CrossMark

ORIGINAL PAPER

The Observation of Child Behavior During Parent-Child Interaction: The Psychometric Properties of the Crowell Procedure

Laurie Loop[1] · Bénédicte Mouton[1] · Elise Brassart[1] · Isabelle Roskam[1]

Abstract The observation of child behavior has to be made in ecologically valid contexts. Parent-child interaction was thought to be the most suitable context, since empirical evidence displayed strong associations with children's behavioral outcomes, psychopathology, social relationships and well-being. Using clinical data from 137 caregiver-child dyads, the main goal of the current study was to test the psychometric properties of an adapted version of the Crowell Procedure among preschoolers. Despite the interest that the Crowell Procedure has aroused, its psychometric properties remain relatively under-documented. This research aimed to study (1) the association between parental and child behavior, (2) the discriminant properties of the Crowell Procedure between preschoolers with a clinical level of externalizing behavior and non-clinical children and (3) the correlation between the Crowell Procedure and a behavioral checklist. The results support the consistency of both tasks and scales, the discriminant properties, external validity and fidelity of the coding system. The Crowell Procedure can therefore be used as an observational paradigm to assess both child and parent behavior in clinical and research contexts. The discriminant analyses revealed that the procedure was effective at differentiating children displaying a clinical level of externalizing behavior from normally developing ones.

Keywords Child behavior · Observational paradigm · Parent-child interaction · Psychometric properties · Preschoolers

✉ Laurie Loop
  laurie.loop@uclouvain.be

[1] Psychological Sciences Research Institute, University of Louvain, 10 Place du Cardinal Mercier, Louvain-la-Neuve 1348, Belgium

Introduction

The valid assessment of child behavior remains a central concern for both clinicians and researchers. It is often achieved with questionnaires consisting of behavioral checklists completed by the caregivers, such as the widely used Child Behavior Checklist (Achenbach and Rescorla 2000a) or the Strengths and Difficulties Questionnaire (Goodman 2001). However, the use of caregiver reports is somewhat limited by informant bias and recall (De Los Reyes and Kazdin 2005). An alternative to questionnaires is provided by observational paradigms administered in a standardized manner. Observational paradigms are intentionally structured to increase the likelihood that a range of relevant behaviors will emerge. Strong arguments have been put forward for observing child behavior in ecologically valid contexts (Lieberman et al. 2000; Sameroff and Fiese 2000; Sroufe 1989). However, what is the most relevant context for conducting such observations? According to the epigenetic view of child development, the most relevant context is parent-child interaction, since it is highly predictive of all subsequent social relationships (Cassidy 2008; Roskam et al. 2015), children's behavioral outcomes (Belsky 1984; Criss et al. 2003), psychopathology (Lieberman 2004) and well-being (Zeanah et al. 2000). In line with these methodological and theoretical recommendations, Crowell and colleagues (Crowell and Feldman 1988; Crowell et al. 1988) proposed the Crowell Procedure, designed to observe child as well as parent behavior during parent-child interaction.

The Crowell Procedure (Crowell and Feldman 1988) was used initially for children from 12 to 60 months of age (Miron et al. 2009; Osofsky et al. 2007). The procedure involves unstructured tasks (such as free play) and structured tasks (teaching tasks like puzzles), allowing the

dyad to exhibit their regular relationship patterns and interactional routines. The procedure also ends with a separation-reunion phase. After the meeting, the parents fill in a brief questionnaire to test two things: the extent to which the parents' and the children's behavior in the lab is comparable with daily life, and the extent to which the parents have confidence in the experimenter. Confidence in the experimenter is supposed to limit participants' resistance to the procedure and to increase the probability of observing behaviors which are representative of authentic parent-child interaction. As suggested by Miron et al. (2009), the unstructured free play episode may reveal the familiarity of the dyad with play and their use of the time as fun-oriented. The procedure also involves transitions designed to induce variety in observational settings and stress in the interaction (Sprang and Craig 2014). The procedure is videotaped, permitting children's and parents' behavior to be coded under blind conditions by independent trained coders. The initial coding system was based on the assessment of nine variables relating to child behavior, i.e. enthusiasm, persistence, self-reliance, affection, negativity, avoidance, controlling behavior, anxiety and compliance. In addition to the child's behavior, the mother's help and support were estimated (Crowell et al. 1988). This coding system has been adapted on several occasions (for an example see Osofsky et al. 2003), in particular by Heller et al. (1999). Their version is based on standardized rating scales, in particular seven child scales, i.e. positive affect, withdrawal/indifference, irritability/anger, non-compliance, aggression, persistence and enthusiasm, and six parent scales, i.e. behavioral responsiveness, emotional responsiveness, positive affect, withdrawal/disinterest, irritability/anger and aggression. The Crowell Procedure scales provide an assessment of both child and parent behavior in a particular interactional setting.

Crowell's work has inspired numerous subsequent studies: it has been cited in more than 170 studies published in peer-reviewed journals (source PsycINFO). These studies have been conducted with toddlers and their parents in community samples (Coleman and Karraker 2003; Conway et al. 2014; Miller et al. 2002), high-risk samples such as disadvantaged (Aoki et al. 2002) or maltreated children (Malik et al. 2002; Osofsky et al. 2007; Robinson et al. 2009) and children in foster care (Zeanah et al. 2003). Research has also been conducted among preschoolers in community samples (Mouton and Roskam 2014; Verschueren et al. 2006) and high-risk samples such as disadvantaged populations (Brassart and Schelstraete 2015a), children with post-traumatic stress (Scheeringa et al. 2004) exposed *in utero* to psychotropic medications (Misri et al. 2006) and incest survivors (Fitzgerald et al. 2005). Likewise, the Crowell Procedure has been used with referred for

depression (Luby et al. 2006) or externalizing behavior (Brassart and Schelstraete 2015b; Roskam et al. 2015). As well as having been received with enthusiasm by researchers, the Crowell Procedure is also popular for clinical use (Miron et al. 2009; Sprang et al. 2004). Despite the interest that the procedure has aroused, its psychometric properties remain relatively under-documented (Miron et al. 2009), although it has been shown to discriminate between clinical and non-clinical parent-child dyads (Crowell and Feldman 1988) or maltreated and healthy children (Smyke 2000). To the best of our knowledge, only one recent study has reported a psychometric analysis of the Crowell Procedure (Sprang and Craig 2014). This was conducted with 151 caregiver-child dyads referred by the child welfare system in the United States. The results of this study demonstrated the usefulness of the procedure and its coding system in a clinical setting. They gave support to the reliability of both child and parent rating scales and suggested that scores could be summarized in two global scales, one for the child and the other for the parent's behavior.

The validation of standardized observational measures such as the Crowell Procedure is very important for both research and clinical purposes. Without it, views on how to carry out observation of parent-child interaction and what should be assessed would be subjective (Budd et al. 2001; Hynan 2003). A lack of agreement would persist about which aspects of child and parent behavior are the most relevant. In the absence of a validated coding system, observation measures would also remain qualitative, conditioned by the rater's skills, knowledge and experience (Schmidt et al. 2007). Finally, validation studies are very helpful for identifying the most discriminant tasks and reliable rating scales in order to define the most effective and focused observation scheme. Without such research, observational procedures tend to be long and time-consuming, making them impossible to implement for both clinicians and researchers.

Therefore, the current study provides a psychometric analysis of an adapted version of the Crowell Procedure through (1) the association between parental behavior and child behavior, (2) the discriminant properties of the Crowell Procedure between preschoolers with clinical levels of externalizing behavior and non-clinical children and (3) the correlation between the Crowell Procedure and a behavioral checklist (Crowell et al. 1988). In line with Sprang and Craig (2014), it also tests the reliability of the child and parent rating scales. Finally, it aims to go further than these previous psychometric findings by testing the inter-task relations, test-retest fidelity, inter-rater agreement and correlations between the parental scales of the Crowell Procedure and one self-reported measure of parenting behavior.

Method

Participants

The study is part of the Hard-t(w)o-Manage (H2M) Children research program conducted at the Université catholique de Louvain and approved by the ethics committee. Data were collected from 137 parent-child dyads from a community sample. Children were between 36 and 72 months old ($M = 53.96$, SD = 8.08) and 58.4 % were boys. The parents were aged between 26 and 54 ($M = 37.25$, SD = 5.20) and 83.9 % were mothers. The educational level of the parent participating was calculated as the number of years of education he/she had completed, counting from first grade onwards. Some had completed 6 years (2.2 %); others had completed 12 years, corresponding to the end of secondary school and compulsory education in Belgium (16.8 %); others had completed 3 more years (corresponding to undergraduate studies) (22.6 %); others had gained a 4-year degree or more (58.4 %). Monthly incomes were less than €2500 for 24.8 % of the families, between €2500 and €4000 for 35.8 % and higher than €4000 for 39.4 %.

Procedure

Information about the study was given in two different ways. First, participants were informed through leaflets, posters and a website and Facebook page created for this study. Parents who were interested in taking part in a study on child development were invited to contact the research team. Second, the information was given out in schools and to pediatricians. Parents who had concerns about their relationship with their child or about their child's behavior were invited to take part in the research. A total of 137 parents who were interested completed an online questionnaire consisting of socio-demographic information and a child behavioral checklist. All participants completed an informed consent with regard to their participation in the videotaped observation procedure. A subsample of 87 participating parents was also asked to report their childrearing behavior. Observation data were collected from the sample of 137 families during one home visit by extensively trained researchers in the case of 24 families or in a laboratory at the university in the case of the remaining 113. A standardized observation of the parent-child interaction was conducted and video-recorded in order to rate how and to what extent the child and parent were behaviorally adjusted when they played and solved problems together. For 70 dyads, the observation was conducted twice at 8-week intervals in order to evaluate test-retest fidelity.

Measures

Parent and child behavior was observed using the Crowell Procedure (Crowell et al. 1988). This initially involved a series of eight episodes including a 10 min free play session, a 5 min clean-up, a bubble blowing episode, four increasingly difficult problem-solving tasks, and a 2 min separation-reunion episode. In this study, five of the eight episodes were taken into account. The bubble blowing and the separation-reunion were eliminated because of their lack of relevance to the study and the problems raised by their implementation at home. Three rather than four increasingly difficult problem-solving tasks were also presented to children. Our objective was to use the most relevant tasks, i.e. those increasing the likelihood of externalized behaviors, and to make sure that the procedure did not exceed half an hour, so that it could be easily used in both research and clinical settings. The episodes, task type, duration and instructions are presented in Table 1. The first free play episode included a standard set of toys which is described in Table 1. At the end of this first episode, a buzzer sounded to instruct the child to tidy up. Each toy had to be put away in its corresponding box. After 5 min, the experimenter came into the room, even if the clean-up was not finished, and proposed three puzzles. These were selected to move from slightly below to significantly above the child's developmental level, inducing positive and negative emotions and increasing the child's need to rely on the caregiver for help. The first puzzle was easy for all children (six pieces in 2D). The second puzzle was achievable with the adult's support (six easily identifiable pieces in 3D) and the third was hard to perform even for adults (ten difficult to identify pieces in 3D). The following instructions were given to the parents prior to the observation. "We have various activities for you to do together. First you will play freely for 10 min with this set of toys. Try to be as natural as possible and play as if you were at home. Then, when you hear a buzzer, your child must put away the toys. Every toy must be put in the right box. You have 5 min to put everything away. Then I will present three puzzles of increasing difficulty. You can help your child if you think it is necessary. Act normally with him/her. After 10 min, a new buzzer will sound, meaning that the game is over".

In the current study, we used the coding system adapted by Heller et al. (1999). Parent behaviors were scored on a seven-point Likert scale for emotional responsiveness (creating a positive emotional context through encouragement and praise), behavioral responsiveness (providing instrumental support adapted to the child's developmental level through well-timed cues), positive affect (smiling and laughing), irritability (frustration with the child), withdrawal/indifference (disinterest in the child) and aggression towards the child. Child behaviors were scored on a seven-

Table 1 The episodes, task type, duration, instructions and materials of the Crowell procedure

Tasks	Type	Duration	Instructions	Materials
Free-play	Unstructured	10 min	*You will initially play freely for 10 min with this set of toys. Try to be as natural as possible and play as if you were at home*	Standard set of toys including animals, furniture, males, females, children, accessories, cars
Clean-up	Structured	5 min	*When you hear a buzzer, it means that you must put away toys. For storage, you must put every toy in the box that corresponds to it. You have 5 min to store everything*	A specific box divided into 7 sections with tags (animals, furniture, males, females, children, accessories, large objects)
3 puzzles	Structured	10 min	*I present you three puzzles of progressive difficulty. You can quite sure help your child do if you think necessary. Do as you normally do with him. After 10 min a new ringtone will sound and will mean that the game session is over*	3 puzzles of progressive difficulty

point Likert scale for positive affect (smiling and laughing), withdrawal/indifference (disinterest in the parent, depression or sadness), irritability/anger (fussing, pouting, punitive behavior directed towards the parent), non-compliance, aggression towards the parent, persistence (orientation and focus on tasks), and enthusiasm for the task. Each scale was coded for each task. A mean score was computed. Coding was done by three independent trained coders, two of whom were certified by Tulane University (USA), with an intercoder reliability of .92 calculated with the weighted Kappa coefficient on 25 % of the sample.

Parenting behaviors were also reported by parents with the Preschool Parent Form of the Evaluation of Parental Practices (EPEP-PPSF; Meunier and Roskam 2009). The EPEP-PPSF is a 40-item instrument yielding nine factors: Positive Parenting, Monitoring, Rules, Discipline, Inconsistent Discipline, Harsh Punishment, Ignoring, Material Rewarding, and Autonomy. A five-point Likert-type scale is provided for each item, ranging from "never" to "always". This instrument has been validated on 565 French-speaking mothers and fathers and shows good psychometric properties (nine-factor solution explaining 61.36 % of the variance, α ranging from .59 to .90). Confirmatory factor analyses in the validation study showed that two second-order factors covering the supportive and controlling dimensions of parenting emerged from the initial factor solution (CFI = 0.94, RMR = 0.03, and RMSEA = 0.05). These two second-order factors were used in the current study in order to limit the number of constructs under consideration. The supportive factor was composed of Positive Parenting, Autonomy, and Rules, and included items such as "When my child seems to have a problem, I discuss with him/her what is wrong". The controlling factor was composed of Discipline, Harsh Punishment, and Ignoring, and included items such as "When my child does something that is not allowed, I only talk to him/her again when he/she behaves better".

Child Behavior the preschool version of the Child Behavior Check List (Achenbach and Rescorla 2004) was used as another assessment of child behavior and was completed by the participating parent. For the external validation of the Crowell Procedure, we particularly focused on 32 items consisting of the externalizing behavior (EB) scale, encompassing attention problems and aggressive behavior, and the anxious-depressed items of the internalizing behavior scale. CBCL provides three-point Likert scales: "not at all present", "moderately present", or "often present". Scores are computed in each scale by summing item scores. The psychometric properties of the initial version of the scale were good, with α of .92 for "externalizing problems" and of .89 for "internalizing problems" (Achenbach and Rescorla 2000b). Similar psychometric properties have been reported for the French version.

According to the norms of the second-order EB scale, 40.1 % of the children in our sample were in the normal range of EB (<21), 18.2 % were in the borderline clinical range (21–24), and 41.6 % were in the clinical range (>24).

Results

Inter-Task Relations

In order to test the relations between the tasks, t-tests, inter-task correlations and reliability analyses with Cronbach's alphas (α) were computed for each parent and child scale. For child scales, they ranged from $r = .54$ to .73 for positive affect ($\alpha = .89$); from $r = .42$ to .87 for withdrawal ($\alpha = .88$); from $r = .35$ to .77 for irritability ($\alpha = .83$); from $r = .44$ to .82 for non-compliance ($\alpha = .85$); from $r = .31$ to .73 for persistence ($\alpha = .81$); and from $r = .38$ to .75 for enthusiasm ($\alpha = .84$). However, coefficients were low for the aggression scale ($\alpha = .34$). A limited number of occurrences were found in several tasks leading to inconsistent results and $\alpha = .34$. In view of this result, the child aggression scale was excluded from further analyses. For the parent scales, coefficients ranged from $r = .48$ to .79 for behavioral responsiveness ($\alpha = .86$); from $r = .39$ to .72 for emotional responsiveness ($\alpha = .84$); from $r = .49$ to .82 for positive affect ($\alpha = .90$); from $r = .24$ to .93 for withdrawal ($\alpha = .82$); and from $r = .23$ to .76 for irritability ($\alpha = .82$). However, coefficients were low for the aggression scale ($\alpha = .63$). A limited number of occurrences were found in several tasks leading to inconsistent results. In view of this result, the parent aggression scale was excluded from further analyses.

The results of correlations and t-tests between free play, clean-up and Puzzle 3 are presented in Table 2. Change was observed in child and parent behavior throughout the procedure as the level of stress increased. For parent scales, responsiveness and positive affect were seen to decrease from free play to clean-up and Puzzle 3, whereas irritability increased. The same was observed for child scales, with positive affect, persistence and enthusiasm decreasing, and non-compliance and irritability increasing.

Inter-Scale Consistency

As the high correlations found for inter-task consistency suggested that a mean score encompassing all tasks could be computed in each scale, correlations between mean scores were computed. They are displayed in Table 3 for the child scales and in Table 4 for the parent scales. They coherently show moderate to high positive relations among the three positive child scales, i.e. positive affect, persistence and enthusiasm, and the three negative ones, i.e. withdrawal, irritability and non-compliance. Moderate to high negative relations are also displayed between positive and negative scales. The same conclusions were drawn for the parent scales. As expected, the reliability analyses showed that the child scales were consistent, with $\alpha = .88$ (including reverse scores for irritability, non-compliance and withdrawal). The same was reported for parent scales, with $\alpha = .80$ (including reverse scores for irritability and withdrawal).

Table 2 Descriptive statistics, correlations and *t*-test between scales between free-play, tidy-up and puzzle 3 tasks

	Free play (FP)		Tidy-up (TU)		Puzzle 3 (P3)		Correlations FP-TU	*t*-test FP-TU	Correlations FP-P3	*t*-test FP-P3
	M	SD	*M*	SD	*M*	SD		*t(135)*		*t(135)*
Parent' scales										
1. Behavioral responsiveness	5.26	.81	4.99	1.05	4.96	1.25	.55	2.78**	.47	2.52*
2. Emotional responsiveness	5.03	.91	4.76	1.01	4.88	1.07	.56	2.82**	.46	1.42
3. Positive affect	5.47	.79	5.21	.94	5.20	.94	.54	3.51***	.49	3.45***
4. Withdrawal	1.10	.44	1.07	.39	1.15	.69	.66	1.00	.25	−.84
5. Irritability	1.22	.53	1.44	.94	1.41	.88	.28	−2.59*	.41	−2.65**
Child' scales										
1. Positive affect	5.36	.88	4.90	1.25	4.89	1.04	.61	5.23***	.57	5.98***
2. Withdrawal	1.14	.56	1.15	.51	1.11	.41	.48	−.31	.53	.70
3. Irritability	1.42	.79	1.65	1.19	1.76	1.20	.47	−2.43*	.40	−3.42***
4. Non-compliance	1.59	.96	2.05	1.50	1.84	1.50	.44	−3.87***	.54	−2.24*
5. Persistence	5.71	.94	4.80	1.37	4.89	1.47	.31	7.39***	.41	6.76***
6. Enthusiasm	5.61	.91	4.49	1.29	4.62	1.29	.38	10.15***	.38	8.79***

*$p < .05$; **$p < .01$; ***$p < .001$

Table 3 Correlations between child scales (mean scores)

	2	3	4	5	6
1. Positive affect	−.49***	−.63***	−.47***	.56***	.69***
2. Withdrawal	–	.08	.14	−.35***	−.39***
3. Irritability		–	.68***	−.53***	−.56***
4. Non compliance			–	−.80***	−.76***
5. Persistence				–	.89***
6. Enthusiasm					–

***p < .001

Table 4 Correlations between parent scales (mean scores)

	2	3	4	5
1. Behavioral responsiveness	.65***	.43***	−.36***	−.37***
2. Emotional responsiveness	–	.76***	−.23**	−.48***
3. Positive affect		–	−.25**	−.53***
4. Withdrawal			–	.04
5. Irritability				–

**p < .01; ***p < .001

Association between Parent and Child Behavior

The correlations between child and parent scales (mean scores) are presented in Table 5. Coherently, high correlations were found for common scales, i.e. child and parent positive affect (r = .78), withdrawal (r = .57) and irritability (r = .57). The moderate to high correlations found between child and parent scales confirm the ability of the Crowell Procedure to capture the bidirectional relationships in the dyad. Child and parent behaviors are seen to be interdependent. The more positive the child's behavior towards the parent, the more positive the parent's behavior is and *vice versa*.

Discriminant Properties of the Crowell Procedure

T-test comparisons were computed in order to test whether the child and parent scales varied according to the level of child externalizing behavior as well as between mothers and fathers. The mean scores of each scale were compared between mothers (N = 115) and fathers (N = 22), with a t-test for independent samples. However, this analysis remains exploratory due to the limited number of fathers in the current study. No significant difference was displayed between mothers and fathers for the five parent scales. In contrast, significant differences were found between children according to their level of externalizing behaviors evaluated with the CBCL externalizing scale. For the purposes of analysis, children (N = 19) who were in the borderline clinical range (between 21 and 24 on the CBCL

externalizing scale) were excluded. Only normally developing children (<21) and children with a clinical level of externalizing behavior (>24) were compared. Preliminary analyses demonstrated that these two groups (normally developing versus clinical level of externalizing behavior) were comparable with regard to socio-demographic data. Five out of the six child scales discriminated between children displaying a clinical level of externalizing behavior from those in the normal range. Descriptive statistics and results of t-tests are presented in Table 6. Children displaying a clinical level of externalizing behavior showed less positive affect, persistence and enthusiasm but higher irritability and non-compliance than normally developing children.

Significant differences were also found for two of the five parent scales. Descriptive statistics and the results of t-tests are presented in Table 7. Parents of children displaying a clinical level of externalizing behavior displayed less positive affect but higher irritability towards their child than parents of normally-developing children.

External Validation

For the child scales, external validation was tested with the CBCL externalizing and the CBCL internalizing behavior scales. For the parent scales, it was tested with the support and control dimensions of the EPEP-PPSF scale. Coefficients are presented in Table 8 for the child scales and in Table 9 for the parent scales.

Overall, moderate correlations support the validity of the Crowell Procedure child scales. Higher externalizing behavior was associated with lower positive affect, persistence and enthusiasm, and higher irritability and non-compliance. Coherently, withdrawal that relates to internalizing behavior was not associated with the externalizing behavior scale of the CBCL.

Overall, the correlations support the validity of the Crowell Procedure parent scales. Controlling behaviors were associated with low behavioral and emotional responsiveness as well as with low positive affect and high irritability. Coherently, parent withdrawal was not associated with controlling behavior. Support was not

Table 5 Correlations between child and parent scales (mean scores)

	Parent behavioral responsiveness	Parent emotional responsiveness	Parent positive affect	Parent withdrawal	Parent irritability
1. Child positive affect	.42***	.62***	.78***	−.23**	−.50***
2. Child withdrawal	−.36***	−.42***	−.44***	.57***	.12
3. Child irritability	−.35***	−.41***	−.49***	.00	.57***
4. Child non-compliance	−.39***	−.44***	−.39***	.01	.57***
5. Child persistence	.35***	.48***	.48***	−.13	−.47***
6. Child enthusiasm	.39***	.56***	.58***	−.18	−.47***

**$p < .01$; ***$p < .001$

Table 6 Means and standard deviations of child scales for children in the normal vs. clinical range of the CBCL-externalizing behavior scale and t-tests

	Normally-developing children $N = 55$		Children with clinical level of externalizing behavior $N = 63$		t-test
	M	SD	M	SD	$t(116)$
1. Positive affect	5.38	.74	4.92	.78	3.19**
2. Withdrawal	1.12	.29	1.06	.19	1.10
3. Irritability	1.28	.50	1.77	.97	−3.37***
4. Non-compliance	1.38	.61	2.05	1.30	−3.43***
5. Persistence	5.56	.76	5.02	1.08	3.03**
6. Enthusiasm	5.31	.81	4.78	.94	3.23**

**$p < .01$; ***$p < .001$

Table 7 Means and standard deviations of parent scales for parents of children in the normal vs. clinical range of the CBCL-externalizing behavior scale and t-tests

	Parents of normally developing children $N = 55$		Parents of children with clinical level of externalizing behavior $N = 63$		t-test
	M	SD	M	SD	$t(116)$
1. Behavioral responsiveness	5.08	.97	5.13	.86	−.26
2. Emotional responsiveness	5.18	.76	5.02	.76	1.16
3. Positive affect	5.51	.67	5.19	.71	2.45*
4. Withdrawal	1.14	.48	1.03	.10	1.77
5. Irritability	1.19	.37	1.47	.75	−2.51*

*$p < .05$

significantly associated with parent scales, whereas the coefficients were in the right direction, i.e. positive support with responsiveness and positive affect but negative support with withdrawal and irritability.

Test-Retest Reliability

Test-retest reliability was studied in a subsample of 70 dyads after an 8-week interval. For the child scales, the correlations were moderate, ranging from $r = .32$, $p < .01$, for the enthusiasm scale, to $r = .57$, $p < .001$, for the persistence scale. Only the correlation for the withdrawal scale was found to be low, with $r = .12$. For the parent scales, correlations were moderate, ranging from $r = .28$, $p < .05$, for the withdrawal scale, to $r = .45$, $p < .01$, for behavioral responsiveness.

Discussion

The aim of the current study was to analyze the psychometric properties of an adapted version of the Crowell Procedure among preschoolers interacting with their mother or father. Overall, the results support the consistency of both tasks and scales, and the discriminant properties,

Table 8 Correlations between child scales and the CBCL-externalizing behavior scale (N = 137)

	Positive affect	Withdrawal	Irritability	Non compliance	Persistence	Enthusiasm
Externalizing behavior	−.34***	−.02	.33***	.31***	−.28***	−.32***
Internalizing behavior	−.04	.17†	.03	.00	−.07	.04

†p < .10; ***p < .001

Table 9 Correlations between parent scales and EPEP support and control scales (N = 87)

	Support	Control
1. Behavioral responsiveness	.12	−.20†
2. Emotional responsiveness	.02	−.25*
3. Positive affect	.11	−.30**
4. Withdrawal	−.13	.07
5. Irritability	−.10	.33**

†p < .10; *p < .05; **p < .01

external validity and fidelity of the coding system. The Crowell Procedure can therefore be used as an observational paradigm to assess both child and parent behavior. However, we failed to validate the aggression scales. A first explanation could be that in the observational context, parent aggression could be inhibited. Adult participants knew they were being video-recorded and observed by experts in parent-child relationships. In such a context, social desirability could play an important role in lowering inappropriate behavior towards the child. Also, child aggression could be inhibited due to the unfamiliar context and the presence of an unknown experimenter. This unfamiliarity could also lower the probability of extreme behavior occurring in children. A second explanation could relate to the specific nature of this scale and hence the nature of observed aggressive behaviors. Unlike the other child and parent behaviors that were measured during the procedure, aggression was assessed with a frequency scale, indicating whether this kind of behavior was present or absent in the parent-child interaction. To report aggressive behavior, some coders therefore expected to observe clear violence against the partner. Measuring aggression on a continuum may be difficult in this setting. A better coding system would be to record the presence of specific aggressive behaviors such as throwing or pulling a puzzle piece or speaking aggressively to the child or to the parent, even if these behaviors are low in intensity.

The inter-task comparisons suggested that a mean score could be computed for each scale through the entire procedure. This was the case for children's positive affect, withdrawal, irritability, non-compliance, persistence and enthusiasm as well as for parents' behavioral and emotional responsiveness, positive affect, withdrawal and irritability.

Inter-task consistency is important, since a certain coherence in the child's behavior in interaction with his/her parent is expected. However, whereas high consistency may suggest that each task is representative to some extent of the entire procedure, clinicians and researchers must be aware that using single tasks may lead to a reduction in the variety of situations where relevant behaviors can be observed. It may increase the probability of misinterpretation of child's and parent's behaviors. The use of single tasks may also lead to the loss of the transitions between the tasks, which was designed to elicit stress in the interaction, allowing problem-solving and stress reduction capabilities to be observed (Sprang and Craig 2014). The use of different tasks is therefore recommended to gather information on representative parent-child interactions and on the way the dyad handles stressful situations, for example when the child does not want to perform a task. Single tasks should be used only in particular contexts, for example when time constraints arise in clinical settings or when rapid screening for preventive purposes is needed. In these specific cases and in line with the suggestion of Miron et al. (2009), the most relevant task might then be free play, for three reasons. First, it seems to be comparable to what happens in the family setting due to its relatively unstructured nature. In the absence of any achievement expectations, parents and children are free to display their interactional routines. Second, in contrast to teaching tasks, free play can be easily implemented with diverse populations including low-educated parents or mentally disabled children. Third, it has been argued that free play serves as the most revealing observational window into children's emotional lives, revealing the familiarity of the dyad with play and their use of fun-oriented time (Miron et al. 2009).

In line with previous psychometric analysis (Sprang and Craig 2014), scores emerging from the different child and parent scales could be summarized in two globally consistent scores. Our results suggest that the six child scales assess a second-order construct that we could label "child behavior". The same is true for the parent scales. These results give support to the conceptual validity of the coding system by displaying moderate to high correlations in the expected directions between the scales. However, the validity of the second-order scales needs to be validated with larger sample sizes and appropriate statistical procedures such as confirmatory factor analysis. These second-order scales could

for example be used in studies where the number of variables has to be limited for the purposes of analysis or in studies where no hypothesis is made regarding specific first-order behaviors. However, first-order scales remain necessary and relevant for studies focusing on specific outcomes such as externalizing problems in children (Roskam et al. 2015). They are also useful for studies documenting the effects of intervention, by helping to identify which specific behavior improves (Loop and Roskam 2016).

The correlations between the child and parent scales illustrate the interdependence that has been theoretically described in the transactional model (Sameroff 2009) and empirically shown in bidirectional effect studies (Barnett et al. 2012; Pardini et al. 2008; Verhoeven et al. 2010). These correlations suggest that the Crowell Procedure may be relevant to capturing dyadic aspects of the parent-child relationship. Also, correlations between parent and child scales assume that behaviors are limited to this specific interactional context. Child behavior has to be understood and interpreted according to parent behavior and *vice versa*. This may explain the moderate correlations (around $r = .30$) with the CBCL. They illustrate that child behavior is only partly the same in various real-life situations as in a particular parent-child interaction. This is why, in the absence of a gold-standard measure of child behavior, strong arguments have been put forward for considering a multi-informant and multi-method assessment and diagnostic procedure (Kraemer et al. 2003; Noordhof et al. 2008; Roskam et al. 2013).

For the purposes of diagnosis, the discriminant analyses revealed that, as expected in the initial study of Crowell et al. (1988), the procedure was effective at differentiating children displaying a clinical level of externalizing behavior from normally developing ones. Differences were significant for five out of six child scales. Only the withdrawal scale was not significant, probably due to the internalizing orientation of this scale. Correlations found between the CBCL internalizing scale and child scales were consistent with this view. Also, two out of the five parent scales displayed good discriminant properties, i.e. positive affect and irritability. Whereas children's externalizing problems seem to preserve parents' responsiveness, they result in less fun-oriented interaction, in particular lower positive affect and higher irritability from the parent. Such emotional reactions illustrate that caring for externalized children is often described as more challenging and less rewarding than caring for other children, leading to lower levels of satisfaction, negative feelings, and higher criticism in caregivers (Coleman and Karraker 2003; Meunier et al. 2011; Slagt et al. 2012). Finally, as was the case for the child scales, withdrawal was not discriminant.

To conclude regarding the external validation of the parent scales, the analyses demonstrated low to moderate correlations with controlling childrearing behavior.

However, no association was displayed with supportive behavior. Supportive behavior in fact refers to positive parenting, autonomy demands and rules, the last two of which are not strictly coded in the Crowell Procedure. By contrast, controlling behaviors such as discipline and ignoring are appraised to some extent through the responsiveness and withdrawal parent scales.

Finally, the reliability of the procedure was found to be moderate for all child and parent scales, but low for child withdrawal. Compared to questionnaire-based assessment, observations are more sensitive to transient conditions related to children's and parents' mood state, tiredness or willingness, as well as to co-occurring events resulting in lower moment-to-moment reliability. An additional important point is the delay between the two waves. It may be that eight weeks is too long to ensure the reliability of the coding system because of dyadic dynamics. New research should be conducted with a shorter interval to provide additional support for the reliability of the procedure.

In sum, in line with previous findings (Crowell et al. 1988; Sprang and Craig 2014), the current study supports the validity of the Crowell Procedure and the coding system adapted from Heller's work (Heller et al. 1999). Several advantages can be highlighted. First, this procedure is highly flexible. It can be implemented in diverse settings such as the laboratory, as in Mouton and Roskam (2014) and Loop and Roskam (2016), health services or at home (Roskam et al. 2015). It can also be used on several different occasions in order to appraise the stability of the constructs over time or to evaluate the effect of interventions for children and their parents, as in Roskam et al. (2015) and Loop and Roskam (2016). The materials such as the free-play set of toys or the teaching task puzzles could also easily be adapted to avoid learning and test-retest effects. Finally, the procedure may also be implemented with different caregivers in foster care (Zeanah et al. 2003), allowing the testing of the core behavioral functioning of children and the extent to which behaviors vary according to the interactional partner. No differences between mothers and fathers appeared in the current study, but this analysis remains exploratory due to the underrepresentation of fathers (only 16 % of the sample). Future research should study the differences between dyads in greater depth. A second advantage is that thanks to its development and validation, the Crowell Procedure is not time-consuming, which makes it useful in both clinical and research settings. In particular, free play could be used in a self-standing manner. Third, the Crowell Procedure is not only a valid assessment tool, but also a therapeutic tool that can, for example, be used for video feedback purposes (Velderman et al. 2006). Also, playing with the child is already a therapeutic step for some dyads (Wettig et al. 2011).

Despite these strong advantages, the Crowell Procedure nevertheless has some drawbacks. Every video-recorded observational paradigm elicits a certain level of stress in participants. The ecological validity of such observation may therefore be questioned. The structured tasks, i.e. clean-up and puzzles, were designed to elicit frustration and negative emotions in both children and their parent, but the possibility cannot be excluded that such negative feelings are also present during free play among participants, especially for dyads for whom play interaction is unusual. Also, some child and parent behaviors may have been inhibited due to the specific context of observation or to the presence of an unknown experimenter. Moreover, although the procedure was highly standardized the possibility cannot be excluded that the location where the dyads were observed, i.e. at home or in the laboratory, slightly influenced the results. Finally, this validation is neither representative nor relevant for all parent-child dyads, as observation paradigms like the Crowell Procedure are very dependent on the age and the characteristics of children. For instance, the procedure may be less relevant for older children who are expected to be more autonomous both in task achievement and emotion regulation. Future validation studies are also needed to replicate the findings in diverse cultural settings and with children displaying internalizing problems, for example.

Compliance with Ethical Standards

References

Achenbach, T. M., & Rescorla, L. A. (2000a). Manual for the ASEBA preschool forms & profiles: An integrated system of multi-informant assessment; child behavior checklist for ages 1 1/2-5; language development survey; caregiver-teacher report form. University of Vermont.

Achenbach, T. M., & Rescorla, L. A. (2000b). *Mental health practitioner's guide for the achenbach system of empirically based assessment (ASEBA)*. Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T. M., & Rescorla, L. A. (2004). *The achenbach system of empirically based assessment (ASEBA) for ages 1.5 to 18 years the use of psychological testing for treatment planning and outcomes assessment: Volume 2: Instruments for children and adolescents.* 3rd edn. Mahwah, NJ: Lawrence Erlbaum Associates Publishers 179–213.

Aoki, Y., Zeanah, C. H., Heller, S. S., & Bakshi, S. (2002). Parent-infant relationship global assessment scale: A study of its predictive validity. *Psychiatry and Clinical Neurosciences, 56*, 493–497.

Barnett, M. A., Gustafsson, H., Deng, M., Mills-Koonce, W. R., & Cox, M. (2012). Bidirectional associations among sensitive parenting, language development, and social competence. *Infant and Child Development, 21*, 374–393.

Belsky, J. (1984). The determinants of parenting: A process model. *Child Development, 55*, 83–96.

Brassart, E., & Schelstraete, M. -A. (2015a). Enhancing the communication abilities of preschoolers at risk for behavior problems: Effectiveness of a parent-implemented language intervention. *Infants & Young Children, 28*, 337–354.

Brassart, E., & Schelstraete, M. -A. (2015b). Simplifying parental language or increasing verbal responsiveness, what is the most efficient way to enhance preschoolers' verbal interactions? *Journal of Education and Training Studies, 3*, 133–145.

Budd, K. S., Poindexter, L. M., Felix, E. D., & Naik-Polan, A. T. (2001). Clinical assessment of parents in child protection cases: An empirical analysis. *Law and Human Behavior, 25*, 93.

Cassidy, J. (2008). The nature of the child's ties. In J. Cassidy, & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications.* 2nd edn. (pp. 3–22). New York, NY: Guilford Press.

Coleman, P. K., & Karraker, K. H. (2003). Maternal self-efficacy beliefs, competence in parenting, and toddlers' behavior and developmental status. *Infant Mental Health Journal, 24*, 126–148.

Conway, A., Mcdonough, S. C., Mackenzie, M., Miller, A., Dayton, C., Rosenblum, K., et al. (2014). Maternal sensitivity and latency to positive emotion following challenge: Pathways through effortful control. *Infant Mental Health Journal, 35*, 274–284.

Criss, M. M., Shaw, D. S., & Ingoldsby, E. M. (2003). Mother–son positive synchrony in middle childhood: Relation to antisocial behavior. *Social Development, 12*, 379–400.

Crowell, J., & Feldman, S. (1988). Mothers' internal models of relationships and children's behavioral and developmental status: A study of mother-child interaction. *Child Development, 59*, 1273–1285.

Crowell, J., Feldman, S., & Ginsberg, N. (1988). Assessment of mother-child interaction in preschoolers with behavior problems. *Journal of the American Academy of Child & Adolescent Psychiatry, 27*, 303–311.

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendtations for further study. *Psychological Bulletin, 131*, 483–509.

Fitzgerald, M. M., Shipman, K. L., Jackson, J. L., McMahon, R. J., & Hanley, H. M. (2005). Perceptions of parenting versus parent-child interactions among incest survivors. *Child Abuse & Neglect, 29*, 661–681.

Goodman, R. (2001). Psychmetric properties of the strenghts and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*, 1337–1345.

Heller, S., Aoki, Y., & Sheffner, X. (1999). *Revision of the Crowell parent-child relationship scale.* New Orleans: Tulane University Medical Center.

Hynan, D. J. (2003). Parent-child observations in custody evaluations. *Family Court Review, 41*, 214–223.

Kraemer, H. C., Measelle, J. R., Ablow, J. C., Essex, M. J., Boyce, W. T., & Kupfer, D. J. (2003). A new approach to integrating data from multiple informants in psychiatric assessment and research: Mixing and matching contexts and perspectives. *The American Journal of Psychiatry, 160*, 1566–1577.

Lieberman, A. F. (2004). Traumatic stress and quality of attachment: Reality and internalization in disorders of infant mental health. *Infant Mental Health Journal, 25*, 336–351.

Lieberman, A. F., Silverman, R., & Pawl, J. H. (2000). Infant-parent psychotherapy: Core concepts and current approaches. *Handbook of Infant Mental Health, 2*, 472–484.

Loop, L., & Roskam, I. (2016). Do children behave better when parents are reinforced in emotion coaching? A micro-trial research. *Journal of Child and Family Studies, 25*, 2223–2235.

Luby, J. L., Sullivan, J., Belden, A., Stalets, M., Blankenship, S., & Spitznagel, E. (2006). An observational analysis of behavior in

depressed preschoolers: Further validation of early-onset depression. *Journal of the American Academy of Child and Adolescent Psychiatry*, 45, 203–212.

Malik, N. M., Lederman, C. S., Crowson, M. M., & Osofsky, J. (2002). Evaluating maltreated infants, toddlers, and preschoolers in dependency court. *Infant Mental Health Journal*, 23, 576–592.

Meunier, J. C., & Roskam, I. (2009). Validation of the preschool and primary school form of a questionnaire assessing parents' child-rearing behavior. *Journal of Clinical Child and Adolescent Psychology*, 38, 166–175.

Meunier, J. C., Roskam, I., & Browne, D. T. (2011). Relations between parenting and child's behavior: Exploring child's personality and parental self-efficacy as third variables. *International Journal of Behavioral Development*, 35, 246–259.

Miller, A. L., McDonough, S. C., Rosenblum, K. L., & Sameroff, A. (2002). Emotion regulation in context: Situational effects on infant and caregiver behavior. *Infancy*, 3, 403–433.

Miron, D., Lewis, M., & Zeanah, C. H. (2009). Clinical use of observational procedures in early childhood relationship assessment. In C. H. Zeanah (Ed.), *Handbook of infant mental health*. New York: Guilford Press.

Misri, S., Reebye, P., Kendrick, K., Carter, D., Ryan, D., Grunau, R. E., & Oberlander, T. F. (2006). Internalizing behaviors in 4-year-old children exposed in utero to psychotropic medications. *American Journal of Psychiatry*, 163, 1026–1032.

Mouton, B., & Roskam, I. (2014). Confident mothers, easier children: A quasi-experimental manipulation of mothers' self-efficacy. *Journal of Child and Family Studies*, 24, 1–11.

Noordhof, A., Oldehinkel, A. J., Verhulst, F. C., & Ormel, J. (2008). Optimal use of multi-informant data on co-occurrence of internalizing and externalizing problems: The TRAILS study. *International Journal of Methods in Psychiatric Research*, 17, 174–183.

Osofsky, J., Bosquet, M., Kronenberg, M., & Hammer, J. (2003). *Revision of parent-child relationship coding manual*. New Orleans: Louisiana State University Health Sciences Center.

Osofsky, J., Kronenberg, M., Hammer, J., Lederman, J. C., Katz, L., Adams, S., & Hogan, A. (2007). The development and evaluation of the intervention model for the Florida Infant Mental Health Pilot Program. *Infant Mental Health Journal*, 28, 259–280.

Pardini, D., Fite, P., & Burke, J. (2008). Bidirectional associations between parenting practices and conduct problems in boys from childhood to adolescence: The moderating effect of age and African-American ethnicity. *Journal of Abnormal Child Psychology*, 36, 647–662.

Robinson, L. R., Morris, A. S., Heller, S., Scheeringa, M. S., Boris, N. W., & Smyke, A. T. (2009). Relations between emotion regulation, parenting, and psychopathology in young maltreated children in out of home care. *Journal of Child and Family Studies*, 18, 421–434.

Roskam, I., Brassart, E., Loop, L., Mouton, B., & Schelstraete, M. -A. (2015). Stimulating parents' self-efficacy beliefs or verbal responsiveness: Which is the best way to decrease children's externalizing behaviors? *Behavior Research and Therapy*, 72, 38–48.

Roskam, I., Meunier, J. C., & Stievenart, M. (2013). The comparison and combination of multi-informant and multi-method data on preschoolers' externalizing behaviour. *The International Journal of Educational and Psychological Assessment*, 14, 79–93.

Roskam, I., Meunier, J.C., & Stievenart, M. (2015). From parents to siblings and peers. *The Wonderful Story of Social Development*, 5, 1–13

Sameroff, A. (2009). *The transactional model of development: How children and contexts shape each other*. Washington, DC: American Psychological Association.

Sameroff, A., & Fiese, B. H. (2000). Transactional regulation: The developmental ecology of early intervention. *Handbook of Early Childhood Intervention*, 2, 135–159.

Scheeringa, M. S., Zeanah, C. H., Myers, L., & Putnam, F. (2004). Heart period and variability findings in preschool children with posttraumatic stress symptoms. *Biological Psychiatry*, 55, 685–691.

Schmidt, F., Cuttress, L. J., Lang, J., Lewandowski, M. J., & Rawana, J. S. (2007). Assessing the parent–child relationship in parenting capacity evaluations: Clinical applications of attachment research. *Family Court Review*, 45, 247–259.

Slagt, M., Deković, M., de Haan, A. D., van den Akker, A. L., & Prinzie, P. (2012). Longitudinal associations between mothers' and fathers' sense of competence and children's externalizing problems: The mediating role of parenting. *Developmental Psychology*, 48, 1554–1562.

Smyke, A.T. (2000). Effects of maternal maltreating status on maternal representation and mother-child interaction. Unpublished doctoral dissertation. University of New Orleans.

Sprang, G., Clark, J., Kaak, O., & Brenzel, A. (2004). Developing and tailoring mental health technologies for child welfare: The Comprehensive Assessment and Training Services (CATS) Project. *American Journal of Orthopsychiatry*, 74, 325.

Sprang, G., & Craig, C. (2014). Crowell problem solving procedure: A psychometric analysis of a laboratory measure of the caregiver-child relationship. *Child and Adolescent Mental Health*, 20, 202–209

Sroufe, L. A. (1989). Relationships, self and individual adaptation. In A. J. Sameroff & R. N. Emde (Eds.), *Relationship disturbances in early childhood* (pp. 70–94). New York: Basic Books.

Velderman, M. K., Bakermans-Kranenburg, M. J., Juffer, F., Van Ijzendoorn, M. H., Mangelsdorf, S. C., & Zevalkink, J. (2006). Preventing preschool externalizing behavior problems through video-feedback intervention in infancy. *Infant Mental Health Journal*, 27, 466–493.

Verhoeven, M., Junger, M., van Aken, C., Deković, M., & van Aken, M. A. G. (2010). Parenting and children's externalizing behavior: Bidirectionality during toddlerhood. *Journal of Applied Developmental Psychology*, 31, 93–105.

Verschueren, K., Dossche, D., Marcoen, A., Mahieu, S., & Bakermans-Kranenburg, M. (2006). Attachment representations and discipline in mothers of young school children: An observation study. *Social Development*, 15, 659–675.

Wettig, H. H., Coleman, A., & Geider, F. J. (2011). Evaluating the effectiveness of theraplay in treating shy, socially withdrawn children. *International Journal of Play Therapy*, 20, 26.

Zeanah, C. H., Larrieu, J. A., Heller, S., & Valliere, J. (2000). Infant-parent relationship assessment. In J. Zeanah (Ed.), *Handbook of infant mental health*. 2nd edn. (pp. 222–235). New York: Guilford Press.

Zeanah, C. H., Nelson, C. A., Fox, N. A., Smyke, A. T., Marshall, P., Parker, S. W., & Koga, S. (2003). Designing research to study the effects of institutionalization on brain and behavioral development: The Bucharest Early Intervention Project. *Development and Psychopathology*, 15, 885–907.