Psychometric Properties of the Dyadic Parent-Child Interaction Coding System in The Netherlands

Mariëlle E. Abrahamse PhD, Larissa N. Niec PhD, David T. Solomon PhD, Marianne Junger PhD & Ramón J. L. Lindauer MD, PhD

Published online: 16 Jul 2019.

Submit your article to this journal ↗

Article views: 39

View related articles ↗

View Crossmark data ↗

Routledge
Taylor & Francis Group

Check for updates

Psychometric Properties of the Dyadic Parent-Child Interaction Coding System in The Netherlands

Mariëlle E. Abrahamse, PhD[a,b], Larissa N. Niec, PhD[c], David T. Solomon, PhD[d], Marianne Junger, PhD[e], and Ramón J. L. Lindauer, MD, PhD[a,b]

[a]Academic Center for Child and Adolescent Psychiatry, De Bascule, Amsterdam, The Netherlands; [b]Department of Child and Adolescent Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands; [c]Center for Children, Families, and Communities, Central Michigan University, Mount Pleasant, MI, USA; [d]Western Carolina University, Cullowhee, NC, USA; [e]Institute for Innovation and Governance Studies (IGS), School of Management & Governance, University of Twente, Enschede, The Netherlands

ABSTRACT

The current study explored the psychometric properties of a parent-child interaction observation system, the Dyadic Parent-Child Interaction Coding System (DPICS) in the Netherlands. Participants included 31 Dutch mother-child dyads and 86 U.S. mother-child dyads for a community sample (2–7 years; 50% boys). Good one-week test-retest reliability was demonstrated among the Dutch sample. Similarities were found between Dutch and U.S. samples on most interaction codes, but mothers in the U.S. sample used more directive behaviors (e.g., commands) in some situations. Findings suggest that the DPICS is a reliable measure of mother-child interactions in the Dutch population. Cultural issues regarding the use of the DPICS are discussed.

Parent management training (PMT) programs, which focus on parenting skills and behaviors as a means of changing child behavior (Kazdin, Siegel, & Bass, 1992), often emphasize consistent, systematic assessment of parent-child interactions across the span of treatment (Niec, 2018). The use of structured measures of parent-child interactions can guide the course of the intervention and measure treatment gains (Aspland & Gardner, 2003; Eyberg & Funderburk, 2011; Roberts & Hope, 2001). Because self-report questionnaires can be impacted by multiple types of bias, and because there are notable differences in scores on subjective report measures across family members (Barbosa, Tannock, & Manassis, 2002), direct-observation coding systems have been considered an important component in the evaluation of parent-child interactions (e.g., Gardner, 2000; Hops, Davis, & Longoria, 1995; Hupp, Reitman, Forde, Shriver, & Kelley, 2008). Intervention studies

show that observational measures of parent-child interaction are more predictive of child outcomes than self-report questionnaires (Patterson & Forgatch, 1995; Zaslow et al., 2006) and are assumed to be "less biased, more objective, and more sensitive" (Prasadarao & Kumaraiah, 1997, p. 278). However, this is only true when recording procedures are consistent, when coders are well-trained, and when the observation systems have a well-validated coding scheme. Furthermore, in order to establish clinical utility, observation systems need the ability to characterize clinical versus nonclinical samples (Reitman, Hummel, Franz, & Gross, 1998). Thus, well-validated assessment techniques utilizing observation of parent-child interactions can play a critical role in research and in the implementation of PMT programs (Pearl, 2009).

Certain parent-child interaction coding systems are being used with more diverse populations as recent years have seen an increase in PMT program research being conducted in other countries and cultures. For instance, recent research has investigated a particular PMT program, parent-child interaction therapy (PCIT), with Latino samples (Matos, Bauermeister, & Bernal, 2009; McCabe, Yeh, Lau, & Argote, 2012) and Chinese samples (Leung, Tsang, Sin, & Choi, 2015), as well as with samples in Australia (Nixon, Sweeney, Erickson, & Touyz, 2004), Germany (Briegel, Walter, Schimek, Knapp, & Bussing, 2015; Schimek, Walter, Bussing, & Briegel, 2014), Norway (Bjørseth & Wichstrøm, 2016), and The Netherlands (Abrahamse et al., 2012; Abrahamse, Junger, Van Wouwe, Boer, & Lindauer, 2016; Niec, 2018; Niec, Abrahamse, Egan, Coelman, & Heiner, 2018). Although these efforts represent a positive trend in disseminating effective parenting programs, the spread of these treatments necessitates further evaluation of the behavioral assessment techniques that are integral to their implementation within new cultural and geographic contexts. The psychometric properties of such assessment instruments may differ across cultural groups; for example, population means may differ (e.g., Kaplan, 1985; Mieloo et al., 2014; Rescorla et al., 2011). Also, cross-cultural differences in parenting practices based on differences in the use of sensitive parenting may affect the psychometric properties of assessment instruments (Mesman, van IJzendoorn, & Bakermans-Kranenburg, 2012).

Researchers have dealt with possible sources of cross-cultural assessment bias using a number of methods, including norming the instruments with new cultural samples (Cheung, Kwong, & Zhang, 2003) and using confirmatory factor analysis with new populations (Garcia-Barrera, Karr, Duran, Direnfeld, & Pineda, 2015), amongst other methods. Although recent research has investigated the psychometric properties of the Eyberg Child Behavior Inventory, a parent-report measure used in the implementation of PMT programs, within a Dutch sample (Abrahamse et al., 2015),

the properties of the Dyadic Parent-Child Interaction Coding System (DPICS; Eyberg, Nelson, Ginn, Bhuiyan, & Boggs, 2013), a parent-child interaction observation system used in both treatment and research, has yet to be investigated in that population.

A number of studies support the psychometric properties of the DPICS with samples from the United States (see Eyberg et al., 2013, for a review). The specific coding categories used in the current version have adequate inter-coder reliability and occur frequently enough for reliable coding. In addition, the DPICS has been found efficient for the screening of disruptive behavior disorders in children in a Norwegian sample (Bjørseth, McNeil, & Wichstrøm, 2015). Given the importance of reliable and valid behavioral observation of parent-child interactions in both the evaluation and the implementation of treatment, and given the prevalence of the use of DPICS in research and treatment (see, e.g., Borden et al., 2014; Niec, Shanley, Barnett, Baker, & Solomon, 2015; Thornberry & Brestan-Knight, 2011), further research with the DPICS in new populations is warranted.

The purpose of the present study was to examine the psychometric properties of the DPICS with a sample of parent-child dyads in The Netherlands. As the assessment of the psychometric properties of the DPICS previously included normative data for specific populations, test-retest reliability, and inter-coder reliability (Eyberg et al., 2013), we examined these psychometric properties in a community sample of Dutch families. Additionally, DPICS scores from a U.S. sample of parent-child dyads were compared to the Dutch sample to examine similarities and differences in interaction styles across cultures and to explore the value of the U.S. norms within The Netherlands. We expected that the robust psychometric properties of the DPICS would be maintained in The Netherlands, including, specifically, test-retest reliability and inter-coder reliability. Although the existing literature does not provide guidance on the similarities or differences to expect on the specific categories used in the DPICS, we hypothesized that the DPICS scores of the Dutch parents would reflect the authoritative parenting style, including autonomy-oriented behavior and emotional warmth, that is commonly found in Dutch parenting (Van der Bruggen, Stams, Bögels, & Paulussen-Hoogeboom, 2010). Therefore, we expected that Dutch parents would use fewer commands, questions, and criticisms during the interactions with their children than U.S. parents.

Methods

Participants and procedure

The present study included a sample of Dutch ($n = 31$) and U.S. ($n = 86$) parent-child dyads. Data from both samples were collected in separate

research studies and merged later for the current research purpose. For the Dutch study, the need for ethical approval was waived by the Medical Ethics Committee of the Academic Medical Center. The approval for the U.S. study was received from the university institutional review board. Informed consent was obtained from all individual participants included in the study.

Dutch sample

Families were recruited with informative flyers distributed to child daycare centers and local schools in Amsterdam, The Netherlands. Parents who were interested in participating contacted the researchers by e-mail or telephone and were subsequently screened during a telephone interview. For example, parents were asked if the common language used in interaction with their child was Dutch or English. The telephone interview also included the administration of a standardized parent rating scale for child disruptive behavior, the Eyberg Child Behavior Inventory (ECBI; Eyberg & Pincus, 1999). The inclusion criterion for participation in the study was a score below the clinical cutoff point on the ECBI Intensity Scale (<132). All families who were interested in participation met this criterion. Also, all parents interested in participation in the study were mothers.

After the telephone screening, mothers and children were visited at home to complete the DPICS observation as well as a number of questionnaires. After a week, the family was visited a second time to complete the DPICS once again. Families received a small compensation for their participation, including a present for the child and a €10 ($11) gift card for each assessment.

The Dutch sample included 31 mother-child dyads. Sample characteristics are presented in Table 1. Children's ages ranged from 2–7 years. Twenty-five (80.6%) children were reported as being of Dutch origin. Six (19.4%) children had one or two parents born in another country (Congo, France, Ghana, Nigeria, South Korea, or Surinam). Mothers' ages ranged from 29–50 years old. Most families were two-parent families: 45.2% of the families were married and 48.4% were unmarried living with a partner. Most families (79.4%) had an income higher than €30,000 ($34,000) and mothers were, in general, highly educated (Table 1).

U.S. sample

The U.S. families were recruited in a similar way as the Dutch sample. Flyers including study information were distributed at child day-care centers and schools in a Midwestern community in the United States. Parents who were interested in participating contacted researchers by telephone

Table 1. Demographic characteristics and mean scores per sample.

		Mean (SD) or Percentage			
		Dutch sample (n = 31)		US sample (n = 86)	
	n	M (SD) or %	n	M (SD) or %	p
Child gender (% boy)	31	64.5	86	45.3	.07
Child age (years)	31	4.35 (1.50)	86	4.41 (1.49)	.87
Child physical illness (% such as asthma or diabetes)	31	6.5	86	16.3	.17
Mother age	31	38.94 (4.68)	83	32.46 (5.64)	<.001*
Mother's education (% high school or less)	31	16.1	86	14.0	.77
Family status (% single parent)	31	6.5	86	14.0	.27
ECBI Intensity Scale	29	86.40 (23.46)	39	95.87 (21.55)	.09
ECBI Problem Scale	25	3.64 (4.32)	39	5.59 (5.22)	.13
CBCL / BASC-2 Internalizing (t-scores)	31	41.99 (8.29)[a]	47	50.11 (12.01)[b]	.76[e]
CBCL / BASC-2 Externalizing (t-scores)	31	43.74 (8.16)[a]	47	46.83 (6.71)[b]	.99[e]
PSI-SF	31	41.47 (15.44)[c]	39	78.38 (17.91)[d]	.02[e]

p: probability of differences between samples according to independent samples t-tests and chi-square tests.
  *Significant difference between samples.
[a]CBCL t-scores.
[b]BASC-2 t-scores.
[c]25-item PSI Short Form raw total score.
[d]36-item PSI-SF raw total score.
[e]Fisher's exact tests comparing samples on number of mothers above clinical cutoff.

and the mothers were invited for an individual assessment with their child at a university-based research center. During this assessment, mothers provided informed consent and completed the DPICS observation as part of a larger battery of measures. Families in this sample received $30 or $40 for their participation.

The total U.S. sample included 86 mother-child dyads. These families participated in two different studies: a study examining the relation between parenting-related cognitions and perceptions of children's behavior (n = 39; Shanley, 2008) and a study examining the effect of coaching on parent-child interactions (n = 47; Shanley & Niec, 2010). The DPICS codes of the 86 mother-child dyads have been previously compared to a sample of Mexican American families (McCabe et al., 2013). To be included in our sample in the current study, child behavior had to be below the clinical cutoff score on the ECBI Intensity Scale or the Externalizing Composite of the Behavior Assessment System for Children-Second Edition (BASC-2; Reynolds & Kamphaus, 2004; see Table 1). In this sample, all children and mothers had a Caucasian background, but one father had a Hispanic background. The ages of the children ranged from 2–7 years. Mother's ages ranged from 24–50 years old. Most families were two-parent families: 83.7% of the mothers were married and 2.3% were unmarried living with a partner. The average family income ranged from $30,001–$39,000, though income data were collected for a part of the sample only. Most mothers were highly educated (Table 1); 14% of the mothers had a high-school education or less.

Measures

Dyadic parent-child interaction coding system

The DPICS is a behavioral observational coding system that measures the quality of parent-child interactions. In this study, DPICS-III was used in both samples (Eyberg, Nelson, Duke, & Bogss, 2005). Parents and children are observed in three 5-minute situations that require an increasing degree of parental control. In the first situation, Child-Led Play (CLP), parents are instructed to follow their child's lead and to play along with the activity chosen by the child. In the second situation, the Parent-Led Play (PLP), parents are instructed to tell their child that it is the parent's turn to choose the activity and to play along with the parent according to their rules. In the last situation, parents are instructed to tell their child to Clean Up (CU) the toys without assistance. In both Dutch and U.S. samples, the same standard instructions were used and the observations were videotaped for later coding. In the Dutch sample the observation was recorded at the family's home and in the U.S. sample the observation was at the research center. In this study, the following parent behaviors used in the DPICS manual were included in the analyses: negative talk, direct command, indirect command, labeled praise, unlabeled praise, reflective statement, behavior description, question, neutral talk, positive touch, and negative touch. For these parental behaviors, the frequencies per situations were counted.

In both samples, independent master-level research assistants and undergraduate students conducted DPICS observations and coding. All coders were intensively trained to 80% agreement with an expert coder for all categories. For each mother-child dyad observation, one random situation (CLP, PLP, or CU) was coded again by a second coder to the estimate interrater reliability. In the Dutch sample, the average percent agreed was calculated. This percent agreed was calculated by summing the agreements across mother-child observations (and situations) and dividing by agreements plus disagreements across observations. For all double-coded observations, the overall percent agreement across DPICS categories was .91 (range .78–.98). Across situations the interrater reliability was similar: 0.91 (CLP), 0.93 (PLP), 0.90 (CU). In the U.S. sample, the average kappa also was .91 (range .84–.97).

Eyberg Child Behavior Inventory (ECBI)

All children in the Dutch sample and 39 children in the U.S. sample were screened for eligibility using the ECBI, a 36-item parent-report of child behavior problems (Eyberg & Pincus, 1999). Both English and Dutch versions have good established reliability (Abrahamse et al., 2015; Funderburk,

Eyberg, Rich, & Behar, 2003). The ECBI has two scales: the Intensity Scale, which measures the frequency of child behavior problems using a 7-point Likert scale (1 = *never* to 7 = *always*), and the Problem Scale, which measures parental tolerance for children's misbehavior, asking parents whether or not they view each of the described behaviors as problematic, using a dichotomous scale (1 = *yes*, 0 = *no*). The published cutoff score for the Intensity Scale is ≤132 and ≤15 for the Problem Scale. In the present study, the internal consistency (Cronbach's alpha) for the ECBI Intensity Scale was .90 for the Dutch sample and .89 for the U.S. sample. The internal consistencies (K-R 20) for the ECBI Problem Scale were .79 and .86, respectively.

Behavioral Assessment System for Children, Second Edition (BASC-2)

Forty-nine of the children from the U.S. sample were screened for eligibility for participation with the BASC-2, a parent report for child behavior and emotional problems using a 4-point scale ranging from 1 (*never*) to 4 (*almost always*) (Reynolds & Kamphaus, 2004). Good psychometric properties were found for this questionnaire. In the current study, the composite scales Internalizing Problems and Externalizing Problems were used. The internal consistency for the Externalizing Problems Composite score was .87. According to the professional manual, a T-score ≥70 indicated clinically significant behavior.

Child Behavior Checklist (CBCL)

In the Dutch sample, the CBCL was used as an additional questionnaire to measure the level of the child's internalizing and externalizing behavior problems (Achenbach & Rescorla, 2000, 2001). Two versions were used; the CBCL for ages 1.5–5 (100 items) and the CBCL 6–18 years (113 items) both using a 3-point scale (0 = *not true*, 1 = *somewhat or sometimes true*, 2 = *very true or often true*). The Dutch translations of the CBCL have good psychometric properties (Verhulst, van der Ende, & Koot, 1996) and the Cronbach's alphas in present study were .75 for the Internalizing Scale and .89 for the Externalizing Scale. To combine the two CBCL age versions, t-scores were calculated using the professional manual, with $t \geq 60$ indicating clinical behavior.

Parenting Stress Index–Short Form (PSI-SF)

In both samples, the PSI-SF was administered to assess parents' perceptions of stress in the parent-child relationship (Abidin, 1995). The English version contains 36 items using a 5-point scale ranging from 1 (*strongly agree*) to 5 (*strongly disagree*). The Dutch translation and adaptation of the PSI-SF

(De Brock, Vermulst, Gerris, & Abidin, 1992) contains 25 items which were rated on a 6-point scale ranging from 1 (*completely disagree*) to 6 (*completely agree*). Reliability and validity for both English and Dutch versions were described as satisfactory to good. In the present study, the sum of all items as an overall parenting stress scale was used with internal consistencies of .95 for the Dutch sample and .93 for the U.S. sample. According to published norms, a raw Total Stress Score above 74 (Dutch version) and 90 (English version), or greater than the 89th percentile, indicated clinically significant levels of parenting stress.

Data analysis

All analyses were performed in SPSS version 22. One-week test-retest reliability of the DPICS categories in the Dutch sample were calculated using Pearson correlations and paired samples *t*-tests. Second, independent samples *t*-tests and chi-square tests were conducted to compare the Dutch and U.S. samples on demographic characteristics and the ECBI means. Also, the number of children with clinical levels of problem behavior and parenting stress for both samples were reported and compared between samples using chi-square tests. Finally, multivariate analyses of covariance (MANCOVAs) were conducted to compare the Dutch and U.S. sample on each category on the DPICS situation (CLP, PLP, and CU).

Results

Test-retest reliability of the DPICS in a Dutch sample

Families in the Dutch sample were visited for a second time one week after the first assessment in order to evaluate the test-retest reliability of the Dutch version of the DPICS. Test-retest reliability was calculated for the individual parent categories per each situation and for the total sum of the categories over the three situations. Table 2 presents the Pearson correlations of the one-week test-retest reliability of the DPICS categories, which were significant for most categories, except for negative talk, behavior descriptions, positive touch, and negative touch. Additionally, paired *t*-tests using the total sum of the categories over the three situations revealed no significant differences between the mean frequencies for almost all parent categories (Table 2). The mean frequencies for each behavior were calculated by the average between the CLP, PLP, and CU scores. Mothers used significantly more questions, $t(30) = 2.26$, $p = .03$, and less positive touch, $t(30) = -2.37$, $p = .03$, during the first assessment.

Table 2. One-week test-retest reliability for Dutch sample ($n = 31$).

	CLP	PLP	CU	Total scales		
	r	r	r	M (SD) T₁	M (SD) T₂	r
Negative talk	.19*	.25	.06	1.58 (13.34)	1.19 (1.25)	.19
Direct command	−.03	.38*	.58**	16.35 (10.44)	14.58 (8.34)	.77***
Indirect command	−.07	.34	.65***	18.19 (9.09)	17.84 (11.34)	.69***
Labeled praise	−.07	.06	.11	0.57 (0.86)	0.60 (1.13)	.50**
Unlabeled praise	.31	.17	.56**	9.50 (6.79)	9.03 (7.10)	.50**
Reflective statement	.34	.20	.37*	6.81 (5.48)	6.48 (5.21)	.70***
Behavior description	−.11	−.14	.26	0.42 (0.62)	0.65 (1.05)	.08
Question	.56**	.48**	.55**	47.51 (17.17)	42.00 (19.38)*	.73***
Neutral talk	.52**	.39*	.62***	95.51 (24.41)	89.32 (27.88)	.64***
Positive touch	.16	.74***	.05	0.97 (1.43)	2.32 (3.23)*	.25
Negative touch	−.05	.21	.23	0.48 (1.00)	0.58 (0.89)	.24

Note. CLP: Child-Led Play; PLP: Parent-Led Play; CU: Clean Up; T₁ First assessment; T₂ Second assessment.
*$p < .05$;
** $p < .01$;
*** $p < .001$.

Demographic differences across U.S. and Dutch samples

Table 1 presents the percentages and means of demographic characteristics and questionnaires. Independent samples t-tests and chi-square tests revealed no significant differences on child age and gender. However, mothers in the Dutch sample were significantly older than the mothers in the U.S. sample, $t (112) = 5.69$, $p \leq .001$.

The means of the additional questionnaires, the ECBI, CBCL, BASC-2, and PSI-SF, are also reported in Table 1. Since the questionnaires differed between the samples, the numbers of children with a score within the clinical level were compared using Fisher's exact tests. For the internalizing scale (CBCL and BASC-2) in both samples, only one mother reported her child's behavior above the clinical cutoff, indicating no significant differences between samples. In the U.S. sample, no children had a score above the clinical cutoff on the BASC-2. In the Dutch sample only one child had a score within the clinical level of the CBCL externalizing scale, again indicating no significant differences between samples. A significant difference was found between the samples on the PSI-SF when the number of mothers with clinically significant levels of parenting stress were compared (scores > 89th percentile; $p = .02$, two-tailed Fisher's exact test). In the U.S. sample, 8.6% of the mothers had clinical levels of parenting stress compared to 3.1% of the mothers in the Dutch sample.

Differences in mother-child interactions across U.S. and Dutch samples

Because maternal age significantly differed between samples, this variable was included as a covariate in the multivariate analyses (MANCOVA). Although not all assumptions (outliers and normality) for this test were satisfied, we have still chosen to use this test, because it is expected that the test is robust enough

to deal with these violations. Table 3 shows the mean scores and standard deviations per DPICS parent category in each situation (CLP, PLP, and CU). For the Dutch sample, we used the first DPICS observation in the analysis. For all DPICS situations, the overall MANCOVA was significant; (CLP, Wilks' lambda (11, 110) = .702, $p = .001$; PLP, Wilks' lambda (11, 110) = .474, $p \leq .001$; CU, Wilks' lambda (11, 110) = .576, $p \leq .001$). Although significant differences between samples were found in the three situations, the mean frequencies for a number of categories were small. For example, for all mothers and all situations, the mean frequencies of labeled praises, behavior descriptions, and negative touch were smaller than one.

In regard to differences between samples on individual DPICS categories, in the Child-Led Play situation, U.S. mothers used significantly more frequent questions than Dutch mothers. In this situation, however, no significant differences were found between the samples in the frequency of negative talk, commands, labeled and unlabeled praises, reflective statements, behavior descriptions, neutral talk, and positive and negative touch.

In the Parent-Led Play situation, U.S. mothers used significantly more frequent negative talks, commands (direct and indirect), and behavior descriptions. Dutch mothers used reflective statements more frequently. No significant differences were found on praises, questions, neutral talk, and positive and negative touch.

Finally, during the Clean Up situation, few differences between samples were found. However, U.S. mothers more frequently used behavior descriptions and questions during the CleanUp situation, while Dutch mothers again used more negative touch. In this situation, there were no significant differences between samples on negative talk, commands, praises, reflective statements, neutral talk, and positive touch.

Discussion

The purpose of this study was to examine the psychometric properties of an observational assessment for mother-child interactions, the DPICS, within a community sample in The Netherlands and to compare the DPICS findings in this Dutch sample to a similar sample in the United States. High one-week test-retest reliability was found for most parent categories, including commands, praise, reflective statements, questions, and neutral talk. Thus, mothers' verbal interactions with their children were generally stable over a one-week period. Negative talk, behavioral descriptions, and nonverbal behavior (e.g., positive and negative touch) of the mother, however, were not significantly correlated between the two observations. An explanation may be that inappropriate maternal behavior such as negative talk and negative touch are influenced by the behavior of the

Table 3. Differences between Dutch and U.S. samples on the DPICS categories for each situation using MANCOVA.

Parent categories	CLP-Dutch (n = 30) M (SD)	CLP-US (n = 80) M (SD)	F	PLP-Dutch (n = 31) M (SD)	PLP-U.S. (n = 79) M (SD)	F	CU-Dutch (n = 31) M (SD)	CU-U.S. (n = 79) M (SD)	F
Negative talk	0.40 (0.86)	0.77 (1.23)	2.60	1.29 (1.70)	2.51 (3.13)	6.44*	1.58 (2.85)	4.10 (6.82)	5.27*
Direct command	2.60 (1.65)	2.56 (3.13)	0.41	5.87 (5.64)	8.72 (7.62)	3.97*	7.81 (6.65)	9.61 (7.47)	1.13
Indirect command	2.90 (2.45)	1.98 (1.87)	2.92	6.00 (3.81)	10.19 (7.00)	8.12**	9.39 (6.09)	11.11 (6.32)	0.26
Labeled praise	0.03 (0.18)	0.18 (0.41)	3.23	0.19 (0.75)	0.16 (0.44)	0.42	0.32 (0.54)	0.61 (1.20)	0.93
Unlabeled praise	2.27 (2.56)	2.45 (2.33)	0.25	2.35 (2.75)	2.49 (2.57)	0.001	4.68 (4.18)	4.19 (4.10)	1.09
Reflective statement	2.57 (2.91)	2.32 (2.60)	0.46	2.52 (2.41)	1.65 (2.05)	5.42*	1.81 (2.18)	1.63 (2.53)	0.50
Behavior description	0.17 (0.46)	0.30 (0.62)	2.29	0.10 (0.30)	0.30 (0.70)	4.63*	0.10 (0.30)	0.32 (0.71)	5.19*
Question	18.33 (7.74)	27.77 (13.14)	13.81**	17.45 (7.39)	24.43 (14.31)	6.22*	11.97 (7.07)	23.95 (11.24)	33.68***
Neutral talk	27.27 (9.03)	27.45 (11.78)	0.76	38.45 (12.13)	35.53 (11.23)	2.08	29.84 (13.23)	29.75 (10.97)	0.08
Positive touch	0.17 (0.46)	0.15 (0.55)	0.01	0.13 (0.43)	0.22 (0.61)	0.02	0.65 (1.28)	1.61 (2.65)	1.99
Negative touch	0.07 (.25)	0.03 (0.16)	1.54	0.10 (0.30)	0.03 (0.16)	3.45	0.32 (0.87)	0.08 (0.31)	5.25*

CLP: Child-Led Play; PLP: Parent-Led Play; CU: Clean Up; Dutch: Dutch sample; U.S.: United States sample.
*p <.05;
**p <.01;
***p <.001.

child, and therefore is more dependent on specific situations. Also, our findings were similar with previous findings on the one-week test-retest reliability found by Brinkmeyer (as cited in Eyberg et al., 2013), where inappropriate maternal behavior (critical statements) had lower correlations than positive maternal behavior (praise). Based on our findings, the mothers' positive verbalizations seemed more stable, except for behavior descriptions and positive touch. The lack of significant findings for these categories may be explained in part by range restriction and lack of variance, because these were categories that occurred very infrequently. With respect to other psychometric properties of the DPICS, the interrater reliability was found to be high among the Dutch coders, further supporting the utility of the coding system in The Netherlands.

With regards to the comparison of the Dutch mother-child dyads and the U.S. mother-child dyads, the overall comparison suggested differences between samples, but the DPICS scores on individual categories were largely similar between populations. Dutch and U.S. mothers showed similar frequencies of praises, neutral talk, and positive touch during the interaction with their child. Some behaviors, however, were significantly different between samples, particular during the Parent-Led Play where parents are instructed to tell the child to play according their rules. Dutch mothers gave fewer commands, used less negative talk (e.g., criticism, negative commands), and used more reflective statements, suggesting that Dutch mothers are less directive in their interaction with their children. Other significant differences found between behaviors of Dutch and U.S. mothers were that there were more behavior descriptions for U.S. mothers and more negative touch for Dutch mothers, but these behaviors were limited ($M < 1$) in all situations. Consistent with our hypothesis, one of the main differences across situations was that Dutch mothers used fewer questions. In PMT programs, negative parental leading (including questions, commands and negative talk) are discouraged for strengthening the parent-child relationship in order to address conduct problems in children. The limited use of questions, indirect commands and negative talk by the nonclinical Dutch mothers may indicate that the approach of PMT programs is a good fit with Dutch families and also reflects the sensitive and authoritative parenting style found in previous studies among Dutch parents. Previous research also demonstrated less directive behavior and more autonomy-oriented parenting behavior in Dutch mothers, including high levels of authoritative control (e.g., praises, understanding behavior (Gerrits, Dekovic, Groenendaal, & Noom, 1996; Yaman, Mesman, Van IJzendoorn, Bakermans-Kranenburg, & Linting, 2010). Also, a large body of research has shown that cultures may vary in their tolerance and perception of disruptive behavior in children (Timimi & Taylor, 2004; Weeland,

van Aar, & Overbeek, 2017). Therefore, parenting behavior in response to their child's behavior may differ between cultures as well. In summary, the findings show that the parent-child interactions of Dutch mother-child dyads are somewhat similar to the interactions of U.S. mother-child dyads, although they differ in specific ways. An important next step, however, is to study the utility of the DPICS in assessing parent-child interactions within a clinical population in The Netherlands, in particular, examining treatment sensitivity (i.e., the measure's ability to detect pre- to post-treatment changes) because the DPICS is primarily used to assess treatment outcomes for young children and their parent participating in PMT programs.

Our study addresses an important gap in the literature for the use of the DPICS within a Dutch population. By comparing this standardized behavior observation measure in Dutch and U.S. mother-child dyads, knowledge of the usefulness in clinical practice and treatment outcome studies is expanded. Also, because well-validated assessment techniques utilizing observation of parent-child interactions play an important role in the implementation of PMT programs, our study is relevant to the global dissemination of PMT.

While our findings support the validity of the Dutch version of the DPICS, there are limitations to this study that should be considered. First, the sample was smaller than other DPICS studies on psychometric properties (Eyberg et al., 2013) and primarily included highly educated mothers. Also, our study lacks a power analysis and sample-size estimation. These limitations may increase the chance for Type I and II errors, leading to limited generalizability of our findings. New research using DPICS in The Netherlands should address this limitation and study the psychometric properties in less highly educated parents, since these families are usually present in clinical populations. It is also a limitation that the parent-child interactions were coded in the home (Dutch sample) and in a research center (U.S. sample), which could affect child behavior in a familiar vs. unfamiliar environment and necessitate the need for few directives in the Dutch sample and more directives in the U.S. sample. A study using a very small sample with four children and their parents showed evidence suggesting that parent-child interactions are similar in the clinic and the home (Shriver, Frerichs, Williams, & Lancaster, 2013). However, it should be noted that so far there is also a lack of consensus in literature related to observation in different settings. A final limitation worth mentioning is that almost 20% of the participating children had one or two parents born outside The Netherlands. Although parents were asked about the language they use to interact with their child during the initial telephone screening, this language may not be the parent's first language. Therefore, limited

vocabulary or fluency could impact interactions between these mothers and children, requiring additional research on this issue.

Conclusion

The findings of our study provide evidence that the DPICS is a psychometrically sound observational measure to assess mother-child interactions in The Netherlands. The one-week test-retest reliability, the normative scores, and the similarities between the DPICS scores of Dutch mother-child dyads and mother-child dyads in a U.S. sample support the usefulness of this behavioral assessment technique in The Netherlands. Although further evaluation of the psychometric properties of the DPICS in different Dutch samples is recommended, the current results are promising for the use of the DPICS with mother-child dyads in The Netherlands.

Acknowledgements

Funding

References

Abidin, R. R. (1995). *Parenting Stress Index: Professional manual* (3rd ed.). Odessa, FL: Psychological Assessment Resources.

Abrahamse, M. E., Junger, M., Chavannes, E. L., Coelman, F. J. G., Boer, F., & Lindauer, R. J. L. (2012). Parent-Child Interaction Therapy for preschool children with disruptive behaviour problems in the Netherlands. *Child and Adolescent Psychiatry and Mental Health*, 6(1), 24. doi:10.1186/1753-2000-6-24

Abrahamse, M. E., Junger, M., Leijten, P. H., Lindeboom, R., Boer, F., & Lindauer, R. J. (2015). Psychometric properties of the Dutch Eyberg Child Behavior Inventory (ECBI) in a community sample and a multi-ethnic clinical sample. *Journal of Psychopathology and Behavioral Assessment*, 37(4), 679–691. doi:10.1007/s10862-015-9482-1

Abrahamse, M. E., Junger, M., Van Wouwe, M. A., Boer, F., & Lindauer, R. J. (2016). Treating child disruptive behavior in high-risk families: A comparative effectiveness trial from a community-based implementation. *Journal of Child and Family Studies*, 25(5), 1605–1622. doi:10.1007/s10826-015-0333-4

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms and profiles*. Burlington, VT: University of Vermont.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school age forms and profiles*. Burlington, VT: University of Vermont.

Aspland, H., & Gardner, F. (2003). Observational measures of parent-child interaction: An introductory review. *Child and Adolescent Mental Health*, 8(3), 136–143. doi:10.1111/1475-3588.00061

Barbosa, J., Tannock, R., & Manassis, K. (2002). Measuring anxiety: Parent-child reporting differences in clinical samples. *Depression and Anxiety*, 15(2), 61–65. doi:10.1002/da.10022

Bjørseth, Å., McNeil, C. B., & Wichstrøm, L. (2015). Screening for behavioral disorders with the Dyadic Parent-Child Interaction Coding System: Sensitivity, specificity, and core discriminative components. *Child and Family Behavior Therapy*, 37(1), 20–37. doi:10.1080/07317107.2015.1000228

Bjørseth, Å., & Wichstrøm, L. (2016). Effectiveness of Parent-Child Interaction Therapy (PCIT) in the treatment of young children's behavior problems. A randomized controlled study. *PLoS One*, 11(9), e0159845. doi:10.1371/journal.pone.0159845

Borden, L. A., Herman, K. C., Stormont, M., Goel, N., Darney, D., Reinke, W. M., & Webster-Stratton, C. (2014). Latent profile analysis of observed parenting behaviors in a clinic sample. *Journal of Abnormal Child Psychology*, 42(5), 731–742. doi:10.1007/s10802-013-9815-z

Briegel, W., Walter, T., Schimek, M., Knapp, D., & Bussing, R. (2015). Parent-Child Interaction Therapy im in-room-coaching: Ergebnisse einer ersten Deutschen fallstudie. *Kindheit Und Entwicklung*, 24(1), 47–54. doi:10.1026/0942-5403/a000158

Cheung, F. M., Kwong, J. Y. Y., & Zhang, J. (2003). Clinical validation of the Chinese Personality Assessment Inventory. *Psychological Assessment*, 15(1), 89–100. doi:10.1037/1040-3590.15.1.89

De Brock, A. J. L. L., Vermulst, A. A., Gerris, J. R. M., & Abidin, R. R. (1992). *Nijmeegse Ouderlijke Stress Index*. Lisse, The Netherlands: Swets & Zeitlinger.

Eyberg, S. M., & Funderburk, B. (2011). *Parent-Child Interaction Therapy protocol*. Gainesville, FL: PCIT International.

Eyberg, S. M., Nelson, M. M., Duke, M., & Bogss, S. R. (2005). *Manual for the Dyadic Parent-Child interaction coding system*. Gainesville, FL: University of Florida.

Eyberg, S. M., Nelson, M. M., Ginn, N. C., Bhuiyan, N. N., & Boggs, S. R. (2013). *Manual for the Dyadic Parent-child interaction coding system*. Gainesville, FL: PCIT International.

Eyberg, S. M., & Pincus, D. B. (1999). *Eyberg Child Behavior Inventory and Sutter-Eyberg Behavior Inventory-Revised: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Funderburk, B., Eyberg, S. M., Rich, B. A., & Behar, L. (2003). Further psychometric evaluation of the Eyberg and Behar rating scales for parents and teachers of preschoolers. *Early Education and Development*, 14(1), 67–81. doi:10.1207/s15566935eed1401_5

Garcia-Barrera, M. A., Karr, J. E., Duran, V., Direnfeld, E., & Pineda, D. A. (2015). Cross-cultural validation of a behavioral screener for executive Functions: Guidelines for clinical use among Colombian children with and without ADHD. *Psychological Assessment*, 27(4), 1349. doi:10.1037/pas0000117

Gardner, F. (2000). Methodological issues in the direct observation of parent-child interaction: Do observational findings reflect the natural behavior of participants? *Clinical Child and Family Psychology Review*, 3(3), 185–198. doi:10.1023/A:1009503409699

Gerrits, L., Dekovic, M., Groenendaal, J., & Noom, M. (1996). Opvoedingsgedrag. In J. Rispens, J. M. Hermanns, & W. H. Meeus (Eds.), *Opvoeden in Nederland* (pp. 41–69). Assen, The Netherlands: Van Gorcum.

Hops, H., Davis, B., & Longoria, N. (1995). Methodological issues in direct observation: Illustrations with the Living in Familial Environments (LIFE) coding system. *Journal of Clinical Child Psychology*, *24*(2), 193–203. doi:10.1207/s15374424jccp2402_7

Hupp, S. D. A., Reitman, D., Forde, D. A., Shriver, M. D., & Kelley, M. L. (2008). Advancing the assessment of parent-child interactions: Development of the parent instruction–giving game with youngsters. *Behavior Therapy*, *39*(1), 91–106. doi:10.1016/j.beth.2007.05.004

Kaplan, R. M. (1985). The controversy related to the use of psychological tests. In B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurements, and applications* (pp. 465–504). New York, NY: Wiley.

Kazdin, A. E., Siegel, T. C., & Bass, D. (1992). Cognitive problem-solving skills training and parent management training in the treatment of antisocial behavior in children. *Journal of Consulting and Clinical Psychology*, *60*(5), 733–747. doi:10.1037/0022-006X.60.5.733

Leung, C., Tsang, S., Sin, T. C. S., & Choi, S. (2015). The efficacy of Parent-Child Interaction Therapy with Chinese families: Randomized controlled trial. *Research on Social Work Practice*, *25*(1), 117–128. doi:10.1177/1049731513519827

Matos, M., Bauermeister, J. J., & Bernal, G. (2009). Parent-Child Interaction Therapy for Puerto Rican preschool children with ADHD and behavior problems: A pilot efficacy study. *Family Process*, *48*(2), 232–252. doi:10.1111/j.1545-5300.2009.01279.x

McCabe, K. M., Shanley, J. R., Niec, L. N., Naaf, M., Yeh, M., & Lau, A. S. (2013). Cultural differences in the parenting of young children: An observational study of low-income Mexican American and European American families. *Child and Family Behavior Therapy*, *35*(4), 307–326. doi:10.1080/07317107.2013.846680

McCabe, K. M., Yeh, M., Lau, A., & Argote, C. B. (2012). Parent-child interaction therapy for mexican americans: Results of a pilot randomized clinical trial at follow-up. *Behavior Therapy*, *43*(3), 606–618. doi:10.1016/j.beth.2011.11.001

Mesman, J., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2012). Unequal in opportunity, equal in process: Parental sensitivity promotes positive child development in ethnic minority families. *Child Development Perspectives*, *6*(3), 239–250. doi:10.1111/j.1750-8606.2011.00223.x

Mieloo, C. L., Bevaart, F., Donker, M. C., Van Oort, F. V., Raat, H., & Jansen, W. (2014). Validation of the SDQ in a multi-ethnic population of young children. *European Journal of Public Health*, *24*(1), 26–32. doi:10.1093/eurpub/ckt100

Niec, L. N. (Ed.). (2018). *Handbook of Parent-Child Interaction Therapy: Innovations and applications for research and practice*. New York, NY: Springer Publications.

Niec, L. N., Abrahamse, M. E., Egan, R., Coelman, F., & Heiner, W. (2018). Global dissemination of Parent-Child Interaction Therapy: The perspectives of Dutch trainees. *Children and Youth Services Review*, *94*, 485. doi:10.1016/j.childyouth.2018.08.019

Niec, L. N., Shanley, J. R., Barnett, M. L., Baker, S. E., & Solomon, D. T. (2015). Optimal or typical performance? The Impact of instructional set on the behavioral assessment of parent-child interactions. *Child and Family Behavior Therapy*, *37*(2), 105–113. doi:10.1080/07317107.2015.1035981

Nixon, R. D. V., Sweeney, L., Erickson, D. B., & Touyz, S. W. (2004). Parent-Child Interaction Therapy: One- and two-year follow-up of standard and abbreviated treatments for oppositional preschoolers. *Journal of Abnormal Child Psychology*, *32*(3), 263–271. doi:10.1023/B:JACP.0000026140.60558.05

Patterson, G. R., & Forgatch, M. S. (1995). Predicting future clinical adjustment from treatment outcome and process variables. *Psychological Assessment*, 7(3), 275–285. doi:10.1037/1040-3590.7.3.275

Pearl, E. S. (2009). Parent management training for reducing oppositional and aggressive behavior in preschoolers. *Aggression and Violent Behavior*, 14(5), 295–305. doi:10.1016/j.avb.2009.03.007

Prasadarao, P., & Kumaraiah, V. (1997). Behavioral assessment: Recent trends and implications for interventions in behavioral medicine. *Nimhans Journal*, 15(4), 273–284.

Reitman, D., Hummel, R., Franz, D. Z., & Gross, A. M. (1998). A review of methods and instruments for assessing externalizing disorders: Theoretical and practical considerations in rendering a diagnosis. *Clinical Psychology Review*, 18(5), 555–584. doi:10.1016/S0272-7358(98)00003-8

Rescorla, L. A., Achenbach, T. M., Ivanova, M. Y., Harder, V. S., Otten, L., Bilenberg, N., … Verhulst, F. C. (2011). International comparisons of behavioral and emotional problems in preschool children: parents' reports from 24 societies. *Journal of Clinical Child and Adolescent Psychology*, 40(3), 456–467. doi:10.1080/15374416.2011.563472

Reynolds, C. R., & Kamphaus, R. W. (2004). *BASC-2: Behavior assessment system for children* (2nd ed.). Circle Pines, MN: Erlbaum.

Roberts, M. W., & Hope, D. A. (2001). Clinic observations of structured parent-child interaction designed to evaluate externalizing disorders. *Psychological Assessment*, 13(1), 46–58. doi:10.1037/1040-3590.13.1.46

Schimek, M., Walter, T., Bussing, R., & Briegel, W. (2014). Parent-Child Interaction Therapy und multiple familiäre Belastungen: Ein fallbericht. *Verhaltenstherapie*, 24(1), 40–46. doi:10.1159/000360198

Shanley, J. R. (2008). *The impact of feedback on parent's reports of self-efficacy and observed skills acquisition* (Vol. Unpublished doctoral dissertation). Mt. Pleasant, MI: Central Michigan University.

Shanley, J. R., & Niec, L. N. (2010). Coaching parents to change: The impact of in vivo feedback on parents' acquisition of skills. *Journal of Clinical Child and Adolescent Psychology*, 39(2), 282–287. doi:10.1080/15374410903532627

Shriver, M. D., Frerichs, L. J., Williams, M., & Lancaster, B. M. (2013). Comparing parent-child interactions in the clinic and at home: An exploration of the validity of clinical behavior observations using sequential analysis. *Child and Family Behavior Therapy*, 35(1), 1–24. doi:10.1080/07317107.2013.760418

Thornberry, T., Jr., & Brestan-Knight, E. (2011). Analyzing the utility of Dyadic Parent-Child Interaction Coding System (DPICS) warm-up segments. *Journal of Psychopathology and Behavioral Assessment*, 33(2), 187–195. doi:10.1007/s10862-011-9229-6

Timimi, S., & Taylor, E. (2004). ADHD is best understood as a cultural construct. *British Journal of Psychiatry*, 184(1), 8–9. doi:10.1192/bjp.184.1.8

Van der Bruggen, C. O., Stams, G. J. M., Bögels, S. M., & Paulussen-Hoogeboom, M. C. (2010). Parenting behaviour as a mediator between young children's negative emotionality and their anxiety/depression. *Infant and Child Development*, 19(4), 354–365. doi:10.1002/icd.665

Verhulst, F. C., van der Ende, J., & Koot, H. M. (1996). *Handleiding voor de CBCL/4-18*. Rotterdam, The Netherlands: Erasmus University, Department of Child and Adolescent Psychiatry, Sophia Children's Hospital.

Weeland, J., van Aar, J., & Overbeek, G. (2017). Dutch norms for the eyberg child behavior inventory: Comparisons with other Western Countries. *Journal of Psychopathology and Behavioral Assessment*, *40*(2), 224–234. doi:10.1007/s10862-017-9639-1

Yaman, A., Mesman, J., Van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., & Linting, M. (2010). Parenting in an individualistic culture with a collectivistic cultural background: The case of Turkish immigrant families with toddlers in the Netherlands. *Journal of Child and Family Studies*, *19*(5), 617–628. doi:10.1007/s10826-009-9346-y

Zaslow, M. J., Weinfield, N. S., Gallagher, M., Hair, E. C., Ogawa, J. R., Egeland, B., … De Temple, J. M. (2006). Longitudinal prediction of child outcomes from differing measures of parenting in a low-income sample. *Developmental Psychology*, *42*(1), 27–37. doi:10.1037/0012-1649.42.1.27