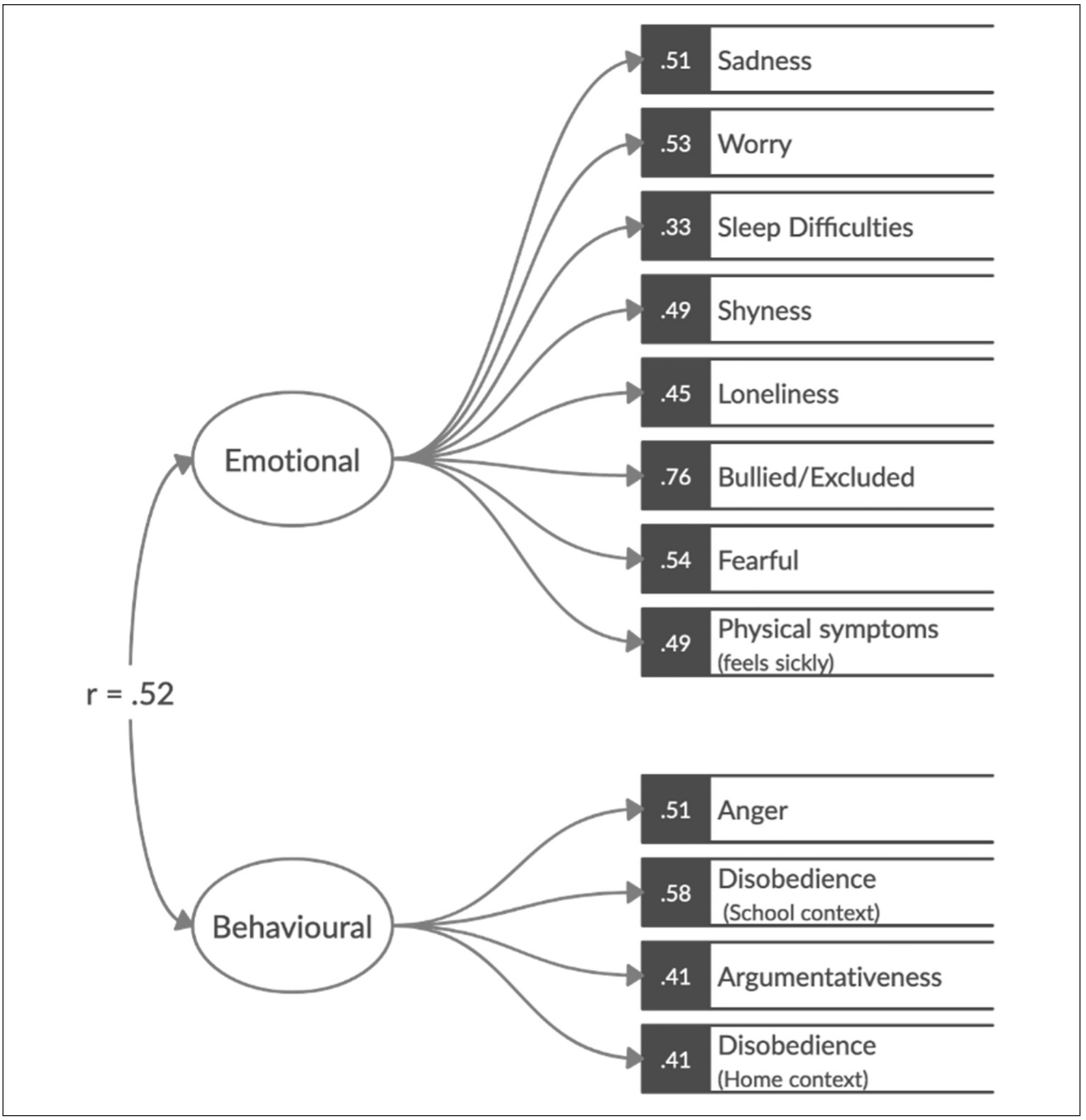

Figure 3. Correlated Two-Factor Model With Items 8 "Hyperactive," 13 "Distracted," and 14 "Physical Aggression" Deleted.

Data presented in Figure 4 indicate that the digital presentation mode of the ICDS is highly acceptable when compared with the written format of the M&MS. Across the whole sample and within age-group levels (5–7, 8–10, and 11 years), at least 75% of children in each age group stated they would recommend the ICDS to other children. When asked which measure they would prefer to do again, results across the whole sample were mixed with 44.9% preferring to complete the ICDS, 33.4% preferring the M&MS, and 22.8% stating they had no preference. Comprehension was high with 72.8% of all participants reporting that the ICDS was easier to understand than the M&MS. Those who had no preference stated it was because they thought both measures were easy to understand. Overall, at least 80% of children in each age group stated they preferred the ICDS over the M&MS.

Table 5. Correlations Between the ICDS Factors and Equivalent Subscales on M&MS, BPM-Y, BPM-P, and SDQ-P as a Function of Age-Group Levels.

	Youth completed		Parent completed	
ICDS factors (*N*)	M&MS	BPM-Y	BPM-P	SDQ-P
Full sample (*N* = 266)	*n* = 266	*n* = 103	*n* = 265	*n* = 258
Behavioral	.498***	.484***	.362***	.401***
Emotional	.470***	.575***	.348***	.305***
Total score	.494***	.587***	.372***	.393***
5–7 years (*N* = 127)	*n* = 127	*n* = 43	*n* = 126	*n* = 125
Behavioral	.415***	.317*	.320***	.407***
Emotional	.370***	.463**	.223*	.138
Total score	.394***	.503***	.227*	.276***
8–10 years (*N* = 103)	*n* = 103	*n* = 25	*n* = 103	*n* = 99
Behavioral	.562***	.549**	.424***	.441***
Emotional	.562***	.729***	.471***	.446***
Total score	.582***	.736***	.501***	.508***
11 years (*N* = 36)	*n* = 36	*n* = 35	*n* = 36	*n* = 35
Behavioral	.672***	.704***	.333*	.281
Emotional	.608***	.625***	.463**	.498**
Total score	.678***	.632***	.532***	.526**

Note. Equivalent M&MS subscales = behavioral and emotional; BPM subscales = externalizing and internalizing; SDQ subscales = conduct and emotional; ICDS = Interactive Child Distress Screener; M&MS = Me and My School Questionnaire; BPM-Y = Brief Problem Monitor–Youth; BPM-P = Brief Problem Monitor-Parent; SDQ-P= Strengths and Difficulties Questionnaire–Parent.
*$p < .05$. **$p < .01$. ***$p < .001$.

The satisfaction ratings of the ICDS were high across the whole sample with respect to the instructions and response options. Younger children <8 years ($n = 68$) reported the highest satisfaction with the characters (83.8%) and with how fun the ICDS was (88.2%). While the majority of 11-year-old children (86.7%) were satisfied with how "fun" the ICDS was, less were satisfied with its length (73.3%) and two thirds were satisfied with the cartoon characters (66.7%). Satisfaction ratings for the whole sample and per age-group level are presented in Figure 5. Nine parents and 73 child participants provided predominantly positive qualitative feedback on the ICDS describing their experiences using the application and their engagement with the animated videos. Qualitative data may be requested from the authors.

ICDS Utility

All participants ($N = 74$, 54% male, $M_{age} = 7.43$ years, $SD = 1.90$) seemed accustomed with using an iPad as none of the participants required help to use it. Zero children who were above 7 years (62.2%) required any assistance to complete any section of the ICDS. Of the seventy-nine 5- and 6-year-old children, only nine (11.4%) required some form of help at least one time. Specifically, six of these children (50% male) asked for help to complete the demographic portion of the ICDS application (i.e., required direction to click on the correct numeral and to click on a cartoon image of a boy or girl to choose their gender) and two boys clicked on the in-app "Buddy" helper assistant at least once. Following this, the same boys verbally asked, "What do I do now?" and required prompting to click the "next" arrow and "play" button to begin the next cartoon animation. One boy (aged 5 years) clicked the "Buddy" helper for assistance when completing the response portion of the questionnaire to respond to the question "Which one is like you?" and one 6-year-old boy asked how to re-watch a video. Overall, the ICDS application was highly functional with 88.6% of participants able to use the ICDS accurately and without any assistance.

Discussion

In contrast to standard text-based measures, the ICDS offers a novel tool that overcomes many of the shortcomings of existing instruments by offering a digital presentation of item content via web application on internet-enabled devices and is optimized for smartphones and handheld tablets. The ICDS utilizes contrasting pairs of audio-visual animations that serve as both the stimulus and the response options and is intended for children as young as 5 years of age. The current analysis first scrutinized the factor structure of 15 items via iterative EFA, which yielded three initial factors. However, Item 14 (physical aggression) performed poorly due to cross-loadings and Factor 3 contained only two items. Item 14 is likely to have performed poorly due to content issues. This animation depicted one

Table 6. Correlations Between ICDS and M&MS, BPM-Y, BPM-P, and SDQ-P Subscales and Total Scores Across the Whole Sample (N = 266).

Scale	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. ICDS Beh	—															
2. ICDS Emo	.422***	—														
3. ICDS Tot	.719***	.934***	—													
4. MMS Beh	.498***	.279***	.411***	—												
5. MMS Emo	.284***	.470***	.473***	.631***	—											
6. MMS Tot	.402***	.437***	.494***	.849***	.946***	—										
7. BPMY Ext	.484***	.416***	.499***	.648***	.455***	.590***	—									
8. BPMY Int	.189	.575***	.528***	.410***	.797***	.754***	.520***	—								
9. BPMY Att	.264**	.401***	.361***	.423***	.494***	.532***	.488***	.514***	—							
10. BPMY Tot	.378***	.570***	.587***	.600***	.716***	.767***	.814***	.837***	.808***	—						
11. BPMP Ext	.362***	.177**	.279***	.512***	.218***	.363***	.277**	.173	.210*	.267**	—					
12. BPMP Int	.173**	.348***	.335***	.332***	.459***	.452***	.214*	.332***	.171	.294**	.404***	—				
13. BPMP Att	.343***	.202***	.291***	.400***	.283***	.360***	.184	.210*	.373***	.309**	.624***	.374***	—			
14. BPMP Tot	.369***	.295***	.372***	.518***	.390***	.482***	.283**	.298**	.326***	.368***	.851***	.713***	.844***	—		
15. SDQP Ext	.409***	.233***	.341***	.510***	.353***	.454***	.265**	.205*	.434***	.366***	.732***	.358***	.859***	.823***	—	
16. SDQP Int	.251***	.319***	.344***	.404***	.519***	.522***	.184	.290**	.168	.264**	.428***	.760***	.470***	.675***	.519***	—
17. SDQP Tot	.385***	.313***	.393***	.529***	.494***	.558***	.263**	.287**	.357***	.370***	.678***	.623***	.779***	.865***	.891***	.850***

Note. Sample size for ICDS and M&MS $N = 266$, BPM-Y $N = 100$, BPM-P $N = 265$, and SDQ-P $N = 258$. Beh = behavioral subscale; Emo = emotional subscale; Tot = total scale score; Ext = externalizing subscale; Int = internalizing subscale; Att = Attention subscale; ICDS = Interactive Child Distress Screener; M&MS = Me and My School Questionnaire; BPM-Y = Brief Problem Monitor-Youth; BPM-P = Brief Problem Monitor–Parent; SDQ-P = Strengths and Difficulties Questionnaire–Parent.
*$p < .05$. ** $p < .01$. *** $p < .001$.

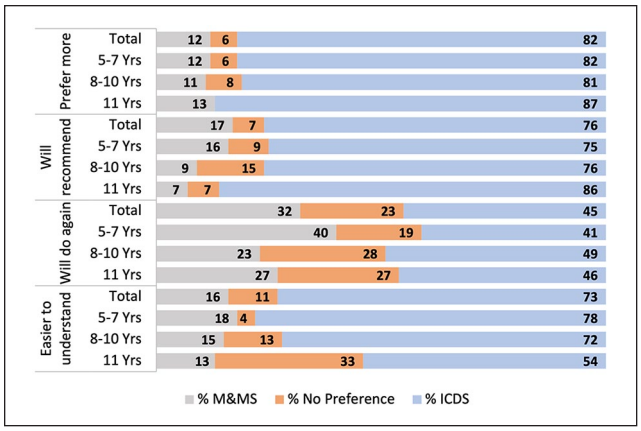

Figure 4. Participant Preference Rating for ICDS and M&MS Measures.
Note. ICDS = Interactive Child Distress Screener; M&MS = Me and My School Questionnaire.

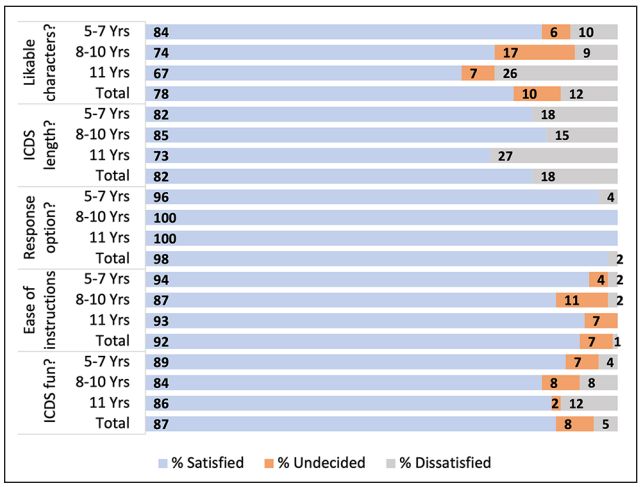

Figure 5. Participant Satisfaction With ICDS.
Note. ICDS = Interactive Child Distress Screener.

child being physically aggressive to another child and it may have been confusing for children as to which actor they were meant to identify with (i.e., aggressor or victim). The two-item third factor depicted attention-related problems that are likely to be relevant for children with a diagnosis of attention-deficit disorder. That the EFA modeled these items as a distinct factor suggests that these types of difficulties are distinct although overlapping with emotional and behavioral issues. The lack of a robust third factor meeting minimal criteria of at least 3 items highlights the need for additional item development for this subdomain to properly capture attention/hyperactivity problems. After the removal of these three items, a clean two-factor solution was supported. The final factor solution fit the data well and produced two interpretable, internally consistent, and correlated factors representing emotional and behavioral difficulties. These results are promising given this is the first attempt to examine child self-reported responses to animated video items via factor analyses.

The converging construct validity of the two ICDS factors was also supported by the pattern of correlations with previously validated measures. Overall, correlations examining the construct validity of the ICDS revealed a pattern of stronger associations as a function of participant age and there was greater convergence for youth report (versus parent-reported) responses at each age group. This pattern is not unique to the ICDS and likely due to well-known difficulties in capturing variation via psychological constructs with younger children. While self-reported information is known to be more challenging to collect from younger children (<11 years), proxy reports from parents and teachers may also be unreliable and inconsistent with child reports, especially for internalizing information (Jardine et al., 2014). In light of such cross-informant variance, child-reported data clearly need to be given due consideration when making diagnostic or treatment decisions, and to do so, valid and reliable child-report instruments are needed.

The results of this study also establish that the digitally animated format of the ICDS instrument had high acceptability to school-aged children. That is, they liked it, understood it, mostly preferred it to pen-and-paper measures, and indicated they would recommend it to others. Satisfaction ratings were exceptional with most children within each age-group level rating the ICDS favorably on both satisfaction and preference questions. Utility results further demonstrated that the design of the ICDS was highly functional and that the digital format was straightforward for even the youngest children to complete on their own. Given the growing digital literacy skills of today's children and their widespread use of such devices, open access to an engaging and innovative digital assessment instrument such as the ICDS is a viable option for universal application.

The present study utilized a convenience sample of parents from one school in the Brisbane region along with an online sample. Although online recruitment did increase the geographical variation in our participant pool, this group was self-selecting, which may introduce some sampling bias regarding the acceptability or utility data. Furthermore, some respondents were asked to provide self-reported data for measures that were outside the intended age range for those instruments. While this may have increased measurement error and possibly attenuated the observed correlations for the younger age groups, this approach was necessary as there were no validated self-report measures available for corroborating child-report scores across the entire sample age range. It is possible that some children's answers may have been influenced by the presence of a parent; however, the fact that 58% of the sample indicated scores that exceeded elevated cut-offs on at least one or

more sub-scales suggests that child participants as a group were able to report on distress. Ultimately, the pattern of correlations within the youngest age group does not deviate substantively in meaningfulness or interpretability from that observed with the oldest age group (although associations are consistently weaker, the pattern converges). This provides some evidence that administering measures to younger-than-recommended cohorts did not overly influence our results. Second, the greater convergence between child-reported responses versus parent and child responses suggests that the possible influence of parents on children's responses also did not overly influence the pattern of results. Despite these issues, the test battery was varied, incorporated both parent and child informants, and utilized well-established self-report scales for determining convergent validity.

Future research will focus on the development of additional items to examine the potential for a third factor focused on attention and hyperactivity as well as confirming the factor structure of the emotional and behavioral factors established here. While this study demonstrated initial promising psychometric properties of this tool, a critical next step will be to evaluate its utility as a screening instrument and its sensitivity to detect emotional and behavioral problems. Thus, the capacity of the ICDS to differentiate between clinical and nonclinical children should be examined to establish clinical norms and cut-off scores to assist with prevention, intervention, and treatment planning. Pre-post reliability and sensitivity of the ICDS to change following intervention is another area of examination required. With respect to the ICDS design and feedback from the older children who were less favorable about the look of the animated characters, the research team has commenced modification of the visual style of the animations to produce a more suitable version for older age groups.

In conclusion, this study aimed to examine the psychometric properties of the ICDS, a digital, animation-based instrument for detecting emotional and behavioral difficulties in children. The ICDS revealed good overall psychometric properties, with a clear two-factor structure, excellent internal consistency, and good construct validity. Furthermore, the digital instrument demonstrated high utility and satisfaction ratings, meaning children understood and enjoyed using it. Given that the ICDS was developed through a series of participatory co-design studies with young people, the instrument is likely to be more effectively implemented and accepted by this population. The ICDS instrument appears to present a promising opportunity for obtaining reliable information from young children under the age of 11 themselves regarding emotional and behavioral difficulties. The prevalence of mental health needs for children worryingly outpaces access to care so it is important to prevent delays to treatment. Universal screening can achieve early identification of problems, alter the trajectory of disorder development and minimize social, emotional, and economic burden.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by an Australian Government Research Training Scheme scholarship.

ORCID iDs

Kirsty Zieschank https://orcid.org/0000-0002-9707-6647
Jamin Day https://orcid.org/0000-0003-4277-4816
Sonja March https://orcid.org/0000-0001-8425-7126

References

Achenbach, T. M., McConaughy, S. H., Ivanova, M. Y., & Rescorla, L. A. (2011). *Manual for the ASEBA brief problem monitor (BPM)* (pp. 1–33). University of Vermont.

Arseneault, L., Kim-Cohen, J., Taylor, A., Caspi, A., & Moffitt, T. E. (2005). Psychometric evaluation of 5- and 7-year-old children's self-reports of conduct problems. *Journal of Abnormal Child Psychology*, *33*(5), 537–550. https://doi.org/10.1007/s10802-005-6736-5

Bevans, K. B., Ahuvia, I. L., Hallock, T. M., Mendonca, R., Roth, S., Forrest, C. B., Blackwell, C., Kramer, J., & Wakschlag, L. (2020). Investigating child self-report capacity: A systematic review and utility analysis. *Quality of Life Research*, *29*(5), 1147–1158. https://doi.org/10.1007/s11136-019-02387-3

Bowen, N. K. (2008). Cognitive testing and the validity of child-report data from the Elementary School Success Profile. *Social Work Research*, *32*(1), 18–28. https://doi.org/10.1093/swr/32.1.18

Carter, A. S., Briggs-Gowan, M. J., & Davis, N. O. (2004). Assessment of young children's social-emotional development and psychopathology: Recent advances and recommendations for practice. *Journal of Child Psychology and Psychiatry*, *45*(1), 109–134. https://doi.org/10.1046/j.0021-9630.2003.00316.x

Centers for Disease Control and Prevention. (2020). *Data and statistics on children's mental health*. https://www.cdc.gov/childrensmentalhealth/data.html

Chambers, C. T., & Johnston, C. (2002). Developmental differences in children's use of rating scales. *Journal of Pediatric Psychology*, *27*(1), 27–36. https://doi.org/10.1093/jpepsy/27.1.27

Comrey, A. L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology*, *56*(5), 754–761. https://doi.org/10.1037/0022-006X.56.5.754

Deighton, J., Croudace, T., Fonagy, P., Brown, J., Patalay, P., & Wolpert, M. (2014). Measuring mental health and wellbeing

outcomes for children and adolescents to inform practice and policy: A review of child self-report measures. *Child and Adolescent Psychiatry and Mental Health*, *8*(1), Article 14. https://doi.org/10.1186/1753-2000-8-14

Deighton, J., Lereya, S. T., Casey, P., Patalay, P., Humphrey, N., & Wolpert, M. (2019). Prevalence of mental health problems in schools: Poverty and other risk factors among 28 000 adolescents in England. *British Journal of Psychiatry*, *215*(3), 565–567. https://doi.org/10.1192/bjp.2019.19

Deighton, J., Tymms, P., Vostanis, P., Belsky, J., Fonagy, P., Brown, A., Martin, A., Patalay, P., & Wolpert, M. (2013). The development of a school-based measure of child mental health. *Journal of Psychoeducational Assessment*, *31*(3), 247–257. https://doi.org/10.1177/0734282912465570

Dowdy, E., Ritchey, K., & Kamphaus, R. W. (2010). School-based screening: A population-based approach to inform and monitor children's mental health needs. *School Mental Health*, *2*(4), 166–176. https://doi.org/10.1007/s12310-010-9036-3

Droit-Volet, S., & Coull, J. (2015). The developmental emergence of the mental time-line: Spatial and numerical distortion of time judgement. *PLOS ONE*, *10*(7), 1–20. https://doi.org/10.1371/journal.pone.0130465

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics. Discovering statistics Using R*. SAGE.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7*(3), 286–299. https://doi.org/10.1037/1040-3590.7.3.286

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research and Evaluation*, *17*(3), 1–13. https://doi.org/10.7275/n560-j767

Gardner, W., Murphy, M., Childs, G., Kelleher, K., Pagano, M., Jellinek, M., McInerny, T. K., Wasserman, R. C., Nutting, P., Chiappetta, L., & Sturner, R. (1999). The PSC-17: A brief pediatric symptom checklist with psychosocial problem subscales. A report from PROS and ASPN. *Ambulatory Child Health*, *5*(3), 225–236.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*(11), 1337–1345. https://doi.org/10.1097/00004583-200111000-00015

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Goodman-Scott, E., Donohue, P., & Betters-Bubon, J. (2019). The case for universal mental health screening in schools. *Counseling Today*, *62*(3), 40–45.

Guzman, M. P., Jellinek, M., George, M., Hartley, M., Squicciarini, A. M., Canenguez, K. M., Kuhlthau, K. A., Yucel, R., White, G. W., Guzman, J., & Murphy, J. M. (2011). Mental health matters in elementary school: First-grade screening predicts fourth grade achievement test scores. *European Child and Adolescent Psychiatry*, *20*(8), 401–411. https://doi.org/10.1007/s00787-011-0191-3

Hawes, D. J., & Dadds, M. R. (2004). Australian data and psychometric properties of the Strengths and Difficulties Questionnaire. *Australian and New Zealand Journal of Psychiatry*, *38*(8), 644–651. https://doi.org/10.1111/j.1440-1614.2004.01427.x

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, *7*(2), 191–205. https://doi.org/10.1177/1094428104263675

Hoe, S. L. (2008). Issues and procedures in adopting structural equation modeling technique. *Journal of Applied Quantitative Methods*, *3*(1), 76–83.

Humphrey, N., & Wigelsworth, M. (2016). Making the case for universal school-based mental health screening. *Emotional and Behavioural Difficulties*, *21*(1), 22–42. https://doi.org/10.1080/13632752.2015.1120051

Ivey, J. (2020). Mental health screening for children and adolescents. *Pediatric Nursing*, *46*(1), 27–31.

Jardine, J., Glinianaia, S. V., McConachie, H., Embleton, N. D., & Rankin, J. (2014). Self-reported quality of life of young children with conditions from early infancy: A systematic review. *Pediatrics*, *134*(4), e1129–e1148. https://doi.org/10.1542/peds.2014-0352

Jeffrey, J., Klomhaus, A., Enenbach, M., Lester, P., & Krishna, R. (2020). Self-report rating scales to guide measurement-based care in child and adolescent psychiatry. *Child and Adolescent Psychiatric Clinics of North America*, *29*(4), 601–629. https://doi.org/10.1016/j.chc.2020.06.002

Kamphaus, M. C., & Reynolds, C. R. (2007). *Behavior assessment system for children – second edition (BASC-2): Behavioral and emotional screening system (BESS)*. Bloomington, MN: Pearson.

Lawrence, D., Hafekost, J., Johnson, S. E., Saw, S., Buckingham, W. J., Sawyer, M. G., Ainley, J., & Zubrick, S. R. (2016). Key findings from the second Australian child and adolescent survey of mental health and wellbeing. *Australian & New Zealand Journal of Psychiatry*, *50*(9), 876–886. https://doi.org/10.1177/0004867415617836

Lorenzo-Seva, U., Timmerman, M. E., & Kiers, H. A. L. (2011). The Hull method for selecting the number of common factors. *Multivariate Behavioral Research*, *46*, 340–364. https://doi.org/10.1080/00273171.2011.564527

Manassis, K., Mendlowitz, S., Kreindler, D., Lumsden, C., Sharpe, J., Simon, M. D., Woolridge, N., Monga, S., & Adler-Nevo, G. (2009). Mood assessment via animated characters: A novel instrument to evaluate feelings in young children with anxiety disorders. *Journal of Clinical Child and Adolescent Psychology*, *38*(3), 380–389. https://doi.org/10.1080/15374410902851655

March, S., Day, J., Zieschank, K., & Ireland, M. (2018). The interactive child distress screener: Development and preliminary feasibility testing. *JMIR MHealth and UHealth*, *6*(4), Article e90. https://doi.org/10.2196/mhealth.9456

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*(3), 519–530. https://doi.org/10.2307/2334770

McCormack, T., & Hoerl, C. (2017). The development of temporal concepts: Learning to locate events in time. *Timing*

and Time Perception, *5*(3–4), 297–327. https://doi.org/10.1163/22134468-00002094

McCrae, J. S., & Brown, S. M. (2018). Systematic review of social–emotional screening instruments for young children in child welfare. *Research on Social Work Practice*, *28*(7), 767–788. https://doi.org/10.1177/1049731516686691

Merikangas, K. R., He, J., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., & Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the national comorbidity survey replication–Adolescent supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(10), 980–989. https://doi.org/10.1016/j.jaac.2010.05.017

Mihalopoulos, C., Vos, T., Pirkis, J., & Carter, R. (2012). The population cost-effectiveness of interventions designed to prevent childhood depression. *Pediatrics*, *129*(3), Article e723. https://doi.org/10.1542/peds.2011-1823

Moffa, K., Wagle, R., Dowdy, E., Palikara, O., Castro, S., Dougherty, D., & Furlong, M. J. (2019). The Me and My School Questionnaire: Examining the cross-cultural validity of a children's self-report mental health measure. *International Journal of School and Educational Psychology*, *9*(1), 1–11. https://doi.org/10.1080/21683603.2019.1650858

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide* (7th ed.) Muthén & Muthén.

National Research Council and Institute of Medicine. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities* (K. E. Warner, T. Boat, & M. E. O'Connell, eds.). The National Academies Press. https://doi.org/10.17226/12480

Newlove-Delgado, T., & Ford, T. J. (2020). Screening methods and when to use them. In E. Taylor, F. C. Verhulst, J. Wong, K. Yoshida, & A. Nikapota (Eds.), *Mental health and illness of children and adolescents* (pp. 17–38). Springer Nature Singapore. https://doi.org/10.1007/978-981-10-2348-4_7

Ogundele, M. O. (2018). Behavioural and emotional disorders in childhood: A brief overview for paediatricians. *World Journal of Clinical Pediatrics*, *7*(1), 9–26. https://doi.org/10.5409/wjcp.v7.i1.9

Patalay, P., Deighton, J., Fonagy, P., Vostanis, P., & Wolpert, M. (2014). Clinical validity of the Me and My School questionnaire: A self-report mental health measure for children and adolescents. *Child and Adolescent Psychiatry and Mental Health*, *8*(17), 1–7. https://doi.org/10.1186/1753-2000-8-17

Piper, B. J., Gray, H. M., Raber, J., & Birkett, M. A. (2014). Reliability and validity of brief problem monitor, an abbreviated form of the child behavior checklist. *Psychiatry and Clinical Neurosciences*, *68*(10), 759–767. https://doi.org/10.1111/pcn.12188

Ringoot, A. P., Jansen, P. W., Rijlaarsdam, J., So, P., Jaddoe, V. W. V., Verhulst, F. C., & Tiemeier, H. (2017). Self-reported problem behavior in young children with and without a DSM-disorder in the general population. *European Psychiatry*, *40*, 110–115. https://doi.org/10.1016/j.eurpsy.2016.08.009

The Royal Australian and New Zealand College of Psychiatrists. (2010). *Prevention and early intervention of mental illness in infants, children and adolescents: Planning strategies for Australia and New Zealand*. https://www.ranzcp.org/news-policy/policy-and-advocacy/position-statements/the-prevention-and-early-intervention-of-mental-il

Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464. https://https://doi.org/10.1214/aos/1176344136

Truman, J., Robinson, K., Evans, A. L., Smith, D., Cunningham, L., Millward, R., & Minnis, H. (2003). The Strengths and Difficulties Questionnaire: A pilot study of a new computer version of the self-report scale. *European Child and Adolescent Psychiatry*, *12*(1), 9–14. https://doi.org/10.1007/s00787-003-0303-9

Whitney, D. G., & Peterson, M. D. (2019). US national and state-level prevalence of mental health disorders and disparities of mental health care use in children. *JAMA Pediatrics*, *173*(4), 389–391. https://doi.org/10.1001/jamapediatrics.2018.5399

Wood, B. J., & McDaniel, T. (2020). A preliminary investigation of universal mental health screening practices in schools. *Children and Youth Services Review*, *112*(March), Article 104943. https://doi.org/10.1016/j.childyouth.2020.104943

Zieschank, K., Machin, T., Day, J., Ireland, M., & March, S. (2020). Children's perspectives on emotions informing a child-reported screening instrument. *Journal of Child and Family Studies*, *30*, 3105–3120. https://doi.org/10.1007/s10826-021-02086-z

Zieschank, K., Day, J., Ireland, M. J., & March, S. (2021). Co-design and qualitative validation of animated assessment item content for a child-reported digital distress screener. *Internet Interventions*, *24*, Article 100381. https://doi.org/10.1016/j.invent.2021.100381