Running head: Source-specific information on social cognition

Source-specific information on social cognition: A matter of context or concept?

Núria de la Osa[1,2]

Eva Penelo[1,3]

J. Blas Navarro[1,3]

Esther Trepat[1,2]

Lourdes Ezpeleta[1,2]

[1]Unitat d'Epidemiologia i de Diagnòstic en Psicopatologia del Desenvolupament

[2]Departament de Psicologia Clínica i de la Salut

[3]Departament de Psicobiologia i Metodologia de les Ciències de la Salut

Universitat Autònoma de Barcelona

The authors declare that they have no conflict of interest.

Acknowledgements

We would like to thank the participating schools and families.

Funding was from the Spanish Ministry of Science, Innovation and Universities (MICINN/FEDER) [grant PGC2018-095239-B-I00].

Mailing address:

Nuria de la Osa – ORCID 0000-0003-4499-0942

Departament de Psicologia Clínica i de la Salut. Edifici B

Universitat Autònoma de Barcelona

08193 Bellaterra (Barcelona). SPAIN

Email: nuria.delaosa@uab.cat

Abstract

This work tackles the measurement invariance of the social cognition construct when different observers, age and participant's age are considered. This is a prior question that needs to be answered before attributing discrepancies in information coming from diverse sources just to the varying behavior occurring across setting, and mainly interpret the discrepancies as indicative of cross-contextual variability. The paper also studies the link between discrepancies and source-specific information and the validity of that information to predict several outcomes.

The measurement invariance across sex, time and informant of a social cognition measure applied to children's parents and teachers was longitudinally tested in a Spanish general population sample, at ages 5 (N = 581) and 10 (N = 438). Full or partial metric and scalar equivalence were found across sex and over time within informants. Partial scalar invariance was not obtained across informants. Latent class analysis identified 2 classes of difficulties in social cognition for both informants at both ages: low social cognition and high social cognition.

Comparison of classes resulting predicting outcomes yielded differential predictions due not only to varying context but also to a different concept of social cognition across informants. In general, significant differences between raters were informant dependent.

We conclude that it is important to consider both teachers' and parents' observations to fully understand the construct of social cognition.

Keywords: Measurement invariance, SCDC, social cognition, source-specific.

Source-specific information on social cognition: A matter of context or concept?

Social cognition (SC) is understood as the set of higher cognitive functions that are involved in understanding social situations by attending to, interpreting, and responding to social cues, thus enabling the planning of appropriate responses (Adolphs, 2001; Happé & Frith, 2014). The fifth edition of the Diagnostic and Statistical Manual for Mental Disorders (DSM-5; APA, 2013) has introduced SC as one of the trans diagnostically impaired factors of neurocognitive functioning. Deficits in SC have been reported in many child and adolescent conditions such as conduct problems (Gilmour, Hill, Place, & Skuse, 2004; Hughes, White, Sharpen, & Dunn, 2000), autism, (Skuse, Mandy, & Scourfield, 2005), attention deficit/hyperactivity disorder (ADHD) (Bora & Pantelis, 2016), oppositional defiant disorder (ODD) (Poletti & Adenzano, 2013) and social anxiety (Tye et al., 2014). Such deficits, are predictors of impaired functional outcomes related to creating and maintaining interpersonal relationships (Arioli, Crespi, & Canessa, 2018). Studies on SC have described a continuous distribution of these deficits in the general population (Skuse et al., 2009), suggesting the convenience of measuring them dimensionally to detect subclinical problems associated with functional impairment in several areas and a large range of childhood developmental disorders (Oliver, Barker, Mandy, Skuse, & Maughan, 2011).

Nearly all the work done relating SC and conduct problems has been carried out relying on parents' information as reported by Gilmour et al. (2004) and Oliver et al. (2011) and metanalysis (Trentacosta & Fine, 2010). Due to the varying contexts in which parents and teachers observe children, their perceptions may be different and the convenience of including teachers in the study of SC would appear to make sense. Teachers observe the children in social situations with peers more frequently than parents do, along time and so, can provide unique information about their interactions (Evans, Pederson, Fite, Blossom, & Cooley, 2015; Mayes & Lockridge, 2018; Sonuga-Barke, Auerbach, Campbell, Daley, & Thompson, 2005).

Furthermore, their interaction with many children may mean they are in a better position than parents to judge typical social ability for developmental stage (Mcneilis, Maughan, Goodman, & Rowe, 2017).

It appears to be of interest to provide data on the accuracy of teachers as informants regarding SC and to check whether there are differences between their information and parents' information and, if so, whether this is simply a matter of the varying context the behavior occurs in or whether discrepancies are due to the different "own norms" parents and teachers may use to inform about it. Discrepancies between raters of children's behavioral/emotional problems have been widely reported (De los Reyes & Kazdin, 2005; De Los Reyes, Henry, Tolan, & Wakschlag, 2009; King et al., 2018; Stuart, 2018). The most common explanation for these discrepancies has been varying contexts. However, regarding the measurement of a construct, it must be established whether it may manifest differently across sex, time or informants, or whether the components of a latent factor are equally interpreted across groups and time (Mcneilis et al., 2017).

As far as we are aware, there are no published data on the agreement of teacher and parent information regarding SC and neither there are any studies on the invariance across different time points and informants using a widely used measure of SC, the Social Cognition Disorder Checklist (SCDC) (Skuse et al., 1997), a 1-factor scale designed to be answered by parents. It was originally developed to measure social behavior deficits in individuals with Turner's syndrome and was later also used to measure social and communication deficits in social reciprocity, non-verbal skills and pragmatic language usage, which are a characteristic triad of the autism spectrum (Skuse, Mandy, & Scourfield, 2005). The underlying premise was that autistic traits are widely present in the general population and that scores in the clinical range are also predictive of the presence of conduct problems in community samples (Constantino & Todd, 2003; Spiker, Lotspeich, Dimiceli, Myers, & Risch, 2002). This

instrument was designed for screening purposes and it can be used to understand the role of sub-threshold autistic trait difficulties (such as understanding the feelings, thoughts and behavior of others) in behavioral disorders such as ADHD and disruptive disorders, included ODD (Skuse et al., 2005). It has also been used and shown to be useful to detect behavioral problems in general populations (Mandy et al., 2013; Oliver et al., 2011; de la Osa, Granero, Penelo, & Ezpeleta, 201). Given the relationship between SC and executive function, in addition to the huge changes that occur at the onset of puberty, age could also be a crucial variable to consider when using the instrument (Blakemore, 2010; Kilford, Garrett, & Blakemore, 2016). Any comparison of test scores based on different informants, or over age, or also for example across sex, should be preceded by measurement invariance analyses, to ensure that responses are comparable. Measurement invariance evaluates whether or not measurements of the same attribute are obtained, in this case when using SCDC, and only if it supported could SCDC scores be meaningfully compared across different settings or groups of responses.

Ha formatat: anglès (EUA)

Codi de camp canviat

Thus, the goals of the present study were to (a) test for measurement invariance of SCDC items across sex, time and informant; (b) study whose information regarding that measure, if any, is most valid to predict emotional or conduct problems, functioning and children's self-perception of the quality of their social relationships. We hypothesized that parents and teachers have a different mental concept when they inform about children and adolescents' SC. Considering teachers' better position compared with parents to observe children in social situations, we anticipate their information would be more predictive to identify the aforementioned difficulties and impairment than parents' ratings.

Method

Participants

The data used in this study correspond to the third and eighth follow ups of a

longitudinal research project into vulnerability to behavioral problems in preschool children. The research was conceived with a two-phase design and was based on an initial random sample of 2,283 children (49.0% female) selected from all registered preschoolers (age 3) in Barcelona (Spain) during the 2009-10 academic year.

All 2,283 families received the invitation to answer the parent's version of the Strengths and Difficulties Questionnaire for 3- and 4-year-olds (SDQ^{3-4}; Goodman, 1997). The proportion of families that agreed to participate in this first phase was 58.7% (N = 1,341 families) and no differences were found when comparing participants and refusals according to sex (p = .95). However, the proportion of refusals was statistically higher for families from low socio-economic groups (p < .001). For the second sampling phase, the minimum sample size of the positive screening cohort was calculated with the software nQuery Advisor 7.0 (Statistical Solutions, 2007). A prevalence of 15% and a multiple correlation between covariates of .40 were assumed. The sample size was determined for detecting OR = 1.8 between psychopathology and risk factors, using a test of hypothesis for risk alpha = .05 and power of .80. It was necessary to also include the negative screening group to obtain unbiased estimates of prevalences and incidences. A sample of 30% of the negative screening group was considered sufficient. As the planned follow-up was 12 years long, the sample size was increased 50% for attrition. Also, the research team established that the final sample's size was manageable.

The final second-phase sample included N = 622 children, with no statistical differences between participants and refusals in sex (p = .820), although more families that agreed to participate belonged to higher SES (Hollingshead, 1975) than those that refused (p =.007). Of the participants, 310 were boys (49.8%), 558 were Caucasian (88.9%), and SES distribution was as follows: 35.4% high, 46.3% medium, and 18.3% low. The final sample for this study, obtained when the children mean age and standard deviation were 5.7 (0.36) and

10.6 (0.34) years-old, included 581 and 434 children respectively (those with both parents' and teachers' SCDC). No differences were found in sex or SES socioeconomic status between the initial sample at age 3 and children at age 5 (female: 49.7%; SES: high 31.8%, medium 46.0%, low 22.2%) or at age 10 (female: 50.0%; SES: high 34.7%, medium 48.6%, low 16.7%) ($p \geq .166$). Most of the children were Caucasian (89.8% at age 5; 91.8% at age 10). Children with intellectual disability or pervasive developmental disorders were not included in the study.

The project was approved by the ethics review committee of the authors' institution. The director of the participating schools and the families received a complete description of the study. The families were recruited at the schools and gave written consent. All the parents of children from P3 (aged 3) were invited to answer the SDQ^{3-4} at home and return it to the schools. The families who agreed and met the screening criteria were contacted by telephone and interviewed yearly at the school throughout the entire 8-year period. The interviewers had previously received training and were blind to the children's screening group.

Measures

The *Social Communication Disorders Checklist* (SCDC; Skuse et al., 1997). This is a 12-item questionnaire with situations rated according to whether the corresponding behavior has been seen during the past 6 months and is responded to on a 3-point Likert-type scale (0: *not true*, 1: *quite true*, and 2: *very true*). Nine of the 12 items serve to measure those aspects of the autistic triad that reflect reciprocal social interaction skills (e.g.: *Does not realize when others are upset or angry*). The remaining items measure behavior problems and reflect functional impairment (e.g.: *Behaviour often disrupts family life;* (Skuse, 2005). It was answered by both parents and teachers. The items were completed using the Spanish validation (de la Osa et al., 2014). Higher scores indicate greater difficulties in perceiving others' feelings and moods and recognizing the consequences of their own behavior in the

environment, poor reciprocity skills in social relations and poor communication skills. In the present sample, ordinal alpha values were .89 and .95 at age 5 and .92 and .96 at age 10 for parents and teachers, respectively.

Strengths and Difficulties Questionnaire (SDQ^{3-4}; and SDQ^{4-16}; Goodman, 1997). The official Spanish parent's versions of the SDQ^{3-4} (Ezpeleta, Granero, de la Osa, Penelo, & Domènech, 2013) was used as a screening tool of children for inclusion in the study. The prosocial scale of SDQ^{4-16} (parents and teachers) was used at both follow ups. In both versions the 25 items (*e.g.: often loses temper*), have three response options (0: *not true*; 1: *somewhat true*; 2: *certainly true*). Ordinal alpha for the parents' scores in the present sample was .75 at age 5 and .83 at age 10, and for the teachers' scores the figures were .89 and .91 at age 5 and 10, respectively.

Inventory of Callous-Unemotional traits (ICU; Frick, 2004). This includes 24 items (*e.g.: does not show his/her emotions*) coded on a 4-point Likert-type scale (0: *not at all true* to 3: *definitely true*) and covers three dimensions: callousness (11 items), which attempts to measure the degree to which the child is unaware of other people's feelings and does not care about behaving in a socially acceptable way; uncaring (8 items), indicating the degree to which the child worries about accomplishing duties and the acceptance of mistakes and their consequences; and unemotional (5 items), which addresses the child's difficulties in sharing emotions or openly expressing feelings. The Spanish version was used (Ezpeleta, de la Osa, Granero, Penelo, & Domènech, 2013). Internal consistency (Cronbach's alpha) for the total score in the sample at ages 5 and 10 was .89 and .90, respectively, and between .75 and .90 for the three scale scores at both follow ups. ICU was answered by teachers.

Children's Aggression Scale (CAS; Halperin & McKay, 2008). The CAS assesses aggressive behavior with 22 items (*e.g.: during the last year, how often has he/she started a fight with another child?*) on a 5-point Likert-type scale (0*: never* to *4: many days*). This

questionnaire was answered by teachers and the total score was used. Cronbach's alpha values were .85 and .88 at 5 and 10 years old respectively.

Diagnostic Interview of Children and Adolescents for Parents of Preschool Children and Young Children (DICA-PPYC). The DICA-PPYC is used to assess children's psychopathology according to the (DSM-IV/DSM-5 taxonomy (American Psychiatric Association, 1994, 2013)). The presence of a diagnosis of ODD, ADHD and any anxiety (separation, generalized social or specific phobia) were considered. This interview, which was answered by parents, has been adapted and validated for the Spanish preschool population, showing good psychometric properties (Ezpeleta et al., 2011). Kappa values for the test-retest and between raters´ concordance was between excellent and good for most DSM-IV diagnoses. The interview discriminated between clinical and non-clinical groups. Also, concordance between interview' diagnosis and clinicians' was studied.

Children's Global Assessment Scale (CGAS; Shaffer et al., 1983). The Spanish adaptation (Ezpeleta, Granero, & de la Osa, 1999) of this scale was used to assess global functional impairment based on children's psychopathology as rated by interviewers. The total score was used, ranging from 0 (*highest impairment*) to 100 (*lowest impairment*). Kappa coefficients for inter-interviewer agreement ranged from moderate to very good (.45-behavior toward others- to .94 -Self-harmful behavior).

Preschool and Early Childhood Functional Assessment Scale (PECFAS) and the *Child and Adolescent Functioning Assessment Scale (CAFAS)* (Hodges, 1995). Both instruments determine the extent to which a subject's functioning is impaired in each of eight psychosocial areas. Four scales were used in this study (School, Home, Conduct and Humor), given that the prevalence of impairment due to autolysis, use of substances, cognition problems and difficulties functioning in the community was extremely low. PECFAS was used for the follow ups at age 5 and CAFAS was used to assess the children in the 10 years

old follow up. Both instruments, are scored based on four levels of impairment (0 = *no or minimal impairment*; 10 = *mild impairment or distress*; 20 = *moderate impairment*; and 30 = *severe impairment*). CAFAS has good psychometric properties in the Spanish population. In this study, we analyzed the dichotomous classification [no impairment (0) versus minimal, mild to severe (10-30)] rated by interviewers.

Child Behavior Checklist (CBCL $1^{1/2}$ – 5 and CBCL 6-18; Achenbach & Rescorla, 2001). The CBCL measures behavioural and emotional problems through 100 items (e.g.: shy; steals at home) with three response options (0: *not true*, 1: *somewhat/sometimes true*, 2: *very true/often true*), and was answered by parents. The withdrawn (age 5), withdrawn/depressed (age 10) and social problems (age 10) empirical scale scores were used. For these three measures, Cronbach's alpha values in the present sample were .87, .89 and .86, respectively.

Risk Factor Interview (UED, 2009). Children self-informed at age 12 (N = 379), answering (yes/no) 3 open questions about friends: *Do you usually have conflicts with friends? Is it difficult for you to make friends?* And *Is it difficult for you to maintain friends?* These three questions were part of a risk factors interview and meant to be used as criteria for the outcome of social functioning.

Statistical Analysis

The analyses were conducted using MPlus 8.4 (Muthén & Muthén, 1998-2019) and Stata 16.1 (StataCorp, 2019). Given that the multistage sampling gave unequal probabilities of being selected depending on the screening group, the analyses were weighted by the inverse of the probability of participants' being selected in the second phase of the project.

First, measurement invariance of the SCDC items was analyzed using Confirmatory Factor Analysis (CFA) for categorical indicators with the Weighted Least Squares Means and Variance adjusted (WLSMV) method of estimation and theta parameterization. For CFA, the

following goodness-of-fit indices were used (Jackson, Gillaspy, & Purc-Stephenson, 2009): chi-square (χ^2), Comparative Fit Index (CFI), Tucker–Lewis Index (TLI) and Root Mean Square Error of Approximation (RMSEA), the latter as the main index of model fit (Hu & Bentler, 1998). We considered the following thresholds: CFI and TLI > .90 and RMSEA < .08 for acceptable fit, and CFI and TLI > .95 and RMSEA < .06 for excellent fit (Brown, 2006). The invariance analyses involved comparing progressively more constrained nested models (from least to most restrictive), following the common sequence: configural (equal forms), metric or weak (equal facto loadings), scalar or strong (equal thresholds), strict (equal uniquenesses) for measurement invariance; further steps can evaluate structural parameters (equal factor variances, factor covariances for multidimensional instruments, and latent means), which can be considered tests of population heterogeneity (Brown, 2006). Table S1 expands details involving each step and model identification. For invariance across sex, a multi-group approach was used; for invariance across age and across informant, a repeated-measure single-group approach was used, and error covariances between analogous items were also freely estimated. The α level for scaled difference chi-square tests between nested models at each step with the "Difftest" option of Mplus for WLSMV method of estimation was set at .01 (e.g., Ferrando, 2000). If full invariance was not accomplished, modification indexes were examined in order to free parameters one at a time to be able to evaluate partial invariance (not attained when more than 20% of parameters are found to be non-invariant; Dimitrov, 2010).

Second, latent class analysis (LCA) was estimated to group children based on response to the 12 SCDC items for each informant and age separately. Given the low endorsement for the *very true* option (coded as 2), and to optimize the obtaining of classes as distinguishable as possible, we reduced the number of categories, by collapsing the two highest categories into one category; then, for LCA all SCDC items will be dichotomous (0 = *not true* vs. 1 or 2

= *quite true* or *very true*), equivalent to a present/absent format. LCAs with one to five classes were conducted using a robust maximum likelihood estimator (MLR). Several goodness of fit indexes were used to determine the optimum number of classes: 1) larger decrement in aBIC; 2) the Vuong-Lo-Mendell-Rubin (VLME) likelihood test and the Lo-Mendell-Rubin (LMO) likelihood ratio test comparing k versus k − 1 number of classes solution; 3) posterior probabilities of belonging to the assigned class above .70; 4) classes with more than 5% of the total sample; and 5) entropy above .80. Additionally, the best clinical interpretability was considered. Moreover, and as a sensitivity analysis with the aim of explore the robustness of our main results, we repeated the analyses by limiting the participants to the subsample conformed by Caucasian children. Given the small percentage of non-white children (10.2% at age 5 and 8.2% at age 10), it was not possible to conduct a separate analysis for this sub-group.

Third, regression models (logistic for binary outcomes and linear for metric outcomes) were estimated at each age. For each outcome, two regressions were conducted, one for each informant (parents and teachers); for each of these regressions, the predictor was the variable class resulting from the solution selected in the corresponding LCA. Next, coefficients between both models, one for each informant at each age, were compared. Outcomes concurrently considered were dimensional and categorical information about psychopathology, impairment, callousness traits, and social functioning. Diagnose of regression models was done by checking normality (Shapiro-Wilk test and standardized normal probability plot) and homogeneity of variances (Breusch-Pagan test and graphics) for linear models (Kleinbaum et al., 2014) and presence of outlier and/or influential cases for both linear (analysis of residuals and Cook distance) and logistic (analysis of residuals and delta chi-square) models (Hosmer et al., 2013).

Results

Invariance Analyses

Measurement invariance analyses showed that there was full (all factor loadings, thresholds, and uniquenesses) or almost full (all the parameters except those for 1 of the 12 items) equivalence across sex within each informant at both age 5 and age 10 (Table S1, models A-D). These results indicate that scores are directly comparable between females and males. For the parent version, latent means did not differ across sex ($d \leq 0.04$; $p \geq .711$), whereas for the teacher version females scored slightly lower than males at both age 5 ($d = 0.35$; $p = .006$) and age 10 ($d = 0.51$; $p = .001$). Regarding invariance over time, there was full measurement equivalence within teachers (Table S1, model F), and partial measurement equivalence (all parameters except 4 of 24 [16.7%] thresholds) within parents (Table S1, model E). Therefore, these results allow the direct comparison of SCDC scores between preschool and school age. Latent means did not differ between age 5 and age 10 ($d \leq 0.13$; $p \geq .122$). Regarding invariance across informants within each age (Table S1, models G-H), all factor loadings were found to be equivalent (full metric invariance), but 9 of the 24 (37.5%) thresholds, which involved items 1, 2, 3, 4, 7, 8, 10 and 11 at age 5, and 6 of the 24 (25.0%) thresholds, which involved items 1, 2, 3, 8, 10 and 11 at age 10, were non-equivalent between parents' and teachers' responses (non-scalar invariance). The parents rated most non-equivalent items more frequently as ''*not true*'' (higher first threshold) than did teachers, thus reporting less problems than teachers, except for item 7 at age 5 and items 7 and 11 at age 10, with a pattern in the opposite direction. These results mean that it was not possible to establish at least partial scalar invariance (≥ 80% of parameters as equivalent) and, therefore, the observed scores are not directly comparable between informants.

Latent Classes and Prediction of Outcomes

Tables 1 and 2 show the goodness of fit indexes for the LCA analyses carried out with one to five classes for each informant at age 5 and age 10, respectively. The solutions with

two classes were selected based on better interpretability and the five statistical criteria mentioned in the method section. In all cases, the two resulting classes could be labelled as high SC difficulties (class 1) and low SC difficulties (class 2), since at age 5 all the items except two of them showed statistically significant OR values when comparing the endorsement between 0-*not true* (absent) and 1-*quite true* plus 2-*very true* (present) options across classes, for both parents and teachers. Specifically, for 10 of the 12 SCDC items, the OR value was statistically above 1, meaning that participants with an item response of present with respect to absent were more likely classified in class 1 (high SC difficulties) than in class 2 (low SC difficulties). A similar profile was observed between the two classes at age 10, with statistically significant OR values for half of the items (data not shown, available on request).

At age 5, 62.8% of children were assigned to the same class according to parents' and teachers' ratings (46.8% low; 16.0% high), 10.3% were assigned to the low SC difficulties class according to parents' ratings and the high SC difficulties class according to teachers' ratings, and 26.9% were assigned to the high SC difficulties class according to parents' ratings and the low SC difficulties class according to teachers' ratings. At age 10, the pattern was similar: 63.5% were assigned to the same class for both ratings (48.2% low, 15.3% high), 14.5% low-parents and high-teachers, and 22.0% high-parents and low-teachers. When the LCA was restricted to only Caucasian children, very similar results were obtained: the 2-class solution was the optimal for both informants and ages, with a maximum of 5 from 507 (parents' ratings) or from 519 (teachers' ratings) children being classified in a different class from that assigned in the main analysis with the whole sample (McNemar test: $p \geq .063$).

Table 3 shows the results for the comparison of predicting outcomes due to belonging to the higher SC difficulties vs the lower SC difficulties class between parents' and teachers' LCA models. At age 5, the parents' ratings better predicted the presence of anxiety ($p = .039$), withdrawal ($p < .001$), and functioning at home ($p = .049$). Belonging to the class with higher

social cognition difficulties is a risk factor for emotional problems. The teachers' ratings better predicted traits of callousness ($p < .001$) and uncaring ($p < .001$), aggressive behavior ($p < .001$), and prosocial behavior as reported by themselves ($p < .001$), indicating that children included in the class of higher difficulties regarding teachers scores presented higher problems with empathy, aggression and social relations. At age 10, the parents' ratings were better predictors of the belonging to the higher difficulties vs. the lower difficulties class in the presence of better functioning at home ($p < .001$), conduct ($p = .022$), general functioning ($p = .009$), prosocial behavior when they are the informants, withdrawal ($p = .029$), and social problems ($p < .001$). The teachers' ratings better predicted the same outcomes at age 5 (callousness, aggression, and prosocial reported by themselves; $p < .001$).

The diagnose of linear regression models showed some significant results of normality and homochedascity statistical tests. Nevertheless, they were mainly due to the high sample size since the checking of the same assumptions by graphical analysis reflected an adequate fit. Some outlier and influential cases were also detected, but it was verified that the influential children were not the same across a high set of statistical models, so the existence of influential cases was ruled out.

Discussion

Differences in source-specific information can be both context and concept related. Parents and teachers rating on a SC measure is fully invariant across sex and partially invariant across age and so measures are comparable between females and males and between 5- and 10-year-old children. However, the measure was not invariant between parent and teacher raters; therefore, as regard informants, and for both developmental moments, the parents' and teachers' mean scores are not comparable. Two latent classes, high SC difficulties and low SC difficulties, were the best solution to represent the children's abilities

in SC. Comparisons of the prediction of belonging to one of these classes as regards different outcomes showed that this was dependent on the informant. The best prediction was informed by the source also informing about the outcome. This is aligned with other studies (Drabrick et al., 2011)

All those items that were non-equivalent in both follow-ups refer to/ask about cognitions, not behaviors, which is interpreted as a major discrepancy regarding non-observable behaviors and a different "norm" for parents and teachers when interpreting children's "cognitions" and social reciprocity. This is a concept that is closer to the Theory of Mind, (ToM) a core aspect of SC (Frith & Frith, 2010). More precisely, based on face validity and factor analysis, items 1, 2, 3 and 8 have been previously identified as a subset of items indexing pure socio-emotional competence (SEC) (Mandy et al., 2013; Oliver et al., 2011). This SEC subset excludes the SCDC items concerned with the behavioral (ex. C*hild is very demanding*) or emotional aspects (ex. *Child is difficult to reason with when upset"*) of social communication difficulties This specific factor includes the two original domains of social reciprocity and non-verbal skills proposed in the original work (Skuse et al., 2005). The other two non-equivalent items (4 and 11) are concerned with family conduct and language use. This could be interpreted as a real different conceptualization of SEC between parents and teachers perhaps due to the different "norms" they use to compare behaviors. This would also support the idea of informant-specific characteristics of children identified as having difficulties regarding SC. Teachers were more likely to endorse items related strictly to the ToM may be due to their classroom salience and greater opportunity to observe the child when they are with friends (King et al., 2018). Hence, we could say that the different information provided by parents and teachers might not be simply a matter of context but the result of different concepts about SC, and while parents report about difficulties based on observable conducts, teachers better report about implicit cognitive components.

The concordance between informants is much higher in the low SC cognition difficulties class and, if we observe the percentages of each informant, it seems that parents are more likely to include children in the higher scores class than are teachers. The latter could be related to the different "own norms" used by teachers with a larger number of children to establish comparison and, subsequently, the possibility of relativizing certain social interactions (Mcneilis et al., 2017). Social desirability could also be implicated and explain part of these results as parents could be reluctant to inform about difficulties.

As regards the validity of the classes created by different informants, no differences were found regarding predicting different outcomes relating to the presence of several disorders. The only exception was anxiety disorders as assessed by means of the structured interview at age 5. The association between parents' information about SC and the presence of any anxiety disorder could be since parents could assimilate avoiding conducts during childhood with SC problems. In fact, has been observed that adolescents at risk of belonging to a high social anxiety trajectory can be discriminated from peers belonging to a lower anxiety trajectory using both cognition and social competence variables (Miers, Blöte, De Rooij, Bokhorst, & Westenberg, 2013). Parents' information as assessed by clinicians is also better for predicting daily functioning at home at any follow up and parents better assess the interference of social cognition difficulties in children's behavior and general functioning at age 10. This indicates that information provided by teachers about SC does not predict the parents' point of view of children's school functioning.

The teachers' scores placed children presenting difficulties with empathy and caring about other people and children with more aggressive problems in the SC high difficulties class. The general tendency is that better predictions of outcomes are mostly those when the informant is the same for both variables: the predictor and the outcome. These results align with those of (De Los Reyes et al., 2009; Lavigne, Dahl, Gouze, Lebailly, & Hopkins, 2015).

The inclusion of children in one or another class was predictive of the prosocial scale score level of the SDQ, the one most related to the SCDC construct, at age 10 for parents and at both follow ups for teaches. Again, the best prediction was informant dependent.

Neither parents' nor teachers' information were predictor of the children's own perception of the quality of their friendship relationships at age 12. This could be explained by the fact that the children's information was collected two years later, and in the meanwhile a change from elementary to high school had taken place with a possible change in friendships and less parental control on these friendships, in addition to greater distancing between teachers and students.

The results of this study allow us to say that parents and teachers identify different children when they are asked to detect SC deficits, not only because they are informing about different settings, but also because they have different concepts about "norms" and behaviors when they are asking about children.

The current study has multiple strengths, including the use of a large community sample followed over a long time and the use of multiple informants and methods. The fact that we are working with a general sample, might be concealing other relevant differences among source-specific information about SC in clinical settings. Further research needs to be carried out to know more about risk factors and outcomes for source-specific symptoms of SC difficulties and specifically those that include children's own perceptions about their social outcomes. Still , we could conclude that in clinical settings it is important to consider both teachers and parents to fully understand SC, as is the case with many other constructs (Lavigne et al., 2015; Mayes & Lockridge, 2018; Stuart, 2018). Dismissal of the information from one of these sources could result in partial or inadequate detection and interventions that could leave out of clinical attendance children and adolescents with SC difficulties and

components of such a complex behavior. Relevance of measure invariance across development, gender and informants should then be taken into research consideration.

References

Achenbach, T., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age and Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth and Families.

Adolphs, R. (2001). The neurobiology of social cognition. *Current Opinion in Neurobiology*, *11*(2), 231–239. https://doi.org/10.1016/S0959-4388(00)00202-6

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders -IV*. Washington, DC: Author.

American Psychiatric Association. (2013). *DSM-5 Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Press.

Arioli, M., Crespi, C., & Canessa, N. (2018). Social Cognition through the Lens of Cognitive and Clinical Neuroscience. *BioMed Research International*, *2018*. https://doi.org/10.1155/2018/4283427

Blakemore, S.J. (2010). The Developing Social Brain: Implications for Education. *Neuron*, *65*(6), 744–747. https://doi.org/10.1016/j.neuron.2010.03.004

Bora, E., & Pantelis, C. (2016). Meta-analysis of social cognition in attention-deficit/hyperactivity disorder (ADHD): Comparison with healthy controls and autistic spectrum disorder. *Psychological Medicine*, *46*(4), 699–716. https://doi.org/10.1017/S0033291715002573

Brown, T. A. (2006). Confirmatory factor analysis for applied research. Guilford Press.

De Los Reyes, A., Henry, D. B., Tolan, P. H., & Wakschlag, L. S. (2009). Linking informant discrepancies to observed variations in young children's disruptive behavior. *Journal of Abnormal Child Psychology*, *37*, 637–652. https://doi.org/10.1007/s10802-009-9307-3

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 studies. *Psychological Bulletin*, *131*(4), 483–509.

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development, 43*, 121-149. https://doi.org/10.1177/0748175610373459

Drabick DAG, Bubier JL, Chen D, Price J, Lanza H. (2011) Source-specific oppositional defiant disorder among inner-city children: Prospective prediction and moderation. J Clin Child Adolesc Psychol. 2011; 40:23–35. [PubMed: 21229441

Evans, S. C., Pederson, C. A., Fite, P. J., Blossom, J. B., & Cooley, J. L. (2015). Teacher-reported irritable and defiant dimensions of oppositional defiant disorder: social, behavioral, and academic correlates. *School Mental Health*, *8*(2), 292–304. https://doi.org/10.1007/s12310-015-9163-y

Ezpeleta, L, Granero, R., & de la Osa, N. (1999). Evaluación del deterioro en niños y adolescentes a través de la Children's Global Assessment Scale (CGAS). *Revista de Psiquiatría Infanto-Juvenil*, *1*, 18–26.

Ezpeleta, L., de la Osa, N., Granero, R., Penelo, E., & Domènech, J. M. (2013). Inventory of callous-unemotional traits in a community sample of preschoolers. *Journal of Clinical Child and Adolescent Psychology, 42*, 91–105. https://doi.org/10.1080/15374416.2012.734221

Ezpeleta, L., Granero, R., de la Osa, N., Penelo, E., & Domènech, J. M. (2013). Psychometric properties of the Strengths and Difficulties Questionnaire(3-4) in 3-year-old preschoolers. *Comprehensive Psychiatry*, *54*(3), 282–291. https://doi.org/10.1016/j.comppsych.2012.07.009

Ferrando, P. J. (2000). Testing the equivalence among different item response formats in personality measurement: A structural equation modeling approach. *Structural Equation Modeling, 7*, 271-286. https://doi.org/10.1207/S15328007SEM0702_7

Frith, U., & Frith, C. (2010). The social brain: Allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1537), 165–176. https://doi.org/10.1098/rstb.2009.0160

Gilmour, J., Hill, B., Place, M., & Skuse, D. H. (2004). Social communication deficits in conduct disorder: a clinical and community survey. *Journal of Child Psychology and Psychiatry*, *45*(5), 967-978. https://doi.org/10.1111/j.1469-7610.2004.t01-1-00289.x

Goodman, R. (1997). The Strenghs and Difficulties Questionnaire. *Journal of Child Psychology and Psychiatry*, *38*, 581–586.

Happé, F., & Frith, U. (2014). Annual research review: Towards a developmental neuroscience of atypical social cognition. *Journal of Child Psychology and Psychiatry*, *55*, 553–577. https://doi.org/10.1111/jcpp.12162

Hodges, K. (1995). *Child and Adolescent Functional Assessment Scale (CAFAS)*. Ypsilanti, MI: Eastern Michigan University, Department of Psychology.

Hollingshead, A. A. (1975). Four-factor index of social status. Unpublished manuscript, Yale University, New Haven, CT.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. 3rd ed. John Wiley and Sons.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424-453. http://dx.doi.org/10.1037/1082-989X.3.4.424

Hughes, C., White, A., Sharpen, J., & Dunn, J. (2000). Antisocial, angry, and unsympathetic: "Hard-to-manage" preschoolers' peer problems and possible cognitive influences. *Journal of Child Psychology and Psychiatry*, *41*(2), 169–179. https://doi.org/10.1111/1469-7610.00558

Jackson, D. L., Gillaspy, J. A., & Purc-Stephenson, R. (2009). Reporting practices in

confirmatory factor analysis: An overview and some recommendations. *Psychological Methods, 14*, 6-23. https://doi.org/10.1037/a0014694

Kilford, E. J., Garrett, E., & Blakemore, S. J. (2016). The development of social cognition in adolescence: An integrated perspective. *Neuroscience and Biobehavioral Reviews*, *70*, 106–120. https://doi.org/10.1016/j.neubiorev.2016.08.016

King, K. M., Luk, J. W., Witkiewitz, K., Racz, S., McMahon, R. J., & Wu, J. (2018). Externalizing behavior across childhood as reported by parents and teachers: A partial measurement invariance model. *Assessment, 25(6)* 744-758. https://doi.org/10.1177/1073191116660381

Kleinbaum, D. G., Kupper, L. L., Nizam, A., Muller, K., & Rosenberg, E.S. (2014). *Applied regression analysis and other multivariable methods*. (5rd ed.). Cengage Learning, Inc.

Lavigne, J. V, Dahl, K. P., Gouze, K. R., Lebailly, S. A., & Hopkins, J. (2015). Multi-domain predictors of oppositional defiant disorder symptoms in preschool children: Cross-informant differences. *Child Psychiatry Human Development*, *46*(2), 308–319. https://doi.org/10.1007/s10578-014-0472-4

Mandy, W., Skuse, D., Steer, C., St Pourcain, B., & Oliver, B. R. (2013). Oppositionality and socioemotional competence: Interacting risk factors in the development of childhood conduct disorder symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry*, *52*(7), 718–727. https://doi.org/10.1016/j..jaac.2013.04.011

Mayes, S. D., & Lockridge, R. (2018). Brief Report: How Accurate is teacher report of autism symptoms compared to parent report? *Journal of Autism and Developmental Disorders, 48*, 1833–1840. https://doi.org/10.1007/s10803-017-3325-8

Mcneilis, J., Maughan, B., Goodman, R., & Rowe, R. (2017). Comparing the characteristics and outcomes of parent-and teacher-reported oppositional defiant disorder: Findings from a national sample. *Journal of Child Psychology and Psychiatry, 59 (6),* 659-666.

https://doi.org/10.1111/jcpp.12845

Miers, A. C., Blöte, A.W., De Rooij, M., Bokhorst, C. L., & Westenberg, P. M. (2013). Trajectories of social anxiety during adolescence and relations with cognition, social competence, and temperament. *Journal of Abnormal Child Psychology*, *41*, 97–110. https://doi.org/10.1007/s10802-012-9651-6

Oliver, B. R., Barker, E. D., Mandy, W. P. L., Skuse, D. H., & Maughan, B. (2011). Social cognition and conduct problems: A developmental approach. *Journal of the American Academy of Child and Adolescent Psychiatry*, *50*(4) 385-394. https://doi.org/10.1016/j.jaac.2011.01.006

de la Osa, N., Granero, R., Penelo, E., & Ezpeleta, L. (2014). Usefulness of the Social and Communication Disorders Checklist (SCDC) for the assessment of social cognition in preschoolers. *European Journal of Psychological Assessment*, *30*(4), 296–303. https://doi.org/10.1027/1015-5759/a000193

Poletti, M., & Adenzano, M. (2013). *Theory of Mind in non- authistic psychiatric disorders of childhood and adolescence*. *Neuropsychiatry, 10(5),* 188–195.

Shaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., & Bird, H. (1983). A children's global assessment scale (CGAS). *Archives of General Psychiatry*, *40*, 1228–1231.

Skuse, D H, James, R. S., Bishop, D. V. M., Coppin, B., Dalton, P., AamodtLeeper, G., … Jacobs, P. A. (1997). Evidence from Turner's syndrome of an imprinted X-linked locus affecting cognitive function. *Nature*, *387*(6634). https://doi.org/10.1038/42706

Skuse, D H, Mandy, W. P. L., & Scourfield, J. (2005). Measuring autistic traits: Heritability, reliability and validity of the Social and Communication Disorders Checklist. *British Journal of Psychiatry*, *187,* 568-572. https://doi.org/10.1192/bjp.187.6.568

Skuse, David H, Mandy, W., Steer, C., Miller, L. L., Goodman, R., Lawrence, K., …

Golding, J. (2009). Social communication competence and functional adaptation in a general population of children: Preliminary evidence for sex-by-verbal IQ differential risk. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*(2), 128-137. https://doi.org/10.1097/CHI.0b013e31819176b8

Sonuga-Barke, E., Auerbach, J., Campbell, S. B., Daley, D., & Thompson, M. (2005). Varieties of preschool hyperactivity: Multiple pathways from risk to disorder. *Developmental Science*, *8*(2), 141–150. https://doi.org/10.1111/j.1467-7687.2005.00401.x

Stuart, J. (2018). Predicting risky health behaviors 35 years later: Are parents or teacher's reports of childhood behavior problems a better judge of outcomes? *Addictive Behaviors, 84,* 255-262. https://doi.org/10.1016/j.addbeh.2018.04.023

Trentacosta, C. J., & Fine, S. E. (2010). Emotion knowledge, social competence, and behavior problems in childhood and adolescence: A meta-analytic review. *Social Development*, *19*(1), 1–29. https://doi.org/10.1111/j.1467-9507.2009.00543.x

Tye, C., Battaglia, M., Bertoletti, E., Ashwood, K. L., Azadi, B., Asherson, P., … McLoughlin, G. (2014). Altered neurophysiological responses to emotional faces discriminate children with ASD, ADHD and ASD+ADHD. *Biological Psychology*, *103*, 125–134. https://doi.org/10.1016/j.biopsycho.2014.08.013

Unitat d'Epidemiologia i de Diagnòstic en Psicopatologia del Desenvolupament. (2009). Schedule for Risk Factors. Parents Version. Unpublished manuscript. Universitat Autònoma de Barcelona, Departament de Psicologia Clínica i de la Salut