

## Are We Talking About the Same Child?

### Parent-Teacher Ratings of Preschoolers' Social-Emotional Behaviors

Sofia O. Major and Maria J. Seabra-Santos

Faculty of Psychology and Educational Sciences, University of Coimbra, Portugal

Roy P. Martin

Department of Educational Psychology, University of Georgia, USA

#### Author's Note

The present paper is based on data of the first author's PhD dissertation, supported by a Doctoral grant (SFRH/BD/29141/2006) from the Fundação para a Ciência e a Tecnologia (FCT), presented to the Faculty of Psychology and Educational Sciences, University of Coimbra (Portugal).

Correspondence concerning this article should be addressed to Sofia Major, Faculdade de Psicologia e de Ciências da Educação, Rua do Colégio Novo, Apartado 6153, 3001-802 Coimbra, Portugal.

E-mail: [smajor@fpce.uc.pt](mailto:smajor@fpce.uc.pt)

Phone: +351 239851450

Fax: +351 239851465

## Are We Talking About the Same Child?

### Parent-Teacher Ratings of Preschoolers' Social-Emotional Behaviors

Most of the available information regarding children's problem behaviors comes from informants' perceptions, assessed with rating scales or interviews (Hay et al., 1999; Merrell, 2008). In the field of social-emotional assessment, behavior rating scales are the most widely used assessment technique for young children (Merrell, 2008; Myers, 2013). This popularity is based on the flexibility (the same issues can be assessed in different settings) and cost-effectiveness of this assessment tool, as well as decades of research supporting reliability and validity. With preschoolers, the use of behavior rating scales is even more important due to the impossibility of collecting valid self-reports from these younger children (Berg-Nielsen, Solheim, Belsky, & Wichstrom, 2012). However, an issue of concern in both research applications and practice applications regarding rating scales is the consistent finding of modest inter-rater agreement (Rescorla et al., 2012).

The literature available on inter-rater agreement from home and school settings can be summarized in four general findings. First, agreement coefficients between ratings of different informants are weak to moderate, with lower levels of agreement among informants who have different roles in the child's life (e.g., parents/teachers) (Achenbach, McConaughy, & Howell, 1987; Culp, Howell, Culp, & Blankemeyer, 2001; Gagnon, Nagle, & Nickerson, 2007; Grietens et al., 2004; Gross, Fogg, Garvey, & Julion, 2004; Hinshaw, Han, Erhardt, & Huber, 1992; Merrell, 2002; Touliatos & Lindholm, 1981; Winsler & Wallace, 2002). Second, parents (especially mothers) tend to rate their children as having more problem behaviors than teachers (Berg-Nielsen et al., 2012; Culp et al., 2001; Gross et al., 2004; Strickland, Hopkins, & Keenan, 2012; Touliatos & Lindholm, 1981; Winsler & Wallace, 2002). Third, parents and teachers have a tendency to agree more concerning externalizing problem behaviors (more visible) than for internalizing ones, primarily because they are less

easily observed (Achenbach et al., 1987; Gagnon et al., 2007; Hinshaw et al., 1992; Merrell, 2002; Winsler & Wallace, 2002). Finally, there is a higher informant agreement for problem behaviors than for social skills (Winsler & Wallace, 2002).

Low inter-rater agreement indicates that ratings on problem behaviors may be substantially different depending on whether the informant is the father, the mother, the teacher, or someone else (Culp et al., 2001; De Los Reyes & Kazdin, 2005). Nevertheless, the idea that different informants (e.g., parents, teachers, professionals, the child and peers) should be used in the assessment of children is well established (e.g., Achenbach, et al., 1987; Grietens et al., 2004; Kamphaus & Frick, 1996; Treutler & Epkins, 2003; Winsler & Wallace, 2002). The recommendation to collect information from different perspectives is based on the knowledge that differences in child (e.g., gender, age), parent (e.g., educational level, psychopathology), and family characteristics (e.g., number of siblings, family status) (De Los Reyes & Kazdin, 2005), as well as the setting in which the behavior is observed (e.g., home and school) may affect the ratings that are made (De Los Reyes & Kazdin, 2005; Gagnon et al., 2007; Hay et al., 1999; Merrell, 2008; Strickland et al., 2012). As far as the child and informant variables are concerned, inconsistent findings have been the rule rather than the exception, suggesting the need for further investigation (De Los Reyes & Kazdin, 2005).

Unfortunately, most of the studies that assess parent-teacher agreement are in reference to school age children, leaving this issue poorly clarified for preschoolers (Grietens et al., 2004; Gross et al., 2004; Rescorla et al., 2012; Strickland et al., 2012; Winsler & Wallace, 2002). The present study aims to give a contribution to this field by exploring parent-teacher agreement for a large representative sample of preschoolers.

Why does informant agreement matter? Low parent-teacher agreement, if not well understood, casts a shadow of doubt on the reliability and validity of an assessment based on responses from any one source. It is well known that ratings of behavior by caregivers, such

as parents and teachers are influenced by individual differences between sources caused by reasons such as stress, personality differences. Such findings may lead to the exaggerated belief that ratings of parents and teachers are so contaminated by these factors that they say more about the raters than the child.

This is a simplistic understanding of source disagreement. Disagreements between informants may be an outcome of observing the child in different settings (e.g., parents and teachers), differences between informants in the manner in which they interact with the child (e.g., mothers and fathers may elicit different behaviors from the child), or perceptions of the same behavior may differ due to values differences among raters (e.g., parents may value assertive behavior on the part of the child, while some teacher may feel this behavior is disrespectful) (Renk, 2005). In each of these cases, ratings may be perfectly reliable and valid, although different. Understanding the factors that affect rater perception makes it clear that there is no single gold standard for the source of ratings.

Based on this understanding, many authors (e.g., Renk, 2005) recommend the use of several informants from different settings to overcome rater differences. However, this raises other issues. One issue is whether to aggregate the ratings from different sources. Further, if aggregation is seen as desirable, how should it be done? Should raters in the home be aggregated (e.g., fathers and mothers) and compared to ratings in the schools (e.g., two teachers), or should all ratings be aggregated (parents and teachers). Achenbach et al. (2008) have pointed out some of the limitations of aggregation of ratings. They argue that informant disagreement is as instructive as agreement, and cannot be assumed as an error. Such differences may justify the need to have distinct intervention strategies at home and in the school or with different individuals in either setting (Achenbach et al., 1987; Culp et al., 2001).

Despite individual differences that affect perceptions, parents have some advantages over other sources of information on child behavior since they spend more time with the child than anyone else, and have observed the behavior of the child across time and across situations (Culp et al., 2001). Moreover, they are often in a good position to offer unique information about the child's life history, present difficulties, and external factors affecting the behavior (e.g., a death in the family).

Although teachers are well prepared to be informants about the child's academic competencies (Mashburn, Hamre, Downer, & Pianta, 2006), their perceptions about social-emotional functioning has not always been considered relevant in the past (Kamphaus & Frick, 1996). Nevertheless, the growing number of children attending preschools and the amount of time spent there draws attention to the importance of considering teacher and aid-teacher perspectives on the assessment of social-emotional behavior of children in this age group (Hinshaw et al., 1992; Rescorla et al., 2012). One of the reasons contributing to the increased interest in the teachers' perspective has to do with their experience, practice and opportunities to observe the child across time and social situations (e.g., classroom, playground), offering them access to information about the child that may be unavailable to parents (Kamphaus & Frick, 1996). Second, unlike parents, teachers have the clear advantage of being able to observe a large sample of children (the class) with whom they can contrast the child's behavior (Grietens et al., 2004; Kamphaus & Frick, 1996; Mashburn et al., 2006; Strickland et al., 2012). Finally, they can observe the child for longer periods in interaction with peers, which gives them access to a unique set of data regarding the child and his/her social behaviors (Mashburn et al., 2006). For preschoolers, teacher perceptions are particularly important in the identification/validation of problem behaviors, since for this younger age group the diversity of noncompliant and aggressive behaviors is quite broad.

Typically, inter-rater agreement is assessed by correlating the ratings of two types of raters (teachers and parents) across a group of children. This method determines the extent to which the rank ordering of individuals in the group by two types of raters is the same. A different way of addressing this issue is the analysis of differences in scores between different sources (De Los Reyes & Kazdin, 2005). This approach to the analysis of source discrepancies focuses on mean difference between types of raters (De Los Reyes & Kazdin, 2005; Treutler & Epkins, 2003), information that the typical correlation analysis cannot detect. These two methods of comparing informant ratings often produce substantially different outcomes (De Los Reyes & Kazdin, 2005; Treutler & Epkins, 2003).

During the 20<sup>th</sup> century, several behavior rating scales have been translated to other languages with evidence of high reliability and validity for different countries. A major issue in the use of behavior rating scales with preschoolers from different language/culture groups is how translation affects the psychometric properties of the instrument, including inter-rater agreement. This topic becomes relevant due to the globalization of psychology and the lack of instruments available for non-English speaking countries (van-Windenfelt, Treffers, Beurs, Siebelink, & Koudijs, 2005). Most of the world must rely on translations and adaptations of a well-established assessment tool from the English language. However, several challenges in the use of translated instruments have been discussed in the literature (van de Vijver & Hambleton, 1996; van-Widenfelt et al., 2005), once this process is more than just a literal translation of the items. The present study exemplifies that kind of situation for the evaluation of Portuguese preschoolers' social-emotional behavior.

Several behavior rating scales have been translated to other languages with good psychometric properties. The two most common in social-emotional and behavioral assessment are the Achenbach System of Empirically Based Assessment (ASEBA) forms (translated into over 90 languages) and the Strengths and Difficulties Questionnaire (SDQ;

Goodman, 1997; with about 80 SDQ translations available from the web site). The translation of the same instrument in several different languages also allowed for outstanding examples of research on the informant agreement in a cross-cultural context. For example, Rescorla et al. (2012) analyzed data for 7,380 preschoolers from 13 societies using the ASEBA preschool forms. The authors found a modest mean cross-informant agreement and that despite distinct language, religion, or cultural values, a similar cross-informant pattern was found across societies; the mean of Total Problems score was significantly higher for parents than for caregivers/teachers.

The purpose of the present study is to examine parent-teacher agreement on ratings of preschoolers' social skills and problem behaviors and to provide additional evidence of reliability to the Portuguese version of the Preschool and Kindergarten Behavior Scales – Second Edition (PKBS-2; Merrell, 2002). We predict that the level of agreement between home and school settings will be low to moderate. Further, we hypothesize that parents and teachers will be more in agreement for externalizing than for internalizing problem behaviors. In addition, discrepancies between parents' and teachers' ratings will also be studied, with the expectation that parents will rate their children as having more problem behaviors than teachers. In order to explore the impact of several variables on informant agreement, analyses will be replicated for child (gender) and parent (mothers' educational level) characteristics.

## **METHOD**

### *Participants*

The Portuguese normative sample of the PKBS-2 was used in the present study. The sample consisted of 1,000 preschoolers rated both by parents and teachers. A random sampling plan was used stratified by child gender ( $n = 500$  for boys and for girls) and age ( $n = 250$  for 3, 4, 5 and 6 years old), with a mean age of 4.50 years ( $SD = 1.12$ ). In addition, the sample was stratified for type of school (58, 16 and 26% for public, private, and social

assistance, respectively), school location (by regions of Portugal) and community size (i.e., urban, moderately urban and rural), with children randomly selected within these strata. Stratification was based on data available from the Portuguese National Institute of Statistics (INE) and the Office of Statistics and Planning for Education (GEPE).

*Parent sample.* Parents' ratings were provided mostly by mothers (83%), with 12% coming from fathers, and the remaining 5% completed by parents jointly or by another informant such as grandparents. Information was available regarding the educational level for 795 of the 834 responding mothers: 19% had completed less than the national compulsory 9 years (basic school), 29% had completed basic school, 30% had received a high school diploma, and 22% had a higher level of education.

*Teacher sample.* Preschool teachers provided ratings of 875 children (87%) and primary school teachers provided ratings on 125 (13%) children (at age 6, children may attend preschool or the first year of primary school depending on their birth month). Teachers rated an average of eight children each (range 5-15). Of the 131 teachers who provided ratings, 129 were female. Professional experience ranged from 1 to 35 years, with most teachers having more than 5 years of professional experience.

### *Instrumentation*

The PKBS-2 (Merrell, 2002) is a North American behavior rating scale designed to assess social skills and problem behaviors for preschool children. It consists of 76 items, divided into two major scales: Social Skills (34 items) and Problem Behaviors (42 items). Subscales were derived based on exploratory and confirmatory analyses with three Social Skills subscales (Social Cooperation, Social Interaction and Social Independence) and two broadband Problem Behaviors subscales (Externalizing and Internalizing Problems). Second order factor analyses led to the development of five supplemental subscales: three for Externalizing Problems (Self Centered/Explosive, Attention Problems/Overactive and



Antisocial/Aggressive) and two of Internalizing Problems (Social Withdrawal and Anxiety/Somatic Problems). The same set of items is used for parents, teachers or other raters from home and school settings, with items rated on a four-point Likert scale (from 0 “*never*” to 3 “*frequently*”). The psychometric properties of the PKBS-2 are documented in the manual, which presents several studies for reliability (e.g. internal consistency, temporal stability) and validity (e.g., construct, convergent and divergent validity) (Merrell, 2002). Despite some limitations (e.g., overrepresentation of boys in the standardization sample), Allin (2004) recognized the PKBS-2 as a user friendly test and a useful tool for preschoolers’ social, behavioral, and emotional assessment, based on sound theoretical and psychometric evidence.

All the efforts were made in order to keep the wording of the PKBS-2 items as similar as possible from the original, with some adaptation to the Portuguese language, so that the meaning of culture-specific items was not lost by literal translation (van-Widenfelt et al., 2005). The process of translation/adaptation to the Portuguese language and back-translation was completed following permission granted by the PKBS-2 author and editor. Item analyses, as well as evidence of reliability and validity were documented for the Portuguese version (Major, 2011; Major & Seabra-Santos, 2014). In terms of reliability, Cronbach alpha coefficients for the Social Skills scale and subscales ranged from .76 to .88 for the parents’ sample and .89 to .95 for the teachers’ sample. For the Problem Behaviors scale, alpha coefficients ranged from .78 to .95 for the parents’ sample and .85 to .97 for the teachers’ sample. These values are very similar to those obtained from the original North American version. Based on the same exploratory and confirmatory procedures used for the PKBS-2 (factor solution tested for the total sample and for home and school samples independently), the factor structure of the American version was replicated, with minor adjustments for some subscales’ designation, in order to make them more clearly understandable in Portuguese (e.g., Cooperation/Social Adjustment instead of just Social Cooperation), or according to

items' content (e.g., Over Activity/Lack of Attention instead of Attention

Problems/Overactive, once the majority of the items included in that subscale are related to activity problems and not to attention ones). Thus, three Social Skills subscales

(Cooperation/Social Adjustment, Social Interaction/Empathy and Social

Independence/Assertiveness) and two Problem Behaviors subscales (Externalizing Problems and Internalizing Problems) were found. Consistent with findings from the original PKBS-2

factor structure, second order factor analyses led to the development of five supplemental subscales (Major, 2011).

### *Procedure*

Once authorization was obtained from the National Data Protection Commission, principals from several schools from all over the country (randomly selected) were approached. After the principals permission was obtained, teachers from 64 schools were contacted by the researcher. Children were randomly selected based on the classroom list and recruited depending on parental consent as well as the following criteria: age between 3-6 years old (ages for which the PKBS-2 were specifically developed), one Portuguese speaking parent (who could understand and rate the items), no motor impairment (the child may have a distinct approach to peers due to such condition) and attendance of the recruited school for at least 3 months (this is a request to rate the PKBS-2, so that the teacher has had a minimal time of interaction with the child). Parent and teacher participation was voluntary. All questionnaires were collected in collaboration with teachers, who served as mediators in the communication between the researcher and parents. A predetermined number of cases were collected from each school. Questionnaires were given to teachers and a packet containing the parent consent form, a document explaining the goal of the research and a questionnaire was sent home. Teachers were only allowed to complete their questionnaires after parents returned

the packet and agreed to participate. The return rate was 88% for parents and 100% for teachers; that is, after obtaining parental consent, none of the teachers refused to participate.

### *Data Analyses*

Inter-class (Pearson  $r$ s) and intra-class ( $ICC$ , for inter-rater reliability) correlation coefficients were calculated to determine parent-teacher agreement for each of the PKBS-2 scales and subscales. Besides bivariate correlations (Pearson coefficient), the  $ICC$  has been used as a measure of rater reliability for quantitative data (Chen & Barnhart, 2013) as it represents a useful index for cross-informant agreement (e.g., Langberg et al., 2010). Furthermore, for a more general approach, the mean correlation was also calculated according to Fisher's  $r$ -to- $z$  transformation. In order to analyze the discrepancies between parent and teacher ratings, mean scores were examined and differences were tested via paired-sample  $t$  tests. This approach was used once it provides a distinct perspective on the informant agreement from correlation analyses, while comparing scales and subscales scores for parents and teachers. Effect size statistics were provided by the Eta Squared statistic, classified according to Cohen's (1988) guidelines (small = .01-.05, moderate = .06-.13 and large  $\geq$  .14). Lastly, analyses were replicated for PKBS-2 total scores (Social Skills and Problem Behaviors scales) based on the child's gender, and three levels of mothers' education (only for mother ratings).

## **RESULTS**

Table 1 displays correlations calculated for pairs of informants (home and school ratings) for each of the PKBS-2 scores. For the total sample, Pearson correlations were low to moderately low (although all significant for  $p < .01$ ), with weaker correlations for the Social Skills scale ( $r = .28$ ) than for the Problem Behaviors scale ( $r = .34$ ). Moreover, correlations were stronger for externalizing problem behaviors than for the internalizing problems ( $r = .42$  vs  $.27$ , respectively). Intra-class correlations, although higher than Pearson's ( $ICC$  range  $.27-$

.56), present the same general pattern. Mean correlation according to Fisher's  $r$ -to- $z$  transformation, also indicates a moderate but significant association between parent and teacher ratings ( $r = .32, p < .01$ ).

#### Insert Table 1

As shown in Table 1, when considering child's gender and mothers' educational level, the same pattern of results is found: low to moderate correlations and higher agreement for externalizing problem behaviors. Although correlations are quite similar for boys and girls, agreement tends to be higher for boys with respect to social skills (Cooperation/Social Adjustment and Social Interaction/Empathy) as well as for some externalizing problems (Anti-Social/Aggressive and Over-Activity/Lack of Attention). When taking into account the 795 mothers for whom educational data were available, there is a trend towards a higher agreement between parent-teacher when mothers had completed a higher level of education (see Table 1).

When analyzing mean differences in the total sample (see Table 2), all the differences between parent and teacher ratings are statistically significant. Parents systematically rated their children in a more positive direction with regard to social skills than teachers, but in a more negative way for problem behaviors. The only exceptions are for the Cooperation/Social Adjustment subscale and for the Social Withdrawal supplemental subscale, with lower ratings by parents than teachers. Although some of the effect sizes are low (.01 for four PKBS-2 scores), many differences are sizeable (between .08 and .36), with the largest effect occurring for the Opposition/Explosive supplemental subscale.

#### Insert Table 2

When conducting the same analyses (only for total scores) for child's gender and mothers' education, the same pattern of results occurred for the total score of the Problem Behaviors scale. That is, children are rated by parents more negatively than by teachers, with

all differences being statistically significant and of moderate to large effect sizes ( $\eta^2 = .08-.22$ ). However, for the total Social Skills scale it is only for boys and for children with higher educated mothers that the differences between ratings reach a statistical significance with small effect sizes ( $\eta^2 = .03$ ).

## DISCUSSION

Considering parents and teachers as the main informants for preschoolers' social-emotional assessment (Strickland et al., 2012), we analyzed a nationally representative sample of children using two approaches proposed by De Los Reyes and Kazdin (2005): agreement (correlations) and scale-score discrepancies (mean comparisons). Correlation analyses provide evidence of modest parent-teacher agreement, a finding that is consistent with the literature (e.g., Achenbach et al., 1987; Gagnon et al., 2007; Grietens et al., 2004; Gross et al., 2004; Merrell, 2002; Winsler & Wallace, 2002). Parent-teacher correlations did not change substantially when analyzed separately for different child and parent variables. Compared with the results from the classical meta-analysis of Achenbach et al. (1987), the mean correlation of .32 obtained for the parent-teacher agreement in this study was quite similar to the mean coefficient of .28 obtained by Achenbach and colleagues across multiple settings.

Several reasons may explain why parents and teachers have such low correspondence in their rates of social-emotional functioning. One reason is that parents and teachers view behavior in different settings with distinct contextual demands (academic vs. family) (Strickland et al., 2012). The preschool classroom (group setting) is a social environment containing a number of same-age peers, while home environment (family setting) seldom includes a number of same-age peers. Second, because of its group structure and demand characteristics, where learning is the focus even for younger children, the environment in the classroom is quite different from that experienced by children at home. Third, children's interaction with teachers may differ from their interaction with their parents due to several

factors, including the personality or general interaction style of the caregiver. Finally, parents and teachers may have different degrees of tolerance for some behaviors. This may occur, in part because parents may be less familiar with behaviors that are considered developmentally appropriate (Strickland et al., 2012). Thus, these kinds of differences may create divergent perspectives concerning the child's behavior.

Results also replicate the literature review (Achenbach et al., 1987; Gagnon et al., 2007; Hinshaw et al., 1992; Winsler & Wallace, 2002) indicating that correlations for parent-teacher agreement are higher for externalizing problem behaviors, than for internalizing behaviors. This difference in the level of agreement, also observed for the Externalizing and Internalizing supplemental subscales, may be explained by the high visibility and annoyance caused by problems such as over-activity or aggressiveness, compared with behaviors related to social withdrawal or anxiety. These results are very similar to those obtained for the original PKBS-2 studies (Merrell, 2002). Furthermore, consistent with the findings of Touliatos and Lindholm (1981) who found higher agreement coefficients for the boys, our results indicate that boys may elicit higher levels of inter-rater agreement on externalizing problem behaviors. As noted by Berg-Nielsen et al. (2012), parents and teachers may be more aware of boys' misbehaviors while they are more tolerant with girls, not only because it is expected that boys have more problem behaviors than girls, but also more externalizing ones, such as aggression. On the other hand, the higher parent-teacher agreement found for mothers with a higher educational level may be due to a better knowledge of child development and more adjusted expectations that are more similar to those of professionals.

When mean differences were examined, parents assign higher mean ratings on all the Social Skills scales (except for Cooperation/Social Adjustment, rated highly by teachers) and Problem Behaviors scales (except for Social Withdrawal). Thus, when compared to teachers, parents over-rated both the positive and the negative attributes of their children's behavior.

While analyzing the differences by disaggregating the sample according to the children's gender and the maternal level of education, the same pattern of parent ratings that had been seen in the total sample was observed, but only for the Problem Behaviors scale. For the Social Skills scale, modest effects were found for boys (rated by teachers with lower levels of social skills) and for children of higher educated mothers (rated by these mothers with higher levels of social skills).

Overall, these results are consistent with the general literature on parent-teacher agreement which indicates that parents rate their children as exhibiting more problematic behaviors than informants from the school setting (e.g., Culp et al., 2001; Gross et al., 2004; Rescorla et al., 2012; Winsler & Wallace, 2002). Several possible explanations for this effect can be discussed. On the one hand, parents spend more time with their children, have more opportunities to observe them, and therefore may have a better knowledge of what is their typical behavior. On the other hand, it is possible that many parents have inappropriate expectations about normative behavior of the preschool child, tending to evaluate as disruptive behaviors that are normative. However, social withdrawal behaviors were rated as more problematic by teachers than parents, possibly because of the social circumstances of the classroom, more outgoing behavior are viewed by teachers as optimizing opportunities for learning and social integration.

The present study, while providing evidence of the reliability of the Portuguese version of the PKBS-2, highlights some issues related to the use of translated behavior rating scales. Consistent with current thinking about cultural-specificity (van de Vijver & Hambleton, 1996; van-Widenfelt et al., 2005), the Portuguese version is not just a literal translation of the PKBS-2. Attention was given to the meaning of the items by experts in child psychological assessment and the subscales labels were adapted to become more understandable in the Portuguese context. The fact that results are congruent with literature

and with the original version of the PKBS-2 highlights the usefulness of this tool in cross-cultural research. The Portuguese version, therefore can be used to compare results to the English version.

All these findings have significant clinical implications for the assessment of preschool children. Although parents and teachers are the most requested informants for the social-emotional assessment of younger children, some level of disagreement is to be expected. Their different perspectives should be considered useful and are likely to provide valuable information not only about the child, but also about the adults themselves (Hay et al., 1999). Thus, while integrating information in clinical practice, and in order to gain a broader understanding of the child's social-emotional behaviors, the professional has to take into account the unique perspective of each one of the informants, while also considering the characteristics of the child and the settings in which behavior is observed.

The present study also documents inter-rater agreement of the social-emotional assessment of preschoolers using the Portuguese version of the PKBS-2. Further, the study contributes to the limited literature on inter-rater agreement for preschool children (Myers, 2013). Despite the strengths of this study, including a large national community sample of preschoolers, one limitation is evident: the current home sample was mostly composed by mothers' ratings. This could be an important limitation since it has been noted that mothers usually rate their children with more problem behaviors than fathers. Nevertheless, the lower participation of fathers in studies of this nature is common (Merrell, 2002; Treutler & Epkins, 2003).

Some future directions should be considered. First, the perspective of fathers should be more systematically integrated. Also, further research is needed to examine the impact of child, parent and teacher characteristics (e.g., teachers' experience) on inter-rater agreement. This question might be addressed using a more complex approach, such as hierarchical



modeling analysis, to explore the contributions of several specific variables on child's positive and negative behaviors. It could prove useful to have parents and teachers observe the same behaviors in a laboratory situation. This might illuminate differences in perception when viewing the same behaviors.

More than 20 years after the meta-analysis presented by Achenbach et al. (1987), the idea of understanding the perspective of different informants is still a pertinent issue in preschoolers' assessment (Berg-Nielsen et al., 2012). The present paper suggests that, even if parents and teachers may not view the child's behaviors exactly the same way, they are still talking about the same child and they should be recognized as essential informants to understand preschoolers' social-emotional behaviors.

## REFERENCES

- Achenbach, T. M., Becker, A., Döpfner, M., Heiervang, E., Roessner, V., Steinhausen, H., & Rothenberger, A. (2008). Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: Research findings, applications, and future directions. *Journal of Child Psychology and Psychiatry*, *49*, 251-275. doi:10.1111/j.1469-7610.2007.01867.x
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213-232.
- Allin, J. D. (2004). Review of the Preschool and Kindergarten Behavior Scales, 2nd Edition. *Journal of Psychoeducational Assessment*, *22*, 81-86. doi: 10.1177/073428290402200108
- Berg-Nielsen, T. S., Solheim, E., Belsky, J., & Wichstrom, L. (2012). Preschoolers' psychosocial problems: In the eyes of the beholder? Adding teacher characteristics as determinants of discrepant parent-teacher reports. *Child Psychiatry and Human Development*, *43*, 393-413. doi: 10.1007/s10578-011-0271-0
- Chen, C-C., & Barnhart, H. X. (2013). Assessing agreement with intraclass correlation coefficient and concordance correlation coefficient for data with repeated measures. *Computational Statistics & Data Analysis*, *60*, 132-145. doi: 10.1016/j.csda.2012.11.004
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Culp, R. E., Howell, C. S., Culp, A. M., & Blankemeyer, M. (2001). Maltreated children's emotional and behavioral problems: Do teachers and parents see the same things? *Journal of Child and Family Studies*, *10*, 39-50. doi: 10.1023/A:1016631111859

- De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin, 131*, 483-509.
- Gagnon, S. G., Nagle, R. J., & Nickerson, A. B. (2007). Parent and teacher ratings of peer interactive play and social-emotional development of preschool children at risk. *Journal of Early Intervention, 29*, 228-242. doi:10.1177/105381510702900303
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*, 581-586. doi: 10.1111/j.1469-7610.1997.tb01545.x
- Grietens, H., Onghena, P., Prinzie, P., Gadeyne, E., Van Assche, V., Ghesquière, P., & Hellinckx, W. (2004). Comparison of mothers', fathers', and teachers' reports on problem behavior in 5- to 6-year-old children. *Journal of Psychopathology and Behavioral Assessment, 26*, 137-146. doi:10.1023/B:JOBA.0000013661.14995.59
- Gross, D., Fogg, L., Garvey, C., & Julion, W. (2004). Behavior problems in young children: An analysis of cross-informant agreements and disagreements. *Research in Nursing & Health, 27*, 413-425. doi:10.1002/nur.20040
- Hay, D. F., Pawlby, S., Sharp, D., Schmücker, G., Mills, A., Allen, H., & Kumar, R. (1999). Parents' judgements about young children's problems: Why mothers and fathers might disagree yet still predict later outcomes. *Journal of Child Psychology and Psychiatry, 40*, 1249-1258. doi: 10.1111/1469-7610.00541
- Hinshaw, S. P., Han, S. S., Erhardt, D., & Huber, A. (1992). Internalizing and externalizing behavior problems in preschool children: Correspondence among parent and teacher ratings and behavior observations. *Journal of Clinical Child Psychology, 21*, 143-150. doi: 10.1207/s15374424jccp2102\_6

- Kamphaus, R. W., & Frick, P. J. (1996). *Clinical assessment of child and adolescent personality and behavior*. Boston: Allyn and Bacon.
- Langberg, J. M., Epstein, J. N., Simon, J. O., Loren, R. E., Arnold, L. E., Hechtman, L., ... Wigal, T. (2010). Parent agreement on ratings of children's Attention Deficit/Hyperactivity Disorder and broadband externalizing behaviors. *Journal of Emotional and Behavioral Disorders, 18*, 41-50. doi:10.1177/1063426608330792
- Major, S. O. (2011). *Avaliação de aptidões sociais e problemas de comportamento em idade pré-escolar: Retrato das crianças portuguesas [Assessment of social skills and problem behaviors in preschoolers: A picture of Portuguese children]* (Doctoral dissertation). Retrieved from [https://estudogeral.sib.uc.pt/bitstream/10316/17774/5/Tese\\_Sofia%20Major.pdf](https://estudogeral.sib.uc.pt/bitstream/10316/17774/5/Tese_Sofia%20Major.pdf)
- Major, S. O., & Seabra-Santos, M. J. (2014). *Preschool and Kindergarten Behavior Scales – Second Edition (PKBS-2): Adaptação e estudos psicométricos da versão portuguesa. [Preschool and Kindergarten Behavior Scales – Second Edition (PKBS-2): Adaptation and psychometric studies of the Portuguese version]*. *Psicologia: Reflexão e Crítica, 27*, 599-609. doi: 10.1590/1678-7153.201427409
- Mashburn, A. J., Hamre, B. K., Downer, J. T., & Pianta, R. C. (2006). Teacher and classroom characteristics associated with teachers' ratings of prekindergartners' relationships and behaviors. *Journal of Psychoeducational Assessment, 24*, 367-380. doi: 10.1177/0734282906290594
- Merrell, K. W. (2002). *Preschool and Kindergarten Behavior Scales – Second Edition*. Austin, TX: PRO-ED.
- Merrell, K. W. (2008). *Behavioral, social, and emotional assessment of children and adolescents* (3rd ed.). New York: Erlbaum.

- Myers, C. L. (2013). Comparing results from the clinical assessment of behavior and child behavior checklist with referred preschoolers. *Psychology in the Schools, 50*, 1-12. doi: 10.1002/pits.21657
- Renk, K. (2005). Cross-informant ratings of the behavior of children and adolescents: The “Gold Standard”. *Journal of Child and Family Studies, 14*, 457-468. doi:10.1007/s10826-005-7182-2
- Rescorla, L. A., Achenbach, T. M., Ivanova, M. Y., Bilenberg, N., Bjarnadottir, G., Denner, S., ... Verhulst, F. C. (2012). Behavioral/emotional problems of preschoolers: Caregiver/teacher reports from 15 societies. *Journal of Emotional and Behavioral Disorders, 20*(10), 1-14. doi: 10.1177/1063426611434158
- Strickland, J., Hopkins, J., & Keenan, K. (2012). Mother-teacher agreement on preschoolers' symptoms of ODD and CD: Does context matter? *Journal of Abnormal Child Psychology, 40*, 933-943. doi: 10.1007/s10802-012-9622-y
- Touliatos, J., & Lindholm, B. W. (1981). Congruence of parents' and teachers' ratings of children's behavior problems. *Journal of Abnormal Child Psychology, 9*, 347-354. doi:10.1007/BF00916839
- Treutler, C. M., & Epkins, C. C. (2003). Are discrepancies among child, mother, and father reports on children's behavior related to parents' psychological symptoms and aspects of parent-child relationships? *Journal of Abnormal Child Psychology, 31*, 13-27. doi: 10.1023/A:1021765114434
- van de Vijver, F., & Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist, 1*, 89-99. doi: 10.1027/1016-9040.1.2.89
- van Widenfelt, B. M., Treffers, P. D. A., Beurs, E., Siebelink, B. M., & Koudijs, E. (2005). Translation and cross-cultural adaptation of assessment instruments used in

psychological research with children and families. *Clinical Child and Family*

*Psychology Review*, 8, 135-147. doi: 10.1007/s10567-005-4752-1

Winsler, A., & Wallace, G. L. (2002). Behavior problems and social skills in preschool children: Parent-teacher agreement and relations with classroom observations. *Early Education & Development*, 13, 41-58. doi:10.1207/s15566935eed1301\_3

Table 1

*Agreement between Parents and Teachers for PKBS-2 Scores (Total Sample, Child's Gender and Mothers' Educational Level)*

PKBS-2 Score	Total Sample		Child's Gender <sup>a</sup>		Mothers' Educational Level <sup>a</sup>		
	Pearson	Intra-class			Basic	High	Higher
	Correlation (N = 1,000)	Correlation (N = 1,000)	Boys (n = 500)	Girls (n = 500)	School (n = 381)	School (n = 238)	Education (n = 176)
<b>Social Skills Scale</b>							
Cooperation/Social Adjustment	.31**	.45**	.33**	.27**	.31**	.28**	.39**
Social Interaction/Empathy	.19**	.27**	.22**	.16**	.15**	.15*	.28**
Social Independence/Assertiveness	.32**	.46**	.32**	.32**	.29**	.34**	.33**
Total Social Skills	.28**	.41**	.29**	.25**	.27**	.22**	.30**
<b>Problem Behaviors Scale</b>							
Externalizing Problems	.42**	.53**	.42**	.38**	.37**	.45**	.49**
Internalizing Problems	.27**	.40**	.26**	.28**	.23**	.19**	.41**
Total Problem Behaviors	.34**	.46**	.34**	.33**	.29**	.35**	.45**
<b>Problem Behaviors Supplemental Subscales</b>							
Anti-Social/Aggressive	.41**	.56**	.41**	.35**	.36**	.46**	.46**
Over-Activity/Lack of Attention	.44**	.55**	.44**	.38**	.41**	.42**	.51**
Opposition/Explosive	.31**	.36**	.31**	.31**	.25**	.34**	.38**
Social Withdrawal	.25**	.39**	.24**	.25**	.19**	.24**	.33**
Anxiety/Somatic Complaints	.25**	.36**	.23**	.26**	.24**	.12	.37**

a = Pearson Correlations

\*  $p < .05$     \*\*  $p < .01$  (two-tailed)

Table 2

*Differences between Parent and Teacher Ratings (Total Sample, Child's Gender and Mothers' Educational Level)*

PKBS-2 Score	Parents		Teachers		<i>t</i>	<i>p</i>	$\eta^2$
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Total Sample ( <i>N</i> = 1,000)							
Social Skills Scale							
Cooperation/Social Adjustment	25.79	3.80	27.29	4.86	-9.25**	.001	.08
Social Interaction/Empathy	25.62	3.15	23.65	5.15	11.37**	.001	.11
Social Independence/Assertiveness	33.33	3.73	32.90	5.31	2.50*	.013	.01
Total Social Skills	84.74	8.82	83.84	13.33	2.06*	.039	.01
Problem Behaviors Scale							
Externalizing Problems	31.51	13.96	23.39	18.53	14.32**	.001	.17
Internalizing Problems	15.55	7.25	13.39	8.80	6.96**	.001	.05
Total Problem Behaviors	47.06	19.24	36.78	24.34	12.84**	.001	.14
Problem Behaviors Supplemental Subscales							
Anti-Social/Aggressive	7.86	5.59	7.33	7.57	2.27*	.024	.01
Over-Activity/Lack of Attention	10.51	4.66	7.99	5.88	13.97**	.001	.16
Opposition/Explosive	13.14	5.21	8.07	6.26	23.69**	.001	.36
Social Withdrawal	4.89	3.56	5.27	4.40	-2.41*	.016	.01
Anxiety/Somatic Complaints	10.66	4.50	8.12	5.07	13.60**	.001	.16
Subsamples by Child's Gender							
Boys ( <i>n</i> = 500)							
Total Social Skills	84.36	8.92	81.74	13.98	4.12**	.001	.03
Total Problem Behaviors	48.88	19.56	41.03	25.71	6.63**	.001	.08
Girls ( <i>n</i> = 500)							
Total Social Skills	85.11	8.71	85.93	12.30	-1.40	.162	.00
Total Problem Behaviors	45.23	18.75	32.54	22.12	11.92**	.001	.22
Subsamples by Mothers' Educational Level							
Basic School ( <i>n</i> = 381)							
Total Social Skills	83.33	9.06	82.21	13.99	1.52	.130	.00
Total Problem Behaviors	50.22	19.06	39.59	24.73	7.84**	.001	.14
High School ( <i>n</i> = 238)							
Total Social Skills	85.35	8.35	85.63	12.16	-0.33	.740	.00
Total Problem Behaviors	46.20	19.01	36.95	25.68	5.49**	.001	.11
Higher Education ( <i>n</i> = 176)							
Total Social Skills	87.32	8.15	85.11	13.13	2.22*	.028	.03
Total Problem Behaviors	41.44	17.86	31.68	20.87	6.31**	.001	.19

\*  $p < .05$     \*\*  $p < .01$