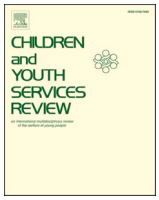
Children and Youth Services Review 143 (2022) 106679

Contents lists available at ScienceDirect

Children and Youth Services Review

journal homepage: www.elsevier.com/locate/childyouth



The dyadic parent-child interaction coding system (DPICS): Negative talk as an indicator of dysfunctional mother-child interaction

María Cañas [a,*], Izaskun Ibabe [b], Ignacia Arruabarrena [a], Joaquín De Paúl [a]

[a] Department of Social Psychology, University College of Psychology. University of the Basque Country, Avda. Tolosa 70, 20018-Donostia-San Sebastián, Spain
[b] Department of Clinical and Health Psychology and Research Methodology, University College of Psychology. University of the Basque Country, Avda. Tolosa 70, 20018-Donostia-San Sebastián, Spain

ARTICLE INFO

Keywords:
Parent–child interaction
Observation
Child maltreatment
Discriminative validity
Child behavior problems

ABSTRACT

Background: The Dyadic Parent-Child Interaction Coding System (DPICS-IV) is a widely used observational instrument that assesses Parent-child interaction (PCI) quality. However, studies specifically examining its psychometric properties published in peer-reviewed journals are scarce. The present study aimed to provide evidence on the discriminative validity of the DPICS-IV to identify indicators of parent–child interaction among clinical mother–child dyads compared to non-clinical.

Method: Participants were 177 mother–child dyads with children aged 4 to 8 years: (1) a clinical sample of 80 dyads where mothers experienced significant difficulties managing their children's behavior problems and identified by Child Welfare and Child Protection Services as at risk for child maltreatment or with substantiated reports, and (2) a non-clinical sample of 97 dyads from the general population.

Results: DPICS Negative talk factor showed high discriminant capacity (AUC = 0.90) between samples, with a cut-off score of 8 that allowed mother–child dyads to be classified with a sensitivity of 82 % and a specificity of 89 %.

Conclusions: Findings of the present study suggested that the DPICS-IV Negative Talk factor is a robust indicator of dysfunctional PCI patterns of families involved with the Child Protection Services. Further research is needed to confirm these findings and to test the accuracy of the cut-off score with a representative sample of the general population.

1. Introduction

The quality of parent–child interaction (PCI) is a cornerstone of healthy child development (Nilsen et al., 2020; Werner et al., 2016). PCI is a dynamic bidirectional process shaped through reciprocal parent-related and child-related effects (Sameroff & Fiese, 2000). Persistent patterns of PCI have been considered objective and observable indicators of the parent–child relationship quality (Larrieu et al., 2018). Functional PCI patterns, characterized by the presence of warm, sensitive, and supportive parenting (Allen et al., 2014; Bocknek et al., 2009), have been robustly linked with optimal outcomes on academic and social competence throughout childhood, adolescence, and adulthood (Fraley et al., 2013; Jeon et al., 2013; Raby et al., 2015). By contrast, dysfunctional PCI patterns, marked by emotionally withdrawn, harsh or over-reactivity discipline (Easterbrooks et al., 2012; Eddy et al., 2001) heightens the risk of child psychopathology, poor academic outcomes, inadequate social competencies, and health problems across the lifespan (Akcinar & Shaw, 2018; Pinquart, 2017; Romanowicz et al., 2019; Stewart-Brown et al., 2005).

Several risk factors for dysfunctional PCI have been extensively documented, showing robust evidence across studies and cultures. Most consistent risk factors include low-socioeconomic status (Belsky et al., 2007; Holstein et al., 2021), caregiver mental health problems (Dietz et al., 2009; Lovejoy et al., 2000; Muzik et al., 2017; Ramsauer & Achtergarde, 2018) and substance abuse disorders (Kelley et al., 2015; Slesnick et al., 2014), and children's difficult temperament or externalizing behavior problems (Eddy et al., 2001; Goldstein et al., 2007).

The prevalence of dysfunctional PCI patterns within the general population ranges from 7.8 % to 10.5 % (Falceto et al., 2012; Holstein et al., 2021; Skovgaard et al., 2007) and can be detected in early stages

* Corresponding author.
E-mail addresses: maria.canas@ehu.eus (M. Cañas), izaskun.ibabe@ehu.eus (I. Ibabe), ignacia.arruabarrena@ehu.eus (I. Arruabarrena), joaquin.depaul@icloud.com (J. De Paúl).

https://doi.org/10.1016/j.childyouth.2022.106679
Received 14 March 2021; Received in revised form 11 May 2022; Accepted 29 September 2022
Available online 6 October 2022

of child development (Skovgaard et al., 2008). In the absence of effective treatment, the deleterious effects of dysfunctional PCI during infancy can extend beyond childhood and have a long-lasting impact throughout adolescence and adulthood (Fraley et al., 2013; Raby et al., 2015). Among clinical populations, such as maltreating families or children with externalizing behavior problems, PCI dynamics are characterized by dysfunctional patterns (Quetsch et al., 2018; Stith et al., 2009). Accumulated evidence across numerous *meta*-analyses indicates that parenting programs whose primary aim is enhancing PCI are effective approaches for the prevention and treatment of child maltreatment (Chen & Chan, 2016; McCoy et al., 2020; Shah et al., 2016; Vlahovicova et al., 2017), as well as child behavior problems (Gardner et al., 2019; Schulz et al., 2019; Werner et al., 2016). Therefore, early detection and effective intervention on dysfunctional PCI are essential to reduce harmful consequences and heighten positive outcomes for children and families.

1.1. Assessment of the Parent-Child interaction

Direct observation of behavior is considered the most effective method for assessing PCI (Lindahl et al., 2019). Guidelines for best practices in the context of child maltreatment (Budd et al., 2011; Office on Child Abuse and Neglect, 2018), as well as in clinical child psychology (Dowdy et al., 2013; Frick et al., 2010), recommend PCI observation as part of the comprehensive assessment of children and families. Observational methods capture unique and detailed information about ongoing PCI patterns that are interpreted as indicators of the parent–child relationship characteristics and quality (Lindahl et al., 2019). Self-reports add valuable information about caregivers' internal psychological processes such as attitudes, attributions, and unrealistic expectations (Gardner, 2000). Nevertheless, when the construct of interest is ongoing behavior, caregiver-reported data are prone to bias due to lack of awareness or to social desirability (Morsbach & Prinz, 2006; Swenson et al., 2016). Therefore, observation is the recommended approach to gather accurate information on behavior during interactions (Baumeister et al., 2007).

The selection of an appropriate observational instrument should be based on the child's age, the purpose of the assessment, and the proven psychometric properties related to the population of interest (Frick et al., 2010). The behaviors of interest will gradually evolve according to the child's stage of development (Budd et al., 2011). Instruments tailored to the assessment of parent-infant interactions are focused on measuring constructs such as sensitivity, responsiveness, or emotional availability through direct observation of, for instance, positive vocal or facial affect expressions (Lotzin et al., 2015). Observation tools targeting pre-school and school-aged children, in addition to warm parenting, shift their attention to behaviors related to discipline patterns, such as coercion or cooperation (Wilson & Durbin, 2012). According to a systematic review by Cañas et al. (2020), the vast majority of observational instruments have been developed to assess mother-infant dyads, hence there is a need for psychometrically sound instruments targeting school-aged children. In this review, the Dyadic Parent-Child Interaction Coding System (DPICS, Eyberg et al., 2013, 2014) was selected as a promising observational instrument.

1.2. Dyadic Parent-Child interaction coding system

The DPICS is a widely used observational instrument for the assessment of pre-school and school-aged children. It is a micro-coding system that assesses parent–child interaction quality by targeting behaviors associated with effective and ineffective parenting practices (Nelson & Olsen, 2018). DPICS primarily focuses on the frequency of parental verbalizations, such as praise, commands, or negative talk (Eyberg et al., 2014). It also targets children's responses to commands -compliance or non-compliance- as an index of children's optimal or disruptive behavior (Eyberg et al., 2014). The procedure involves three semi-structured situations, starting with child-directed play (5 min, coded), after which the parent switches to directing the play (5 min, coded), and concludes with the parent informing the child that it is time to pick up the toys (5 min, coded) (Eyberg et al., 2013). DPICS is in its fourth edition and has been used with different clinical populations (e.g., children with behavior problems or autism spectrum disorders, maltreated and high-risk children) across multiple settings (e.g., clinic, laboratory, home, school) with children aged 3–12 years (Eyberg et al., 2013). It is mainly used to assess families' clinical needs, monitor their progress during treatment, and as an outcome measure for preventive or rehabilitative programs.

Despite the widespread use of DPICS and the continued research conducted for over 30 years to refine and upgrade the instrument to its fourth edition (for a review, see Eyberg et al., 2013), the number of published studies on its psychometric properties is limited (Nelson & Olsen, 2018). Such paucity is a pervasive shortcoming among observational instruments pointed out by several studies in the field of behavioral observation (Aspland & Gardner, 2003; Roberts, 2001). Two recent systematic reviews found only four studies supporting DPICS psychometrics (Cañas et al., 2020; Hurley et al., 2014). The gathered evidence across such studies shows good inter-rater reliability, low-quality evidence of convergent, predictive, and discriminative validity, and lack of evidence of factorial validity. Nevertheless, recent research on mother–child dyads at risk for child maltreatment in Spain has established a DPICS four-factor structure comprising Praise, Questions, Clear commands, and Negative talk (Cañas at al., 2021). In the study, two of the DPICS-IV categories, Behavior Description and Reflections, were not retained in the final model. Both categories showed extremely low or zero variability, resulting in convergence problems or non-significant parameters. To achieve an adequate model fit, these categories were discarded. The yielded four-factor solution showed good fit indicators: ML $\chi 2$ (83, N = 80) = 84.92, p =.42; CFI = 0.959, RMSEA = 0.036, [90 % CI 0.00 - 0.07] (for more details of the process see Cañas at al., 2021). Researchers and clinicians who have used the DPICS within the PCIT assessment protocol will notice the exclusion of such categories. The main problem in obtaining a factor structure that encompasses these categories is that they are rarely observed during the pretreatment assessment phase.

1.3. Discriminative or known-groups validity

Practitioners require accurate instruments to identify families in need of intervention to improve parent–child interaction (Frick et al., 2010). Discriminative or known-groups validity is related to the capacity of any measure to discern between clinical and non-clinical cases (Mokkink et al., 2018). Some studies carried out in the United States have provided information about the DPICS discriminative validity (Aragona & Eyberg, 1981; Borrego et al., 2004; McCabe et al., 2010; Quetsch et al., 2018; Robinson & Eyberg, 1981; Timmer et al., 2002; Webster-Stratton, 1985a, 1985b). On the one hand, it has been found that, compared with non-abusive mothers, physically abusive mothers were more likely to display negative behaviors toward their children (Borrego et al., 2004; Webster-Stratton, 1985b) and to praise less often (Timmer et al., 2002). On the other hand, it has also been found that caregivers of referred children with externalizing behavior problems produced significantly more negative talk than caregivers of non-referred children (McCabe et al., 2010; Quetsch et al., 2018; Webster-Stratton, 1985a). Moreover, referred children exhibited a significantly higher rate of non-compliance than non-referred children (McCabe et al., 2010). Results of these studies were mainly based on mean differences between groups. However, assessment instruments intended for clinical purposes also need to prove discriminative accuracy by providing evidence of their sensitivity, that is, their ability to detect clinical cases (true positives) correctly, and specificity, their ability to detect non-clinical cases (true negatives) correctly (Frick et al., 2010). The preferred statistical approaches for examining an instrument's

discriminative accuracy are the precision indices derived from ROC curve analysis. The area under the ROC curve (AUC) determines the measured variable's strength as predictor of the clinical condition.

In the European context, Bjørseth et al. (2015) explored the discriminative accuracy of DPICS by ROC curve analysis providing AUC, sensitivity, and specificity indicators on a clinical sample of children with behavior problems. Their findings showed that Negative talk, Direct command with child compliance, and Indirect command with no opportunity to comply were the best discriminant categories with an excellent AUC value (0.85), with a good sensitivity index (85 %) but a specificity index (70 %) below the optimal threshold.

1.4. Purpose of the present study

The main objective of the present study was to determine the discriminative or known-groups validity of the DPICS-IV for parent–child interaction patterns in two samples. The current study included a clinical sample of mother–child dyads with substantiated maltreatment reports and identified as at risk for maltreatment by Child Protection Services in which mothers were experiencing significant difficulties managing children's behavior problems, and a non-clinical sample of community mother–child dyads.

The added value of the current study over previous research addressing similar questions is its analytic strategy. The DPICS-IV known-groups validity was examined using a three-stage analytical approach combining structural equation models, logistic regression, and ROC curve analysis. This thorough approach had not been previously applied to analyze the discriminative ability of DPICS and would strengthen the yielded outcomes.

Based on the findings of previous studies, it was expected that DPICS-IV would discriminate effectively between mother–child dyads from clinical and non-clinical samples in the current study. Specifically, it was hypothesized that mothers in the clinical sample would display more Negative talk toward their children than mothers in the non-clinical sample (Bjørseth et al., 2015; Borrego et al., 2004; Timmer et al., 2002; Webster-Stratton, 1985a, 1985b). Moreover, it was expected that children in the clinical sample would exhibit a higher rate of Noncompliance than children in the non-clinical sample (McCabe et al., 2010). The results are expected to hold when comparing the non-clinical sample to the two clinical subgroups separately: mother–child dyads with a substantiated report of maltreatment versus mother–child dyads at risk for child maltreatment.

To our knowledge, no study has so far examined optimal clinical cut-off scores for any DPICS category among maltreating or at risk for maltreatment mother–child dyads. Therefore, we also aimed to establish cut-off scores that significantly discriminate between mother–child dyads from the clinical and non-clinical samples, taking into account the two clinical subgroups. Based on previous studies in which clinical mother–child dyads have shown significantly higher rates of Negative talk than non-clinical mother–child dyads (McCabe et al., 2010; Quetsch et al., 2018; Webster-Stratton, 1985a, 1985b), we expected to find a cut-off score for the Negative talk factor that remained above the 80 % sensitivity and specificity standards (Cicchetti et al., 1995).

2. Method

2.1. Participants and procedure

A total of 177 mother–child dyads with children aged between 4 and 8 years old participated in the study: 80 dyads where mothers experienced significant difficulties managing their children's behavior problems and identified by Child Welfare and Child Protection Services of the region of Gipuzkoa (Basque Country, Spain) as substantiated maltreatment (n = 47, 58.75 %) or at risk for child maltreatment (n = 33, 41.25 %) (clinical sample), and 97 dyads from the general population (non-clinical sample).

Socio-demographic characteristics of participants are shown in Table 1. No statistically significant differences were observed for children's gender and age between samples, but significant differences were found for other sociodemographic variables. Mothers from the non-clinical sample were more frequently born in Spain, reported a higher level of education and better work situation than mothers from the clinical sample. Moreover, families from the non-clinical sample showed fewer economic difficulties and children lived more frequently with two biological parents than families from the clinical sample.

The Ethics Committee of the University of the Basque Country UPV/EHU (Spain) approved the study protocol. In order to be accredited in the use of DPICS it is necessary to undergo training to ensure that the observers code with 80 % agreement. In this study, two coders received intensive training and met the accreditation criteria for the use of DPICS.

2.1.1. Clinical sample

The clinical sample was selected from a larger Randomized Control Trial (RCT) carried out in the region of Gipuzkoa (Spain) examining the effectiveness of Incredible Years, an evidence-based early intervention program for children with significant behavior problems and their families (Arruabarrena et al., 2022). Child Welfare and Child Protection Services caseworkers informed eligible families based on the following

Table 1
Socio-demographic Characteristics of Participants.

	Clinical sample (n = 80)	Non-clinical sample (n = 97)		
Children	*n* (%)	*n* (%)	$\chi 2/t$	Size effect
Child gender			$\chi 2 = 1.95$	Φ =0.10
Male	52 (65.0)	53 (54.6)		
Female	28 (35.0)	44 (45.4)		
Child age [*M* (*SD*)]	5.93 (1.24)	6.30 (1.23)	$t = 1.84$	d = 0.20
Mothers	*n* (%)	*n* (%)	$\chi 2/t$	Size effect
Mother age [*M* (*SD*)]	37.25 (7.23)	40.11 (4.08)	$t = 3.31$**	d = 0.49
Mother origin			$\chi 2$ = 34.22**	V = 0.44
Spanish	50 (62.5)	94 (96.9)		
Immigrant	30 (37.5)	3 (3.1)		
Mother educational level			$\chi 2$ = 79.89**	V = 0.67
Elementary education	26 (32.5)	4 (4.1)		
Professional training/ secondary education	40 (50.0)	11 (11.3)		
University	14 (17.5)	82 (84.5)		
Mother work situation			$\chi 2$ = 26.71**	V = 0.39
Stable employment	36 (45.0)	73 (75.3)		
Temporary employment	10 (12.5)	15 (15.5)		
Unemployed/working at home	34 (42.5)	9 (9.3)		
Families	*n* (%)	*n* (%)	$\chi 2/t$	Size effect
Family income			$\chi 2$ = 32.54**	V = 0.43
Financial difficulties	29 (36.3)	3 (3.1)		
Without financial difficulties	51 (63.7)	94 (96.9)		
Family structure			$\chi 2$ = 49.28**	V = 0.53
Two biological parents	34 (43.5)	88 (90.7)		
One-parent family	4 (5.0)	3 (3.1)		
Divorced	42 (52.5)	6 (6.2)		

Note. M = mean; SD = standard deviation; *df* = degrees of freedom; $\chi 2$ = chi-square; Φ = phi;
V = Cramer's V; d = Cohen's d; **p <.001.

inclusion criteria: (1) there was a substantiated report of child maltreatment or a significant risk of child maltreatment, (2) children exhibited significant behavior problems, and (3) parents struggled significantly with managing their children's behavior. Parents with severe mental health disorders, severe cognitive limitations, or substance abuse were excluded. Children diagnosed with neurodevelopmental disorders (e.g., autism), who had significant developmental delays, or who were undergoing psychotherapeutic or psychiatric intervention were excluded from the study. Sexual abuse cases and children in temporary care were also excluded.

The total RCT participants were 111 mother–child dyads. Observational data were collected by the main researcher during the pretreatment phase from 85 families (76 %) who consented to being videotaped with their children at home with the DPICS-IV procedure. No significant differences were found in sociodemographic characteristics between families who provided consent and those who did not. After excluding siblings and removing videotapes that did not meet the quality criteria for coding (e.g., mother's language; more people in the room), the present study's final clinical sample comprised 80 mother–-child dyads.

2.1.2. Non-clinical sample

Recruitment of the community sample was conducted through six elementary schools located across the region. The school directors distributed the information of the study to families and provided their facilities to conduct the observational sessions with mother–child dyads. The convenience community sample included 97 mothers who agreed to be videotaped with their children. Observation was carried out by the main researcher with the DPICS-IV procedure after participants provided written informed consent. Families in the community sample were asked if they had ever had or currently have contact with Child Protection or child mental health services; none were found to have previous involvement with these services. There was no compensation for participation in the study.

The resulting community sample was not representative of the general population: mothers in this sample had a higher educational level (84.5 % of the community sample had a university degree vs 27.7 % of the Basque population), fewer financial difficulties (3.1 % vs 14.4 %), and were less likely to live in a one-parent household (9.3 % vs 20.61 %) and to be foreign-born (3.1 % vs 12.04 %) than mothers in the Basque population (EUSTAT, 2018).

2.2. Measures

2.2.1. *Dyadic Parent-Child interaction coding System-IV (DPICS-IV clinical version; Eyberg et al., 2014, Cañas et al., 2021)*

The DPICS-IV is a direct behavior observation instrument for assessing the quality of parent–child interaction. This micro-coding system requires videotaping 25 min of dyadic interactions across three tasks: Child-Led Play (CLP, 5 min. warm-up + 5 min. coded segment), Parent-Led Play (PLP, 5 min. warm-up + 5 min. coded segment), and Clean-Up (CU, 5 min. coded segment). These tasks have been designed to elicit the occurrence of the behaviors of interest in a limited length of time by varying parental control levels. For the first 10 min, the child chooses and directs the activity (Child-Led play, CLP); for the next 10 min, the parent takes the lead in the play, gaining more control over the situation (Parent-Led play, PLP). Finally, the parent indicates that it is time to pick up the toys, ending the semi-structured interaction with the maximum parental control level (Clean-Up, CU).

The scoring system of DPICS-IV is based on frequency counts of eight main categories (Neutral talk, Praise -labeled and unlabeled-, Behavior Description, Reflection, Questions, Direct commands, Indirect commands, and Negative talk) across the three tasks. For instance, the Negative talk category is computed through three items, one for each 5-minute coded segment (e.g., Negative talk-CLP, Negative talk-PLP, Negative talk-CU). The Negative talk global score reflects the total number of negative talk instances by a parent within the 15 coded minutes. Therefore, the resulting DPICS-IV variables are quantitative, with values depending on the frequency with which participants exhibit the behavior.

For the present study, we used a Spanish adaptation of the DPICS-IV clinical version with a four-factor structure (Praise, Questions, Clear commands, and Negative talk), with Clear commands being a second-order factor (Compliance and Noncompliance) (Cañas et al., 2021). The Praise factor includes labeled and unlabeled positive evaluations of an attribute, product, or behavior of the child (e.g., The dog you drew is amazing). The Questions factor includes descriptive or reflective comments expressed in the form of a question (e.g., Do you want to draw a dog?). The Clear commands factor contains direct commands (straight request to the child, e.g., Draw a dog in your picture) and indirect commands (suggest a request to the child, e.g., How about drawing a dog?). The Negative talk factor includes critical, sassy or rude statements toward the child (e.g., The dog you drew is a mess) as well as negatively expressed commands telling the child what not to do (e.g., Don't draw on the wall). Finally, the Compliance factor includes commands issued by the caregiver in which the child attempts to obey, while the Noncompliance factor notes commands issued by the caregiver in which the child does not obey. In the present study, the clinical and non-clinical samples received identical instructions based on the standardized DPICS instructions (Eyberg et al., 2014), with minor adaptations to account for the recording environment (homes or schools), which was not the typical PCIT clinical setting (one-way mirror and microphone to deliver instructions to parents via earphone). For example, toys were not in containers or boxes, so the specification of the clean-up situation: "Put all toys in their containers…" was replaced with "Put all toys in their place".

Interrater reliability on DPICS-IV items was completed by the main researcher and a Ph.D. candidate, both with certified training in DPICS-IV, based on the double-coding of 15 % randomly selected videotapes from the clinical sample.

The intraclass correlation coefficients (ICC) are outlined in Table 3 and were all above 0.74. The interpretation of the ICC index establishes the threshold of 0.70 for acceptable ICCs and 0.80 or greater for very good ICCs (Heyman et al., 2014). In this study, the internal consistency for Praise ($\alpha = 0.70$) and Negative Talk ($\alpha = 0.73$) factors was acceptable, while Noncompliance ($\alpha = 0.66$), Clear Commands ($\alpha = 0.62$) and Questions ($\alpha = 0.60$) did not reach the desirable level ($\alpha \geq 0.70$). The Compliance factor showed low internal consistency ($\alpha = 0.40$).

2.2.2. Socio-demographic data

Socio-demographic information of children (gender and age), mothers (age, country of origin, educational level, and work situation) and families (family structure and income) was obtained from mothers' reports.

2.3. Data analysis

As several DPICS-IV items showed moderate to severe non-normal distribution (skewness > 2; kurtosis > 3), the non-parametric Mann Whitney *U* test with rank biserial correlation (small = 0.10, medium = 0.30, and large = 0.50) was used to compare both samples across the DPICS factors and items.

Confirmatory factor analyses (CFA) were conducted to assess the adequacy of DPICS-IV four-factor model (Praise, Questions, Clear commands, and Negative talk) (Cañas et al., 2021). Since some of the DPICS-IV items were severely positively skewed, CFA analyses were performed with root square transformed items to ensure interpretability. This transformation improves the assumptions of normality and homogeneity of variance. Therefore, model goodness-of-fit was assessed by the normal theory maximum-likelihood (ML) chi-square, and the comparative fit index (CFI). To assess the adequacy of model fit, a CFI of 0.90 or above was considered acceptable (Bentler, 2006), and a root mean

square error of approximation (RMSEA) value of 0.06 or less was desirable (Hu & Bentler, 1999). A structural model then positioned clinical status as a predictor of DPICS-IV factors (clinical sample = 1 vs non-clinical sample = 0) in order to identify the DPICS-IV factors that discriminated between clinical and non-clinical samples.

Once best indicators were identified, to provide more evidence of DPICS-IV discriminative or known-groups validity, two separate logistic regression models were conducted. Model 1 was constructed with DPICS-IV factors identified through the previous analyses as significant indicators. The aim was to test the capacity of each factor to correctly classify mother–child dyads. Model 2 included DPICS-IV factors retained by Model 1 and socio-demographic variables as control variables. For this analysis, categorical variables were dichotomized: Mother university degree (0 = No; 1 = Yes), and Employment (0 = No; 1 = Yes). The strength of the relationship between predictor variables and the outcome of interest (clinical sample membership) was analyzed based on the odds ratio (OR) indices. ORs above 1 indicate that the predictor increases the probability that a mother–child dyad would be part of the clinical sample, while ORs below 1 imply that the predictor decreases the likelihood of belonging to that sample. Third, ROC curve analyses were conducted to determine the discriminative accuracy based on the Area Under the Curve (AUC) of the regression model, as well as the individual contributions of each variable retained in the model for clinical status diagnosis. The AUC value is a sound index of discriminative accuracy of a test or a predictor variable, and general guidelines establish the thresholds for excellent (AUC > 0.80), and poor discrimination (AUC < 0.70) (Hosmer et al., 2013).

In addition to showing the diagnostic accuracy of the test, ROC analysis offers the optimal cut-off score for the test. A cut-off score is considered optimal when it classifies most of the individuals correctly. The optimal cut-off score index of sensitivity and specificity should be above the 80 % threshold (Cicchetti et al., 1995). There are several approaches to determine the most appropriate cut-off value in a ROC curve analysis, but the most advisable is the point at which the Youden index is highest (Habibzadeh et al., 2016). Therefore, cut-off scores were calculated based on the Youden index (J = Sensitivity + Specificity – 1) highest value. This method determines the optimal cut-off score by maximizing sensitivity (proportion of clinical sample mother–child dyads classified correctly) and specificity (proportion of non-clinical sample mother–child dyads classified correctly). In the present study the NCSS statistical package was used to generate a complete report of the ROC curve analyses, including a table containing the Youden index for each point on the ROC curve.

Additionally, ROC curve analyses were replicated by splitting the clinical sample into two subgroups: mother–child dyads with a substantiated report of maltreatment and those at risk for child maltreatment, and by comparing each subgroup to the community sample separately.

The univariate data analyses were performed using the SPSS version 26. ROC-curve estimations were calculated using NCSS 2020 statistical software (Number Cruncher Statistical Systems, Kaysville, UT, USA). CFA and structural model analyses were conducted with the EQS (Structural Equation Program) version 6.2 (Multivariate Software Inc., Encino, CA). The transformed DPICS-IV data were only used in the structural equation analyses.

3. Results

3.1. Sample differences on DPICS-IV factors and items

Table 2 summarizes the results of the Mann-Whitney test on DPICS-IV factors and items for both samples. Mothers of the clinical sample exhibited higher rates of Negative talk, across the three items and in the global Negative talk factor, with a large effect size. Also, significant higher levels of Noncompliance were found among clinical mothers-child dyads with a medium effect size across the three items and in the global Noncompliance factor. No differences were found in Praise, Questions and Compliance items and factors between clinical and non-clinical dyads.

3.2. Confirmatory factor analysis

Table 3 reports the ranges, means, standard deviations for transformed DPICS-IV items, and factor loadings in the confirmatory factor analysis. All factor loadings were significant (p <.05). Fit indices for the CFA model which required no model modification were all adequate: ML $\chi 2$ (82, N = 177) = 112.15, p =.41; CFI = 0.932, RMSEA = 0.047, [90 % CI 0.02 - 0.06]. All latent factors and factor loadings were significant (p <.05). The Negative talk factor showed an inverse correlation with the Praise (r = -0.23, p =.02) and Questions factors (r = -0.20, p =.009).

Table 2
Descriptive Statistics for DPICS-IV Items and Mean Differences as a Function of Sample.

Factors and items	Clinical sample (n = 80)				Non-clinical sample (n = 97)				Mean differences test	
	M (*SD*)	Rg	Skw	Krt	*M* (*SD*)	Rg	Skw	Krt	*U*	*r*
PRAISE FACTOR	6.74 (7.81)	0–34	1.77	3.07	6.20 (5.96)	0–28	1.43	1.94	3670.0	0.03
1. Praise CLP	2.84 (4.16)	0–20	1.99	4.10	1.69 (2.10)	0–11	1.77	3.77	3635.5	0.06
2. Praise PLP	2.15 (2.67)	0–11	1.51	1.72	2.36 (3.33)	0–19	2.61	8.95	3854.0	0.01
3. Praise CU	1.92 (3.07)	0–16	2.70	8.82	2.14 (2.33)	0–10	1.28	1.30	3229.5	0.12
QUESTION FACTOR	36.05 (19.99)	0–100	0.92	1.10	35.71 (12.66)	10–73	0.36	0.14	3959.5	0.05
4. Question CLP	15.33 (10.46)	0–43	0.88	0.28	15.05 (7.59)	0–35	0.61	−0.03	4055.0	0.04
5. Question PLP	13.31 (8.47)	0–40	0.86	0.97	13.56 (6.68)	0–34	0.63	0.80	4074.5	0.05
6. Question CU	7.32 (6.23)	0–25	1.00	0.20	7.10 (3.52)	1–17	0.60	0.01	4205.0	0.11
CLEAR COMMAND FACTOR	9.23 (6.81)	0–34	1.44	1.62	7.63 (4.89)	0–24	1.15	1.04	3384.5	0.10
COMPLIANCE FACTOR	5.15 (3.88)	0–18	1.10	1.12	6.05 (3.81)	0–23	1.25	3.13	4417.0	0.16
7. Comply Indirect command-CU	0.70 (1.16)	0–6	0.70	6.75	0.77 (1.04)	0–5	1.70	3.15	4179.5	0.07
8. Comply Direct command-PLP	2.14 (2.17)	0–10	1.60	3.24	1.95 (1.88)	0–7	0.84	0.15	3738.5	0.03
9. Comply Direct command-CU	2.20 (2.08)	0–9	1.36	2.10	2.57 (2.61)	0–17	2.34	9.21	4055.0	0.07
NONCOMPLIANCE FACTOR	4.06 (5.26)	0–29	2.32	6.58	1.59 (2.38)	0–13	2.86	9.66	2452.0**	0.35
10. Noncompliant Indirect command-CU	1.30 (2.12)	0–11	2.70	8.11	0.37 (0.67)	0–3	1.78	2.60	2898.50**	0.25
11. Noncompliant Direct command-PLP	1.22 (1.97)	0–12	2.66	10.64	0.40 (0.90)	0–5	3.12	10.73	2710.0**	0.28
12. Noncompliant Direct command-CU	2.30 (3.33)	0–17	2.13	4.94	0.93 (1.52)	0–8	2.59	7.49	2859.0*	24
NEGATIVE TALK FACTOR	16.69 (12.29)	0–74	1.66	4.99	3.30 (3.71)	0–18	1.66	2.55	762.5**	0.80
13. Negative talk CLP	5.83 (7.85)	0–48	3.00	11.47	0.62 (1.03)	0–4	1.67	2.17	1207.5**	0.69
14. Negative talk PLP	6.06 (5.22)	0–26	1.47	2.30	1.36 (2.13)	0–10	2.38	5.71	1248.0**	0.68
15. Negative talk CU	5.22 (6.40)	0–40	3.20	13.57	1.32 (2.09)	0–9	2.36	5.69	1609.0**	0.57

Note: *Rg* = Range of scores; Skw = Skewness; Krt = Kurtosis; *CLP* = Child-Led Play; *PLP* = Parent-Led Play; *CU* = Clean-Up; *U* = Mann-Whitney *U* test; *r* = rank biserial correlation.
*p <.01; **p <.001.

Table 3
Means, Standard Deviations, Ranges for DPICS-IV Transformed Items, Standardized Factor Loadings and Intra Class Correlation indices.

	Rg	*M*	*SD*	Factor loadings	ICC
PRAISE FACTOR					
Praise CLP	0–3.32	1.02	0.92	0.60	0.93
Praise PLP	0–4.47	1.16	1.11	0.58	0.99
Praise CU	0–4.00	1.05	0.97	0.80	0.99
QUESTION FACTOR					
Question CLP	0–6.56	3.70	1.23	0.59	0.99
Question PLP	0–6.33	3.50	1.10	0.65	0.96
Question CU	0–5.00	2.50	0.97	0.55	0.89
CLEAR COMMAND FACTOR					
COMPLIANCE FACTOR					
Comply Indirect command-CU	0–3.16	0.55	0.66	0.33	0.78
Comply Direct command-PLP	0–2.45	1.16	0.84	0.46	0.81
Comply Direct command-CU	0–4.12	1.31	0.83	0.46	0.82
NONCOMPLIANCE FACTOR					
Noncompliant Indirect command-CU	0–3.74	0.48	0.74	0.57	0.74
Noncompliant Direct command-PLP	0–3.46	0.51	0.72	0.55	0.85
Noncompliant Direct command-CU	0–4.12	0.82	0.93	0.77	0.88
NEGATIVE TALK FACTOR					
Negative talk CLP	0–6.93	1.13	1.30	0.68	0.89
Negative talk PLP	0–5.10	1.41	1.23	0.69	0.80
Negative talk CU	0–6.33	1.28	1.19	0.69	0.98

Note: Rg = range of scores; ICC = Intra class Correlation.

Praise was positively correlated with the Clear commands (r = 0.30, p =.004) and Question factors (r = 0.52, p =.02). Moreover, Clear commands showed a positive correlation with the Questions (r = 0.28, p =.02) and Negative talk (r = 0.61, p =.01) factors.

3.3. Structural model

The structural model is presented in Fig. 1. The DPICS-IV four-factor model was rerun with each latent factor regressed on clinical status (clinical sample = 1 vs non-clinical sample = 0) to analyze the discriminative or known-groups validity. The structural model had good fit statistics: ML, $\chi 2(92, N = 177) = 121.60$, p <.001, CFI = 0.946, RMSEA = 0.04 [90 % CI 0.02 - 0.06]. In this model, clinical mother–child dyads were significantly associated with higher scores on DPICS-IV Negative talk (β = 0.81, p <.001) and Noncompliance (β = 0.42, p <.001) factors.

3.4. Logistic regression analyses

In the first binary logistic regression with DPICS-IV Negative talk and Noncompliance factors which showed predictive capability in structural model analyses, only the Negative talk factor made a unique statistically significant contribution to the model, revealing that the DPICS-IV Noncompliance factor did not add significant information. The Hosmer and Lemeshow test was not significant, showing a good fit of the model [$\chi 2$ (8, n = 170) = 8.76, p =.363] with a medium effect size (Nagelkerke's adjusted value = 0.62).

A second logistic regression model was tested with DPICS-IV Negative talk factor and the socio-demographic variables that showed significant differences between clinical and non-clinical samples on the chi-square test. The stepwise forward Wald method was applied to test variables predicting sample membership. Four variables showed statistical significance: DPICS-IV Negative talk factor, mother origin, mother educational level, and family income (see Table 4). The Hosmer and Lemeshow test indicated a good fit of the estimated model [$\chi 2$ (8, n = 170) = 3.04, p =.932] without multicollinearity problems (VIF = 1.78–1.33), reaching a large effect size in Nagelkerke's adjusted value (r^2 = 0.80). According to the DPICS-IV Negative talk factor OR, each time the Negative talk factor score increased 1 unit, the likelihood that the mother–child dyad belonged to the clinical sample increased by approximately 1.34 times. These results revealed that the Negative talk factor was a robust predictor increasing the probability of being in the clinical sample after controlling for socio-demographic variables.

Table 4
Logistic Regression Predicting clinical or Non-Clinical sample Membership with Pretransform Predictor Variables.

Predictor variables	*B*	*SE*	Wald	*p*	OR	95 % CI
DPICS-IV Negative talk	0.29	0.06	22.99	0.000	1.34	(1.19–1.50)
Mother origin	2.21	0.90	5.99	0.014	9.13	(1.55–53.60)
Mother educational level	2.35	0.59	15.78	0.000	10.44	(3.28–33.23)
Financial difficulties	2.69	1.06	6.48	0.011	14.73	(1.86–116.81)

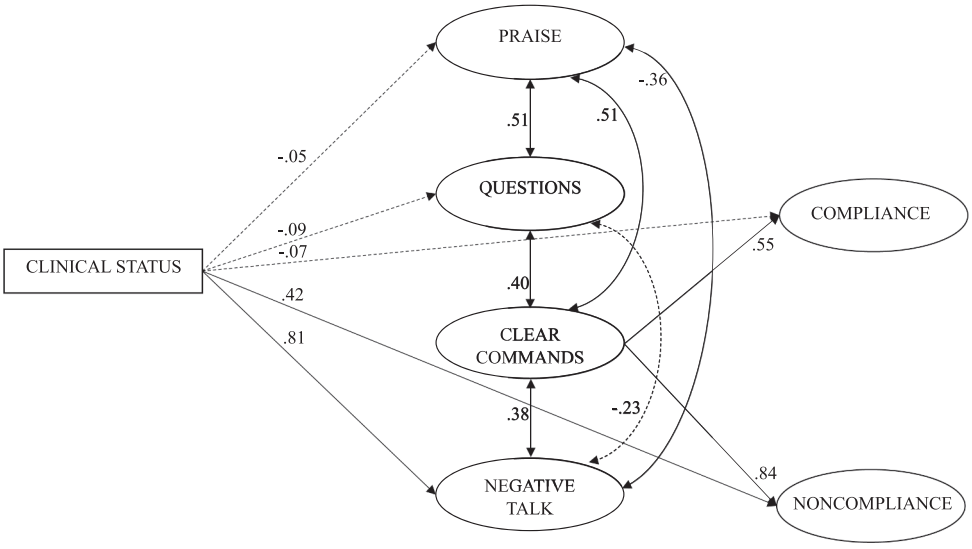


Fig. 1. Structural Model of Clinical Status as a Predictor of DPICS-IV Factors. Note: All estimated parameters are standardized. Latent factor, regression and correlation coefficients are significant at p <.05, except discontinuous lines.

3.5. ROC curve analyses

3.5.1. Area under the curve AUC

The area under the ROC curve for the second regression model yielded an outstanding discrimination level, correctly placing 97 % of the mother–child dyads in their correct sample [AUC = 0.97; 95 % CI = 0.935–0.987; p <.001]. Individual ROC curve analyses for each predictor found only one of the demographic variables, mother educational level, to have an excellent level of discrimination [AUC = 0.83; 95 % CI = 0.76–0.88; p <.001]. However, DPICS-IV Negative talk factor was the predictor with the highest area under the ROC curve, accurately discriminating 90 % of the dyads between non-clinical and clinical samples [AUC = 0.90; 95 % CI = 0.849–0.946; p <.001]. The discriminative capacity of the ROC curve for Negative talk is shown in Fig. 2. According to Hosmer and Lemeshow (2013), this AUC index represents excellent discrimination. Although the Noncompliance factor was not retained in the model, since it had shown a moderate capacity to discriminate between both samples in previous analyses, its area under the ROC curve was calculated to ascertain its discriminant capacity. The results showed an AUC index [AUC = 0.67; 95 % CI = 0.588–0.747; p <.001] at the upper end of the poor discrimination range (e.g., 0.50 < ROC < 0.70).

Additional analyses conducted by splitting the clinical sample into two subgroups showed that the Negative talk factor was able to accurately discriminate the non-clinical mother–child dyads from: (1) mother–child dyads with a substantiated report of maltreatment [AUC = 0.92; 95 % CI = 0.86-0.95; p <.001], and (2) mother child-dyads at risk for child maltreatment [AUC = 0.87; 95 % CI = 0.75-0.96; p <.001]. These results indicate that the Negative talk factor was slightly more accurate with substantiated maltreatment cases, with 92 % of mother–child dyads accurately identified, compared to 87 % of those at risk for maltreatment.

3.5.2. Cut-off score

An optimal cut-off score for the DPICS-IV Negative talk factor was identified using the highest Youden Index (J = 0.68), which is the primary recommended approach (Habibzadeh et al., 2016). Therefore, a value of 8 on Negative talk was identified as the optimal cut-off score. Using the 8-point score, 81 % of mother–child dyads from the clinical sample (sensitivity) and 88 % of mother–child dyads from the non-clinical sample (specificity) were correctly classified. As can be seen in Fig. 3, the overlap between both samples on the distribution of DPICS-IV Negative talk factor scores was small, with 19 % of mother–child dyads from the clinical sample scoring equal or below 8 (false negatives), and 12 % of mother–child dyads from the non-clinical sample scoring above 8 (false positives).

Additional analyses were conducted to determine the optimal cut-off score for the DPICS-IV Negative talk factor with two clinical subgroups separately. The cut-off score of 8 on Negative talk was retained for the substantiated maltreatment condition, with 82 % of mother–child dyads correctly classified (sensitivity) and 88 % correctly classified in the non-clinical sample (specificity). The optimal cut-off score for the mother–-child dyads at risk for child maltreatment was 7, with an 82 % of sensitivity and 85 % of specificity.

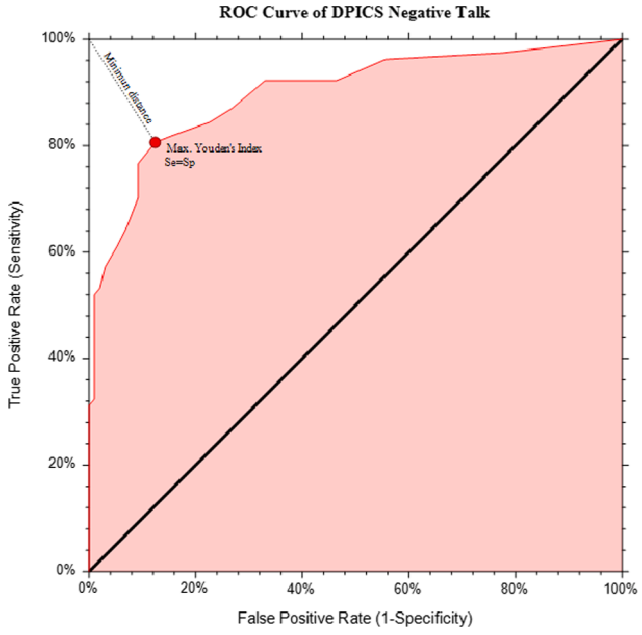


Fig. 2. ROC Curve of the scores on DPICS-IV Negative talk factor for clinical status diagnosis. Note: Area Under the Curve: 0.90.

4. Discussion

The present study aimed to analyze the DPICS-IV's ability to discriminate between parent–child interaction patterns of clinical mother–child dyads with substantiated reports or at risk for child maltreatment, and non-clinical mother–child dyads from a community sample. Findings showed that neither the DPICS-IV as a whole nor four of the six DPICS-IV factors (Praise, Questions, Clear commands, and Compliance) discriminate between both samples. Indeed, only two of the DPICS-IV factors showed statistically significant differences between the two samples: Negative talk (reflecting the mothers negative verbalizations towards the child) and Noncompliance (targeting children disruptive behavior).

As hypothesized, mothers in the clinical sample showed more Negative talk toward their children than mothers in the non-clinical sample, and this indicator presented high accuracy for discriminating between the clinical and non-clinical samples and provided an excellent AUC index [AUC = 0.90]. These results are in line with previous research in which DPICS Negative talk emerged as an indicator of dysfunctional PCI patterns among clinical samples of children with disruptive behavior problems (Bjørseth et al., 2015; McCabe et al., 2010; Quetsch et al., 2018; Robinson & Eyberg, 1981) as well as maltreating families (Aragona & Eyberg, 1981; Borrego et al., 2004; Timmer et al., 2002). Nevertheless, the present study goes beyond previous research by adopting a three-stage analytical approach that strengthened the yielded outcomes. The combination of structural equation models, logistic regression, and ROC curve analysis had not been previously applied to analyzing the discriminative ability of DPICS. The results obtained in each of the analyses were consistent, providing evidence to support the DPICS-IV Negative talk factor's valuable clinical utility for identifying dysfunctional interaction among maltreating and at-risk mother–child dyads. There is empirical evidence from a longitudinal study that supports the present study findings. Webster-Stratton et al. (2011) found that the level of dyad coercion during the post-treatment phase measured with DPICS-R was a significant predictor of maladjustment ten years later. Mothers who showed more than ten critical statements on post-intervention assessments placed their children at greater risk for involvement with the juvenile justice system ten years later. To ensure the best possible outcomes, the authors of that study concluded that interventions should be sustained until the level of dyad coercion falls below the identified critical threshold. These findings suggest that the identification of families with dysfunctional PCI patterns in need of intervention could be improved by including the systematic observation of parent–child interaction though the DPICS-IV Negative talk factor in services working with children and families (e.g., child welfare and child protection services, mental health services).

Also as hypothesized, children in the clinical sample exhibited a higher rate of Noncompliance than children in the non-clinical sample. However, Noncompliance was not a valid discriminatory factor. Indeed, previous studies carried out with samples of children with significant

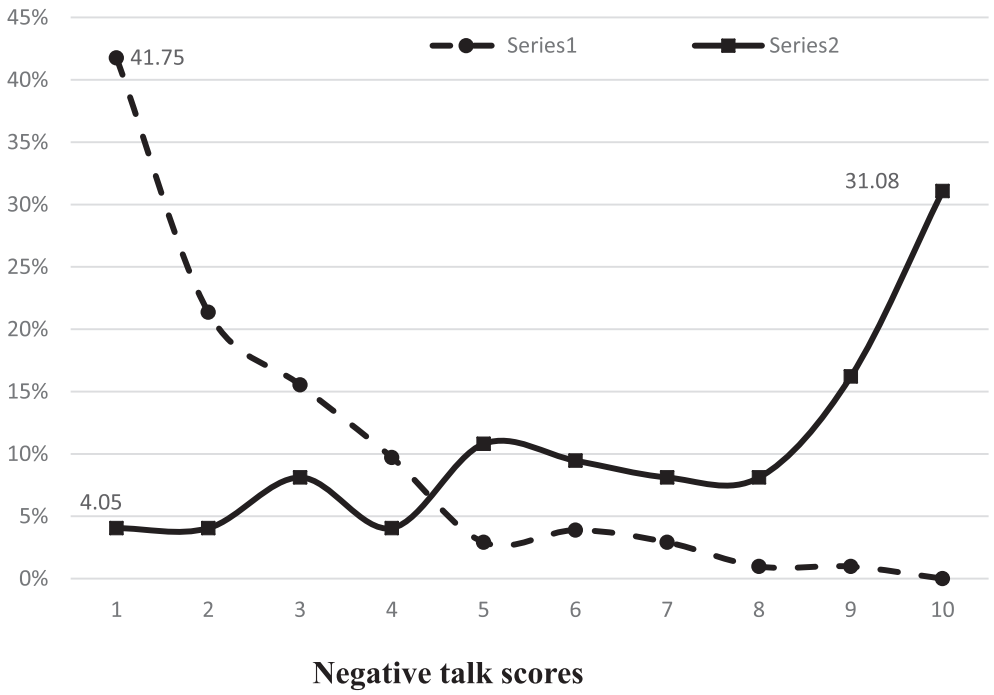


Fig. 3. DPICS-IV Negative Talk Factor Distribution across Clinical and Non-Clinical Samples. Note: This figure is based on equal number of child-mother dyads in the 10 subgroups according to deciles.

behavior problems showed mixed findings related to the DPICS Noncompliance category's discriminant ability. While some studies reported higher rates of Noncompliance among children referred for behavioral problems compared to non-referred (Aragona & Eyberg, 1981; McCabe et al., 2010; Robinson & Eyberg, 1981), other studies found no significant differences (Bjørseth et al., 2015; Quetsch et al., 2018). In the present study, Noncompliance was more frequently observed in children from maltreating and at-risk families than in children from the non-clinical sample, with a medium effect size. Nevertheless, Noncompliance was no longer effective when controlled by the Negative talk factor. Thus, Noncompliance did not add a significant contribution to the discriminant capacity of the model. The results of the current study highlight the reciprocal relationship between these two factors as triggers of dysfunctional PCI patterns. Borrego et al. (2004) pointed out this reciprocity through sequential analyses among physically abusive mother–child dyads measured by DPICS. They found a tendency for abusive mothers to express critical comments after child noncompliance, while receiving criticism from the mother increased the likelihood of child noncompliance. In the present study, although child noncompliance and mother negative talk contributed to dysfunctional patterns of PCI, the mother's effect showed a significantly greater capacity to explain these patterns than the child's disruptive behavior.

Praise has been consistently clustered within the warmth element of parenting skills (McKee et al., 2013), and it could be expected to be associated inversely with mother–child dyads who exhibited patterns of dysfunctional PCI. However, in the present study, no significant differences were found between clinical and non-clinical samples in the DPICS-IV Praise factor. Previous studies with DPICS had shown mixed results in the use of praise across cultures and countries. On the one hand, consistent with our results, studies conducted with Mexican-American parents (McCabe et al., 2010) or carried out in other countries like Norway (Bjørseth et al., 2015) did not find differences between mother–child dyads referred by child behavior problems and non-referred in the use of praise. In contrast, studies conducted among European-American or African-American samples found that clinical mother–child dyads exhibited a lower rate of praise than non-clinical (Aragona & Eyberg, 1981; Borrego et al., 2004; Robinson & Eyberg, 1981; Timmer et al., 2002). Cultural differences might account for the variance in the use of praise by parents. There is some evidence pointing in this direction, finding, for example, that Chinese parents tend to be reluctant to use praise with their children (Leung et al., 2009), and that Mexican-American parents show significantly less praise during interaction than European-Americans (Livas-Dlott et al., 2010; McCabe et al., 2013).

It was expected to identify optimal clinical cut-off scores for the DPICS-IV factors that met sensitivity and specificity standards for mother–child dyads with substantiated maltreatment or at risk for maltreatment. The DPICS-IV Negative talk factor was the best indicator of clinical status. The optimal cut-off score of 8 on Negative Talk accurately discriminated the vast majority of mother–child dyads in this study, with rates of specificity and sensitivity exceeding the recommended standard threshold (80 percent). When the clinical sample was segmented, the cut-off score for substantiated maltreatment cases was sustained, while for at risk for child maltreatment the optimal cut-off score was lowered to 7. Thus, the Negative talk factor emerges as an indicator of dysfunctional interaction in mother–child dyads in both clinical conditions. Nevertheless, these cut-off scores should be considered as preliminary evidence that should be further strengthened in future studies.

4.1. Clinical implications

Best-practice guidelines recommend the inclusion of the parent–child interaction observation as part of the comprehensive assessment of families and children referred to child welfare, child protection, or mental health services (Budd et al., 2011; Frick et al., 2010). Nevertheless, practitioners are reluctant to add direct observational instruments to their assessment protocols due to their high cost in terms of time and resources (Gardner, 2000; Roberts, 2001). Providing less complex and time-consuming measurement instruments could help practitioners to incorporate direct observation of PCI into clinical settings. Thus, the current study's findings provide preliminary evidence of the DPICS-IV Negative talk factor's utility. This factor is a potentially cost-effective method for assessing PCI in mother–child dyads referred

to preventive and early intervention programs by Child Welfare and Child Protection Services. Nevertheless, since DPICS was developed to identify distinct treatment targets for families, the inclusion of all DPICS-IV factors in the assessment would provide more complete information about mother–child interaction patterns. It is essential to point out that a high Negative talk score should not be interpreted as a diagnosis of child maltreatment as it cannot determine that child maltreatment has occurred nor will occur; rather, it should be interpreted as an indicator of dysfunctional mother–child interaction that would add useful information for practitioners within a comprehensive family assessment.

4.2. Limitations and future directions

Findings of the present study should be interpreted in light of some limitations. First, while the observation of the clinical mother–child dyads was conducted in the family home, the observation of the non-clinical dyads was conducted at the children's school. Although both environments are "natural environments" for the children, the differences in the setting in which the observation was conducted were not controlled for in the study design. Future studies should address this issue and control for any potential setting effects. Second, given that the non-clinical sample was not in need of intervention services, it may be more susceptible to volunteer bias than the clinical sample. The willingness to participate in the study of mothers from the non-clinical sample may be driven by a strong interest or commitment to the upbringing of their children. Third, the non-clinical sample was not representative of the general population in socioemographic terms, but the influence of such variables was controlled through statistical analysis. The present study should therefore be replicated with a representative sample of the general population. Fourth, the DPICS's Negative talk factor encompasses conceptually different verbalizations such as mothers critical comments and negatively expressed commands. As a result, the strength with which criticism and negative commands are separately associated with child maltreatment risk status or substantiated child maltreatment cannot be determined. Future studies could address this issue by coding criticism and negative commands separately and examining the extent to which both types of verbalizations serve as indicators of clinical status.

Fifth, while the current study's clinical sample included families at risk for maltreatment or with a substantiated report of maltreatment, no comparisons to other clinical groups were made (e.g., children with behavior problems with no involvement in Child Welfare or Child Protection Services). As a result, we cannot conclude that the observed levels of Negative Talk are unique to maltreating or at-risk families, as such levels may also be indicators of dysfunctional parent–child interaction patterns in other clinical groups. A study conducted in 1985 using the DPCIS-R examined the differences in parent–child interaction patterns between 19 abusive and 21 non-abusive mothers of children with behavioral problems, finding that abusive mothers tended to display significantly higher rates of criticism (17.92) than non-abusive mothers (7.76) (Webster-Stratton, 1985b). Future research could use the DPICS-IV to compare PCI patterns across clinical groups (abusive mothers vs non abusive mothers of children with behavioral problems), replicating the present study's analytic strategy. This could provide evidence to determine the DPICS-IV's capacity to discriminate between specific clinical group PCI patterns.

Sixth, the results could be slightly affected by observer-expectancy bias, given that the primary coder was not blind to the participants' condition. However, observers trained in the use of behavior frequency counting instruments such as DPICS are less prone to observer-expectancy bias than those using global rating systems, which may be particularly susceptible to observer expectations (Aspland & Gardner, 2003; Margolin et al., 1998). Nonetheless, in future studies, it is advisable to ensure that observers are blind to the participants' clinical status to prevent this type of bias. Previous research using the DPICS yielded comparable ICC indices to those obtained in the current study, with values ranging from 0.79 to 0.98 (Thornberry & Brestan-Knight, 2011), from 0.70 to 0.97 (Weeland et al., 2017), from 0.67 to 0.91 (Niec et al., 2016), and from 0.62 to 0.91 (Bjørseth et al., 2015). Nevertheless, inter-rater reliability was established on a random selection of 15 % of DPICS videotapes. Although there is no gold standard percentage of videos to double-code for establishing inter-rater reliability, in the scientific literature, we find that this sub-sample usually ranges between 20 % and 30 % of the total videos (Breitenstein et al., 2012; McCabe et al., 2021; Niec et al., 2016;). Therefore, it is advisable for future studies to establish inter-rater reliability by double-coding around the 25 % of total videotapes.

Finally, because the present study focused exclusively on mother–-child dyads, the results are not generalized to father-child dyads. The paucity of studies analyzing the psychometric properties of observational instruments with father-child dyads has been identified as a critical issue that needs to be addressed (Lotzin et al., 2015). Therefore, this study should be replicated with a broader father-child dyads sample to ascertain whether findings hold or vary by caregiver figure.

5. Conclusions

The findings of the present study contribute to the scientific literature of parent–child interaction assessment in several ways. First, the evidence supporting the DPICS-IV Negative talk factor's discriminant ability was gathered through robust analytic approaches and by correctly controlling for the potential effects of sociodemographic variables. Second, an optimal cut-off score of 8 was set up for the DPICS-IV Negative talk factor, with an area under the ROC curve of 0.90 (95 % CI, 0.85-0.95). Third, the present study addresses the scarcity of studies targeting the strengthening of DPICS-IV discriminative or known-groups validity by adding valuable evidence regarding the instrument's usefulness. However, additional research is required to strengthen the factorial validity of DPICS-IV since only one published study has so far provided evidence analyzing this psychometric property. Also, analysis of DPICS-IV's incremental validity warrants further study. It is therefore essential to determine the DPICS-IV's ability to collect unique and meaningful information that self-reported measures cannot capture. Studies addressing this issue could enhance the use of the DPICS-IV to assist practitioners in conducting comprehensive family assessments. Although further research is still needed to strengthen the evidence of its psychometric properties, the DPICS-IV has the potential to be used in the context of Child Welfare and Child Protection Services to assess parent–child interaction patterns of mother–child dyads, to define intervention goals, and to monitor progress throughout the intervention.

CRediT authorship contribution statement

María Cañas: Conceptualization, Methodology, Software, Data curation, Formal analysis, Writing – original draft, Visualization, Funding acquisition. **Izaskun Ibabe:** Conceptualization, Methodology, Writing – review & editing, Supervision. **Ignacia Arruabarrena:** Writing – review & editing, Supervision, Funding acquisition. **Joaquín De Paúl:** Conceptualization, Supervision, Funding acquisition, Conceptualization, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was financially supported by a grant from the Ministry of Economy and Competitiveness (MINECO) of Spain [Reference: PSI2013-46272-P], and the Basque Government Research Personnel Education

and Training Program scholarship granted to the first author [PRE_2017_2_050].

Funding Sources

This work was financially supported by a grant from the Ministry of Economy and Competitiveness (MINECO) of Spain [Reference: PSI2013-46272-P], and the Basque Government Research Personnel Education and Training Program scholarship granted to the first author [PRE_2017_2_050].

References

Akcinar, B., & Shaw, D. S. (2018). Independent Contributions of Early Positive Parenting and Mother-Son Coercion on Emerging Social Development. *Child Psychiatry & Human Development, 49*(3), 385–395. https://doi.org/10.1007/s10578-017-0758-4

Allen, B., Timmer, S. G., & Urquiza, A. J. (2014). Parent-Child Interaction Therapy as an attachment-based intervention: Theoretical rationale and pilot data with adopted children. *Children and Youth Services Review, 47*(3), 334–341. https://doi.org/10.1016/j.childyouth.2014.10.009

Aragona, J. A., & Eyberg, S. M. (1981). Neglected children: Mothers' report of child behavior problems and observed verbal behavior. *Child Development, 52*(2), 596–602. https://doi.org/10.2307/1129179

Arruabarrena, I., Rivas, G. R., Cañas, M., & De Paúl, J. (2022). The incredible years parenting and child treatment programs: A randomized controlled trial in a child welfare setting in Spain. *Psychosocial Intervention, 31*(1), 43–58.

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

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, *2*(4), 396–403. 10.1111/j.1745-6916.2007.00051.x.

Belsky, J., Bell, B., Bradley, R. H., Stallard, N., & Stewart-Brown, S. L. (2007). Socioeconomic risk, parenting during the preschool years and child health age 6 years. *European Journal of Public Health, 17*(5), 508–513. https://doi.org/10.1093/eurpub/ckl261

Bentler, P. M. (2006). EQS 6 structural equations program manual. In *Los Angeles: BMDP Statistic Software* (Issue 818). http://www.econ.upf.edu/~satorra/CourseSEMVienna2010/EQSManual.pdf.

Bjørseth, Å., McNeil, C., & 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. https://doi.org/10.1080/07317107.2015.1000228

Bocknek, E. L., Brophy-Herb, H. E., & Banerjee, M. (2009). Effects of parental supportiveness on toddlers' emotion regulation over the first three years of life in a low-income African American sample. *Infant Mental Health Journal, 30*(5), 452–476. https://doi.org/10.1002/imhj.20224

Borrego, J., Timmer, S. G., Urquiza, A. J., & Follette, W. C. (2004). Physically abusive mothers' responses following episodes of child noncompliance and compliance. *Journal of Consulting and Clinical Psychology, 72*(5), 897–903. https://doi.org/10.1037/0022-006X.72.5.897

Breitenstein, S. M., Gross, D., Fogg, L., Ridge, A., Garvey, C., Julion, W., & Tucker, S. (2012). The chicago parent program: Comparing 1-year outcomes for african american and latino parents of young children. *Research in Nursing & Health, 35*(5), 475–489.

Budd, K. S., Connell, M., & Clark, J. (2011). Evaluation of parenting capacity in child protection. *Oxford University Press*. https://doi.org/10.1093/med:psych/9780195333602.001.0001

Cañas, M., Ibabe, I., & De Paúl, J. (2020). Promising observational instruments of parent-child (0–12 years) interaction within the child protection system: A systematic review. *Child Abuse & Neglect, 109*, Article 104713. https://doi.org/10.1016/j.chiabu.2020.104713

Cañas, M., Ibabe, I., Arruabarrena, I., & De Paúl, J. (2021). Dyadic parent-child interaction coding system (Dpics): Factorial structure and concurrent validity. *Psicothema, 33*(2), 328–336. https://doi.org/10.7334/psicothema2020.429

Chen, M., & Chan, K. L. (2016). Effects of parenting programs on child maltreatment prevention: A meta-analysis. *Trauma, Violence, and Abuse, 17*(1), 88–104. https://doi.org/10.1177/1524838014566718

Cicchetti, D. V., Volkmar, F., Klin, A., & Showalter, D. (1995). Diagnosing autism using icd-10 criteria: A comparison of neural networks and standard multivariate procedures. *Child Neuropsychology, 1*(1), 26–37. https://doi.org/10.1080/09297049508401340

Dietz, L. J., Jennings, K. D., Kelley, S. A., & Marshal, M. (2009). Maternal depression, paternal psychopathology, and toddlers' behavior problems. *Journal of Clinical Child and Adolescent Psychology, 38*(1), 48–61. https://doi.org/10.1080/15374410802575362

Dowdy, E., Twyford, J., & Sharkey, J. D. (2013). Methods of assessing behavior: Observations and rating scales. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwean (Eds.), *The Oxford handbook of child psychological assessment* (pp. 623–650). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199796304.013.0026.

Easterbrooks, M. A., Bureau, J. F., & Lyons-Ruth, K. (2012). Developmental correlates and predictors of emotional availability in mother-child interaction: A longitudinal study from infancy to middle childhood. *Development and Psychopathology, 24*(1), 65–78. https://doi.org/10.1017/S0954579411000666

Eddy, J. M., Leve, L. D., & Fagot, B. I. (2001). Coercive family processes: A replication and extension of patterson's coercion model. *Aggressive Behavior, 27*(1), 14–25.

EUSTAT - Euskal Estatistika Erakundea - Instituto Vasco de Estadística (2018). Datos estadísticos de la C.A. de Euskadi https://www.eustat.eus/indice.html.

Eyberg, S. M., Chase, R. M., Fernandez, M. A., & Nelson, M. M. (2014). *Dyadic Parent-Child Interaction Coding System (DPICS-IV) Clinical Manual Fourth Edition*. PCIT International.

Eyberg, S. M., Nelson, M. M., Ginn, N. C., Bhuiyan, N., & Boggs, S. R. (2013). *Dyadic Parent-Child Interaction Coding System (DPICS): Comprehensive manual for research and training* (Vol. 2013). PCIT International.

Falceto, O. G., Giugliani, E. R. J., & Fernandes, C. L. C. (2012). Problematic parent-infant relationships in two-parent families: Prevalence and risk factors in a Brazilian neighborhood. *Trends in Psychiatry and Psychotherapy, 34*(3), 139–146. https://doi.org/10.1590/s2237-60892012000300005

Fraley, R. C., Roisman, G. I., & Haltigan, J. D. (2013). The legacy of early experiences in development: Formalizing alternative models of how early experiences are carried forward over time. *Developmental Psychology, 49*(1), 109–126. https://doi.org/10.1037/a0027852

Frick, P. J., Barry, C. T., & Kamphaus, R. W. (2010). *Clinical assessment of child and adolescent personality and behavior* (3rd ed.). Springer Science + Business Media.

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. https://doi.org/10.1023/A:1009503409699

Gardner, F., Leijten, P., Harris, V., Mann, J., Hutchings, J., Beecham, J., … Landau, S. (2019). Equity effects of parenting interventions for child conduct problems: A pan-European individual participant data meta-analysis. *The Lancet Psychiatry, 6*(6), 518–527. https://doi.org/10.1016/S2215-0366(19)30162-2

Goldstein, L. H., Harvey, E. A., & Friedman-Weieneth, J. L. (2007). Examining subtypes of behavior problems among 3-year-old children, Part III: Investigating differences in parenting practices and parenting stress. *Journal of Abnormal Child Psychology, 35*(1), 125–136. https://doi.org/10.1007/s10802-006-9047-6

Habibzadeh, F., Habibzadeh, P., & Yadollahie, M. (2016). On determining the most appropriate test cut-off value: The case of tests with continuous results. *Biochemia Medica*, *26*(3), 297–307. 10.11613/BM.2016.034.

Heyman, R. E., Lorber, M. F., Eddy, J. M., & West, T. V. (2014). Behavioral observation and coding. In H. T. Reis, & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd ed., pp. 345–372). Cambridge University Press.

Holstein, B. E., Pant, S. W., Ammitzbøll, J., & Pedersen, T. P. (2021). Social inequality in parent–infant relations: Epidemiological study of community nurse records. *Scandinavian Journal of Public Health*, *July 2020*, 140349482098313. 10.1177/1403494820983137.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Wiley. 10.1002/9781118548387.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hurley, K. D., Huscroft-D'Angelo, J., Trout, A., Griffith, A., & Epstein, M. (2014). Assessing parenting skills and attitudes: A review of the psychometrics of parenting measures. *Journal of Child and Family Studies, 23*(5), 812–823. https://doi.org/10.1007/s10826-013-9733-2

Jeon, H. J., Peterson, C. A., & DeCoster, J. (2013). Parent-child interaction, task-oriented regulation, and cognitive development in toddlers facing developmental risks. *Journal of Applied Developmental Psychology, 34*(6), 257–267. https://doi.org/10.1016/j.appdev.2013.08.002

Kelley, M. L., Lawrence, H. R., Milletich, R. J., Hollis, B. F., & Henson, J. M. (2015). Modeling risk for child abuse and harsh parenting in families with depressed and substance-abusing parents. *Child Abuse and Neglect, 43*, 42–52. https://doi.org/10.1016/j.chiabu.2015.01.017

Larrieu, J. A., Middleton, M. A., Kelley, A. C., & Zeanah, C. H. (2018). Assessing the relational context of infants and young children. In C. H. Zeanah (Ed.), *Handbook of Infant Mental Health* (4th ed., pp. 279–296). The Guilford Press.

Leung, C., Tsang, S., Heung, K., & Yiu, I. (2009). Effectiveness of Parent—Child Interaction Therapy (PCIT) among chinese families. *Research on Social Work Practice, 19*(3), 304–313. https://doi.org/10.1177/1049731508321713

Lindahl, K. M., Malik, N. M., & Wigderson, S. (2019). Observational methods: Overview. In B. H. Fiese, M. Celano, K. Deater-Deckard, E. N. Jouriles, & M. A. Whisman (Eds.), *APA handbooks in psychology®. APA handbook of contemporary family psychology: Foundations, methods, and contemporary issues across the lifespan* (pp. 281–297). American Psychological Association. https://doi.org/10.1037/0000099-016.

Livas-Dlott, A., Fuller, B., Stein, G. L., Bridges, M., Mangual Figueroa, A., & Mireles, L. (2010). Commands, competence, and cariño: Maternal socialization practices in mexican american families. *Developmental Psychology, 46*(3), 566–578. https://doi.org/10.1037/a0018016

Lotzin, A., Lu, X., Kriston, L., Schiborr, J., Musal, T., Romer, G., & Ramsauer, B. (2015). Observational tools for measuring parent–infant interaction: A systematic review. *Clinical Child and Family Psychology Review, 18*(2), 99–132. https://doi.org/10.1007/s10567-015-0180-z

Lovejoy, M. C., Graczyk, P. A., O'Hare, E., & Neuman, G. (2000). Maternal depression and parenting behavior: A meta-analytic review. *Clinical Psychology Review, 20*(5), 561–592. https://doi.org/10.1016/S0272-7358(98)00100-7

Margolin, G., Oliver, P. H., Gordis, E. B., O'Hearn, H. G., Medina, A. M., Ghosh, C. M., & Morland, L. (1998). The nuts and bolts of behavioral observation of marital and family interaction. *Clinical Child and Family Psychology Review, 1*(4), 195–213.

McCabe, K. M., Sakamoto, M., Rosas, Y. G., Kehoe, K., La, R., Zerr, A., & Yeh, M. (2021). Keeping an "I" on PRIDE: Measuring imitation in parent-child interaction therapy. *Behavior Therapy, 52*(1), 28–38.

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 & Family Behavior Therapy, 35*(4), 307–326. https://doi.org/10.1080/07317107.2013.846680

McCabe, K., Yeh, M., Lau, A., Argote, C. B., & Liang, J. (2010). Parent-Child Interactions Among low-income mexican american parents and preschoolers: Do clinic-referred families differ from nonreferred families? *Behavior Therapy, 41*(1), 82–92. https://doi.org/10.1016/j.beth.2009.01.003

McCoy, A., Melendez-Torres, G. J., & Gardner, F. (2020). Parenting interventions to prevent violence against children in low- and middle-income countries in East and Southeast Asia: A systematic review and multi-level meta-analysis. *Child Abuse and Neglect, 103*, Article 104444. https://doi.org/10.1016/j.chiabu.2020.104444

McKee, L. G., Jones, D. J., Forehand, R., & Cuellar, J. (2013). Assessment of parenting behaviors and style, parenting relationships, and other parent variables in child assessment. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwean (Eds.), *The Oxford handbook of child psychological assessment* (pp. 788–821). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199796304.013.0035.

Mokkink, L. B., de Vet, H. C. W., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Quality of Life Research, 27*(5), 1171–1179. https://doi.org/10.1007/s11136-017-1765-4

Muzik, M., Morelen, D., Hruschak, J., Rosenblum, K. L., Bocknek, E., & Beeghly, M. (2017). Psychopathology and parenting: An examination of perceived and observed parenting in mothers with depression and PTSD. *Journal of Affective Disorders, 207*, 242–250.

Morsbach, S. K., & Prinz, R. J. (2006). Understanding and improving the validity of self-report of parenting. *Clinical Child and Family Psychology Review, 9*(1), 1–21. https://doi.org/10.1007/s10567-006-0001-5

Nelson M.M., Olsen B. (2018). Dyadic Parent–Child Interaction Coding System (DPICS): An adaptable measure of parent and child behavior during dyadic interactions. In Niec, L. (Eds.), *Handbook of Parent-Child Interaction Therapy* (p. 285-302). Springer. 10.1007/978-3-319-97698-3_18.

Niec, L. N., Barnett, M. L., Prewett, M. S., & Shanley Chatham, J. R. (2016). Group parent–child interaction therapy: A randomized control trial for the treatment of conduct problems in young children. *Journal of Consulting and Clinical Psychology, 84*(8), 682–698.

Nilsen, F. M., Ruiz, J. D. C., & Tulve, N. S. (2020). A meta-analysis of stressors from the total environment associated with children's general cognitive ability. *International Journal of Environmental Research and Public Health, 17*(15), 1–34. https://doi.org/10.3390/ijerph17155451

Office on Child Abuse and Neglect (2018). *Child Protective Services : A Guide for Caseworkers*. U.S. Department of Health and Human Services. https://www.childwelfare.gov/pubPDFs/cps2018.pdf.

Pinquart, M. (2017). Associations of parenting dimensions and styles with externalizing problems of children and adolescents: An updated meta-analysis. *Developmental Psychology, 53*(5), 873–932. https://doi.org/10.1037/dev0000295

Quetsch, L. B., Wallace, N. M., McNeil, C. B., & Gentzler, A. L. (2018). Emotion regulation in families of children with behavior problems and nonclinical comparisons. *Journal of Child and Family Studies, 27*(8), 2467–2480. https://doi.org/10.1007/s10826-018-1081-9

Raby, K. L., Roisman, G. I., Fraley, R. C., & Simpson, J. A. (2015). The enduring predictive significance of early maternal sensitivity: Social and academic competence through age 32 years. *Child Development, 86*(3), 695–708. https://doi.org/10.1111/cdev.12325

Ramsauer, B., & Achtergarde, S. (2018). Mothers with acute and chronic postpartum psychoses and impact on the mother-infant interaction. *Schizophrenia Research, 197*, 45–58. https://doi.org/10.1016/j.schres.2018.02.032

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

Robinson, E. A., & Eyberg, S. M. (1981). The Dyadic Parent-Child Interaction coding system: Standardization and validation. *Journal of Consulting and Clinical Psychology, 49*(2), 245–250. https://doi.org/10.1037//0022-006X.49.2.245

Romanowicz, M., Vande Voort, J. L., Shekunov, J., Oesterle, T. S., Thusius, N. J., Rummans, T. A., … Schak, K. M. (2019). The effects of parental opioid use on the parent-child relationship and children's developmental and behavioral outcomes: A systematic review of published reports. *Child and Adolescent Psychiatry and Mental Health, 13*(1), 1–11. https://doi.org/10.1186/s13034-019-0266-3

Sameroff, A. J., & Fiese, B. H. (2000). Transactional regulation: The developmental ecology of early intervention. In J. P. Shonkoff, & S. J. Meisels (Eds.), *Handbook of early childhood intervention* ((2nd ed.), pp. 135–159). Cambridge University Press. https://doi.org/10.1017/CBO9780511529320.009.

Schulz, S., Leijten, P., Shaw, D. S., & Overbeek, G. (2019). Parental reactivity to disruptive behavior in toddlerhood: An experimental study. *Journal of Abnormal Child Psychology, 47*(5), 779–790. https://doi.org/10.1007/s10802-018-0489-4

Shah, R., Kennedy, S., Clark, M. D., Bauer, S. C., & Schwartz, A. (2016). Primary care-based interventions to promote positive parenting behaviors: A meta-analysis. *Pediatrics, 137*(5), e20153393–e. https://doi.org/10.1542/peds.2015-3393

Skovgaard, A. M., Olsen, E. M., Christiansen, E., Houmann, T., Landorph, S. L., & Jørgensen, T. (2008). Predictors (0–10 months) of psychopathology at age 1½ years – a general population study in The Copenhagen Child Cohort CCC 2000*. *Journal of Child Psychology and Psychiatry, 49*(5), 553–562. https://doi.org/10.1111/j.1469-7610.2007.01860.x

Skovgaard, A. M., Houmann, T., Christiansen, E., Landorph, S., Jørgensen, T., Olsen, E. M., … Lichtenberg, A. (2007). The prevalence of mental health problems in children 1½ years of age - the Copenhagen Child Cohort 2000. *Journal of Child Psychology and Psychiatry, 48*(1), 62–70. https://doi.org/10.1111/j.1469-7610.2006.01659.x

Slesnick, N., Feng, X., Brakenhoff, B., & Brigham, G. S. (2014). Parenting under the influence: The effects of opioids, alcohol and cocaine on mother-child interaction. *Addictive Behaviors, 39*(5), 897–900. https://doi.org/10.1016/j.addbeh.2014.02.003

Stewart-Brown, S. L., Fletcher, L., & Wadsworth, M. E. J. (2005). Parent-child relationships and health problems in adulthood in three UK national birth cohort studies. *European Journal of Public Health, 15*(6), 640–646. https://doi.org/10.1093/eurpub/cki049

Stith, S. M., Liu, T., Davies, L. C., Boykin, E. L., Alder, M. C., Harris, J. M., … Dees, J. E. M. E. G. (2009). Risk factors in child maltreatment: A meta-analytic review of the literature. *Aggression and Violent Behavior, 14*(1), 13–29. https://doi.org/10.1016/j.avb.2006.03.006

Swenson, S., Ho, G. W. K., Budhathoki, C., Belcher, H. M. E., Tucker, S., Miller, K., & Gross, D. (2016). Parents' use of praise and criticism in a sample of young children seeking mental health services. *Journal of Pediatric Health Care, 30*(1), 49–56. https://doi.org/10.1016/j.pedhc.2015.09.010

Timmer, S. G., Borrego, J., & Urquiza, A. J. (2002). Antecedents of coercive interactions in physically abusive mother-child dyads. *Journal of Interpersonal Violence, 17*(8), 836–853. https://doi.org/10.1177/0886260502017008003

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.

Vlahovicova, K., Melendez-Torres, G. J., Leijten, P., Knerr, W., & Gardner, F. (2017). Parenting programs for the prevention of child physical abuse recurrence: A systematic review and meta-analysis. *Clinical Child and Family Psychology Review, 20*(3), 351–365. https://doi.org/10.1007/s10567-017-0232-7

Webster-Stratton, C. (1985a). Mother perceptions and mother-child interactions: Comparison of a clinic-referred and a nonclinic group. *Journal of Clinical Child Psychology, 14*(4), 334–339. https://doi.org/10.1207/s15374424jccp1404_11

Webster-Stratton, C. (1985b). Comparison of abusive and nonabusive families with conduct-disordered children. *American Journal of Orthopsychiatry, 55*(1), 59–69.

Webster-Stratton, C., Rinaldi, J., & Reid, J. M. (2011). Long-term outcomes of Incredible Years parenting program: Predictors of adolescent adjustment. *Child and Adolescent Mental Health, 16*(1), 38–46.

Weeland, J., Chhangur, R. R., van der Giessen, D., Matthys, W., Orobio de Castro, B., & Overbeek, G. (2017). Intervention effectiveness of the incredible years: New insights into sociodemographic and intervention-based moderators. *Behavior Therapy, 48*(1), 1–18.

Werner, C. D., Linting, M., Vermeer, H. J., & Van Ijzendoorn, M. H. (2016). Do intervention programs in child care promote the quality of caregiver-child interactions? A meta-analysis of randomized controlled trials. *Prevention Science, 17*(2), 259–273. https://doi.org/10.1007/s11121-015-0602-7

Wilson, S., & Durbin, C. E. (2012). The laboratory parenting assessment battery: Development and preliminary validation of an observational parenting rating system. *Psychological Assessment, 24*(4), 823–832. https://doi.org/10.1037/a0028352