Psychometric validation of a parent-reported measure of childhood alexithymia

Psychometric validation of a parent-reported measure of childhood alexithymia: the Alexithymia Questionnaire for Children – Parent (AQC-P)

Subtitle: Psychometric Properties of the AQC-P

Ruth Harriet Brown[1], Aja Murray[1,2], Mary E. Stewart[3] & Bonnie Auyeung [1, 4]

[1] Psychology Department, The University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, United Kingdom

[2] Institute of Criminology, University of Cambridge, Cambridge, UK

[3] Psychology, Heriot-Watt University, Edinburgh, Scotland, EH14 4AS, United Kingdom

[4] Autism Research Centre, Department of Psychiatry, University of Cambridge, Cambridge, CB2 8AH, United Kingdom

Correspondence to: Ruth Harriet Brown (email: rbrown11@exseed.ed.ac.uk)

Article Type: Original Article

Manuscript Word Count: 4984

Abstract

Alexithymia can be defined as difficulties in describing one's emotions and is of interest within clinical and developmental psychology as a potential mediating and exacerbating factor across multiple forms of psychopathology. Measuring alexithymia via self-reports can be challenging, as those with heightened alexithymia may have difficulties in recognising their alexithymia traits due to impaired metacognitive skills. Thus, there would be considerable benefits to the availability of a psychometrically validated parent-reported alexithymia measure that may circumvent the issue of self-reports. We therefore examined the psychometric properties of a new parent-reported alexithymia measure, the Alexithymia Questionnaire for Children – Parent (AQC-P) in a community sample of 257 child-parent dyads. Furthermore, we examined the level of agreement between the parent-rated AQC-P and its self-rated counterpart, the Alexithymia Questionnaire for Children (AQC). Confirmatory factor analysis found that an oblique three-factor structure provided the best model for both AQC-P and AQC, with this structure showing measurement invariance across child gender. All subscales had omega internal consistency values $>.70$, supporting their reliability. Cross-informant consistency was supported by significant correlations between AQC and AQC-P scores. Results support the use of the AQC-P as a measure of parent-reported alexithymia in children.

Alexithymia refers to difficulties in identifying and describing one's emotions (Sifneos, 1973) and is thought to affect approximately 10% of the adult population (Mattila et al., 2006). While it is not itself recognised as a psychiatric illness (Ricciardi et al., 2015), it is considered to be a significant aggravating factor in many psychiatric conditions (Grabe et al., 2004) and is therefore of considerable interest to developmental and clinical psychologists. However, alexithymia measurement is challenging because those who would be expected to score high on these traits may lack the metacognitive skills required to recognise their own alexithymic traits (Taylor et al., 1999). This issue may be further exacerbated in child populations due to their incomplete cognitive development and more limited abilities to interospect on their emotions (Parker et al., 2010). In this study we report on the validation of a new parent-reported measure of alexithymia to assess the construct in children, thereby potentially circumventing the issue of self-report.

Recognising the possible limitations of self-reported measures of alexithymia, a small number of authors have previously developed observer-rated assessment tools for use in childhood yielding the Alexithymia Scale for Children – Teacher Form (ASC-TF; Fukunishi et al., 1998), the Children's Alexithymia Measure (CAM; Way et al., 2010) and the Alexithymia Questionnaire for Children – Parent (AQC-P; Costa et al., 2017). Only the CAM and the AQC-P were developed for use in parents however, with the former previously shown to correlate non-significantly with the most widely used child measure of alexithymia, the Alexithymia Questionnaire for Children (AQC; Rieffe, et al., 2006; Griffin et al., 2016). Further, only the AQC-P has items that correspond in content to the original adult alexithymia measure, the Toronto Alexithymia Scale-20 (TAS-20; Bagby et al., 1994), making it possible to obtain a multi-informant perspective using comparable sets of items. However, the AQC-P has yet to undergo extensive psychometric validation, with only the measure's internal consistency reported in previous publications (Costa et al., 2017; Costa et

al., 2019). Questions thus remain regarding its factorial validity, level of agreement with its self-reported counterpart the AQC and gender invariance; the latter being important to test given previous evidence of a higher prevalence of alexithymia in males with the TAS-20 (Levant et al., 2009). Likewise, the AQC-P's convergent validity with known correlates of alexithymia (e.g., depressive symptoms; Parker et al., 1991, and empathic/prosocial behaviour; Grynberg et al., 2010) remains unexplored. Thus, in this study we provide an initial psychometric examination of the AQC-P in a community ascertained sample of 257 child-parent dyads.

Method

We report how we determined our sample size, all data exclusions, all data inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all measures in the study, and all analyses including all tested models. If we use inferential tests, we report exact p values, effect sizes, and 95% confidence intervals (CI).

Participants

Recruitment occurred over three waves of questionnaire distribution. First, children from Grade 4 to Grade 6 were recruited from two schools in the UK that agreed to participate; one local-authority and one private. In total, 521 families were approached to take part in the study. Of those approached, 175 families volunteered to take part, producing a response rate of 34%. Second, all the families in a university database of volunteers who were willing to take part in developmental studies and who had a child aged between 8-13 years (N = 45) were contacted 25 families of whom took part, producing a response rate of 55%. Finally, 57 families from a separate University of Edinburgh database of volunteers interested in taking part in research consented to complete the AQC-P. This gave an initial sample size of 257. Five children who had a diagnosis of autism spectrum disorder (ASD)

were excluded from the sample to avoid confounding of ASD and alexithymic traits. This exclusion criterion was established prior to the data analysis.

The sample consisted of 121 boys and 131 girls between 8 to 13 years (mean age = 10.13; SD = 1.06), with 21.6%, 41.6% and 36.8% from Grade 4, Grade 5 and Grade 6, respectively. The majority of parent participants were female (90%) and had a high level of education; 1.2% had no qualifications, 8.0% had school qualifications, 35.5% obtained an undergraduate degree, 26.4% obtained a postgraduate degree and 11.2% obtained a doctoral degree. The remaining 17.6% (44) did not disclose their highest educational attainment. All participants were fluent in English.

Materials

All questionnaires were administered in English and as a child and parent booklet provided to participating families, containing a self-report questionnaire for children and a parent-report questionnaire for the parent(s).

Alexithymia Questionnaire for Children

The Alexithymia Questionnaire for Children (AQC, Rieffe et al., 2006) is a 20-item scale used to self-assess alexithymia in children as young as eight. It is built upon the three original subscales of the TAS-20 (Bagby et al., 1994); 'difficulty identifying feelings' (DIF), 'difficulty describing feelings' (DDF) and 'externally oriented thinking' (EOT), where items are reworded to aid understanding by children. The total AQC has been previously found to have good internal consistency (α >.700; Rieffe et al., 2006). Using Rieffe and colleagues' (2006) scoring system, items were rated on a three-point scale (1 = "Not True" to 3 = "True") in order to simplify the response scale for child participants. Scores ranged from 20 to 60, with higher scores indicating a greater degree of alexithymic traits.

Alexithymia Questionnaire for Children – Parent

The AQC was modified for use in parent(s) of young children by Costa and colleagues (2017). The Alexithymia Questionnaire for Children– Parent (AQC-P) retains the same wording used by Rieffe and colleagues' (2006) (e.g., "I am able to describe my feelings easily" became, "my child is able to describe their feelings easily"), with the same three subscales DIF, DDF and EOT. The measure has been previously found to have good internal consistency ($\alpha > .800$; Costa et al., 2017). Like the AQC, the AQC-P is rated on a 3-point Likert scale in order to alleviate potential score comparison issues. Scores ranged from 20 to 60, with higher scores indicating a greater degree of alexithymic traits.

Depression Self-Rating Scale

The Depression Self-Rating Scale (DSRS; Birleson, 1981) is an 18-item self-reported scale that assesses depressive symptoms in children. The measure's scores were found to have good internal consistency in the current study ($\alpha = .948$), similar to previous investigations (Birleson, 1981). Items are rated on a three-point scale (0 = "Never" to 2 = "Always"). Scores ranged from 0 to 36, with higher scores indicative of greater depressive symptoms.

Empathy Quotient for Children

The Empathy Quotient for Children (EQ-C; Auyeung et al., 2009) is a 27-item parent-reported assessment of empathic behaviour in children. The measure's scores have been previously found to have good internal consistency ($\alpha > .900$, Auyeung et al., 2009), similar to the current study ($\alpha = .856$). Items were rated on a four-point scale, with "Definitely Agree" and "Slightly Agree" responses endorsing empathic behaviour scored as 2 and 1, respectively. "Definitely Disagree" and "Slightly Disagree" responses were scored as 0. Scores ranged from 0 to 54, with high scores indicative of higher empathy.

Strengths and Difficulties Questionnaire

The Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) is a 25-item parent-reported behavioural screening tool composed of five subscales that assess the child's negative and positive behaviour; 'conduct problems', 'inattention-hyperactivity', 'emotional symptoms', 'peer problems' and 'prosocial behaviour'. The SDQ scores were found to have good internal consistency ($\alpha = .722$), similar to previous findings (Goodman et al., 1997). Items were rated on a three-point scale (1 = "Not True" to 3 = "Certainly True"). Scores ranged from 5 to 15 for each subscale, with lower scores indicative of a behavioural strength and lower scores in the subscale 'prosocial behaviour' as a behavioural difficulty.

Procedure

For the school sample, questionnaire packs were distributed to pupils to be completed at home. For the database sample, participating families were sent their questionnaire packs via post. Before taking part, children and parent(s) were asked to read an information sheet and sign consent forms if they wished to participate. Families were informed that the study was voluntary and any information they provided would be anonymous. Parent(s) were additionally asked to provide demographic information on their child, including their age, gender and any developmental difficulties. Both children and parents were asked to complete the questionnaire booklets independently, however the children were informed they could ask their parent to clarify any item(s) that they did not understand. On completion, the questionnaires were either returned to school by the children to be collected by the researcher, sent back to the university via a pre-paid envelope or given back to the researchers during the laboratory study. This study received ethical approval from the authors' research ethics committee (62-1516/6).

Data Analysis

Confirmatory factor analyses were conducted on the AQC and AQC-P scores in order to evaluate factorial validity. Three factor structures were investigated. First, the original three-factor model (DIF, DDF and EOT; Bagby et al., 1994) was explored. Next, a two-factor model proposed by Erni, Lötscher and Modestin (1997) (where DIF and DDF are merged together, known as 'DDIF') has been suggested to be a more suitable specification for the TAS-20. Therefore, this specification was also explored. Lastly, for completeness, a one-factor model (combining DIF, DDF and EOT) was investigated. Given the ordinal response format of the data, a diagonally weighted least squares (DWLS) estimation was used. Following the recommendations of Kline (2010), the suggested cut-offs <.060 for the root mean square error of approximation (RMSEA), <.080 for the standardised root mean square residual (SRMR) and >.900 for the comparative fit index (CFI) were used to identify adequate model-fit. Next, prior to investigating gender effects on AQC and AQC-P scores, measurement invariance was tested. As the data were categorical, the following was tested in the following order; configural invariance; threshold equivalence; threshold and loading invariance; and threshold, loading and intercept invariance. Invariance was assumed to hold if the differences in CFI, RMSEA and SRMR between nested models were less than .005, .010 and .005, respectively (Chen, 2007). In order to assess internal consistency, McDonald's omega values were then calculated for the AQC and AQC-P total and subscales scores. Mean differences in alexithymia ratings between informants were then analysed using paired t-tests. Cohen's *d* values were run in order calculate effect sizes for these differences. Lastly, to assess convergent validity, correlations were calculated between the AQC, the AQC-P and the additional questionnaires. Fisher z-transformations were used in order to identify any significant differences in the correlation coefficients. Analyses were conducted using both SPSS version 22, the R lavaan package for confirmatory factor analysis/measurement invariance (Rosseel, 2012) and psych package for the Fisher z-transformations (Revelle,

2018). Input/output files can be seen in the Electronic Supplemental Materials (ESM) 1 and 2.

Results

Data Cleaning

Missing data were first addressed. Sixteen children and fifteen parents omitted one or two items of their respective alexithymia measures. An Expectation Maximisation (EM) algorithm within SPSS was used to estimate the values of the missing data points, allowing full scores to be generated. Two children were identified as multivariate outliers using Mahalanobis' distances. The child and their parent's data were removed from the dataset, giving a final sample size of 250 child-parent dyads.

Preliminary Analysis

Descriptive statistics for the AQC-P and AQC scores are shown in Table 1. The descriptive statistics for the external measures are shown in ESM 3.

AQC-P Model Fit and Measurement Invariance

In order to investigate the fits of three-factor (model 1), two-factor (model 2) and one-factor (model 3) models, confirmatory factor analyses were conducted (see Table 2). The DWLS estimator produced a non-significant chi square goodness of fit test for the three-factor model [χ^2 (167) = 141.76, p = .922] but not the two-factor model [χ^2 (169) = 375.78, p <.001] nor one-factor model [χ^2 (190) = 523.93, p < .001]. The criteria for adequacy of fit were met for the three-factor model, as satisfactory values for the CFI (>.900), RMSEA (< .060), SRMR (<.080) emerged. Despite an adequate SRMR value, no other goodness-of-fit tests were met in both the two-factor and one-factor model. Measurement invariance across gender was then tested in the three-factor model for the AQC-P. It was found ΔCFI,

ΔRMSEA and ΔSRMR across the configural; threshold invariance; and threshold and loading invariance models were less than .005, .010 and .005, respectively (i.e., ΔCFI= .004, ΔRMSEA = .009 and ΔSRMR <.001) for the addition of threshold constraints; and ΔCFI= .004, ΔRMSEA = .001 and ΔSRMR < .001 for the addition of loading constraints), suggesting that invariance held up to the threshold and loading invariance level. However, ΔCFI was .007 and thus larger than .005, and ΔRMSEA was .012 and thus larger than .010 with the addition of intercept invariance constraints when added to the threshold and loading model (see ESM 4). Partial invariance up to the threshold, loading and intercept level (ΔCFI= .001, ΔRMSEA = .001 and ΔSRMR = .003) could be achieved when constraints were freed on item 20 (see ESM 5).

AQC Model Fit and Measurement Invariance

The above-described CFA and gender invariance analyses were also conducted for the AQC scores. First, the model fit of a one-, two- and three-factor structure was investigated (see Table 2). The DWLS estimator produced a significant chi square goodness of fit tests for the three-factor model [$\chi2$ (167) = 235.53, p <.001], the two-factor model [χ^2 (169) = 316.17, p <.001] and one-factor model [χ^2 (170) = 341.63, p < .001]. However, all other criteria for adequacy of fit were met for the three-factor model, as satisfactory values for the CFI (>.900), RMSEA (< .060), SRMR (<.080) emerged from the analysis. While the one- and two-factor models produced satisfactory RMSEA values, CFI values were below acceptable levels for both models (<.900) and SRMR values were below acceptable levels for the one factor model (<.060). Next, the degree of gender invariance was assessed using the three-factor model for the AQC (see ESM 6). While ΔRMSEA (.007) and ΔSRMR (<.001) were acceptable, ΔCFI was above the cut-off in the threshold invariance model (.006). However, follow-up analyses revealed partial threshold invariance (ΔCFI = .001; ΔRMSEA = .005; ΔSRMR = .002) when constraints were released on item 16 (see ESM 7). Furthermore, ΔCFI,

ΔRMSEA and ΔSRMR were less than .005, .010 and .005 respectively across the threshold and loading (ΔCFI = <.001; ΔRMSEA = .001; ΔSRMR = <.001) and threshold, loading and intercept invariance models (ΔCFI = .003; ΔRMSEA = .001; ΔSRMR = .001).

Reliability and Validity

Internal Consistency

Acceptable omega reliability values were found for the AQC-P total scores (ω = .870) and the DIF (ω = .890), DDF (ω = .700) and EOT (ω = .750) subscales. Likewise, the overall AQC produced acceptable omega values for the total scores (ω = .780) and the DIF (ω = .850) DDF (ω = .760) subscales. However, similar to the findings of Rieffe and colleagues (2006), the EOT subscale did not meet the acceptable level of internal consistency (ω = .560).

Correlations Between Total and Subscale Scores Across Raters.

The total AQC and AQC-P scores were significantly correlated (r = .325, p <.001). At the subscale level, child and parent DIF (r = .401, p <.001), DDF (r = .206, p <.001) and EOT (r = .345, p <.001) all showed significant correlations.

Rating Differences Between the AQC and AQC-P

To asses if there were significant rating differences between the AQC and AQC-P scores, paired t-tests were conducted. Overall, children rated themselves more alexithymic than their parent (t(249) = 6.26, p <.001, *d* = .461, 95% CI [1.87, 3.59)]). At the subscale level, children gave higher DIF (t(249) = 6.25, p <.001, *d* = .434, 95% CI [-1.67, -.870)]) and DDF (t(249) = 7.11, p <.001, *d* = .573, 95% CI [-1.62, -.918)]) ratings. However, there was no significant difference between parent and child EOT scores (t(249) = .954, p = .341, *d* = .071, 95% CI [-.598, .208)]).

Convergent Validity

Fisher z-transformations suggested that the AQC-P correlated significantly stronger with the EQ-C ($Z_{observed} = 2.75$, p = .006); and the 'prosocial behaviour' ($Z_{observed} = 3.21$, p = .001), 'inattention-hyperactivity' ($Z_{observed} = 2.17$, p = .030) and 'conduct problems' ($Z_{observed} = 3.21$, p = .036) subscales of the SDQ, compared to the AQC scores. Conversely, the AQC were found to correlate significantly stronger with DSRS scores ($Z_{observed} = 4.87$, p <.001) compared to the AQC-P (see Table 3).

Discussion

The aim of the current study was to assess the psychometric properties of the recently developed AQC-P to evaluate whether it is able to meet the current need for a parent-report measure of alexithymia, alongside the self-reported AQC. Analyses suggested that the AQC-P scores showed factorial validity with the instrument's hypothesised three-factor structure; good internal consistency; partial gender invariance up to the threshold, loading and intercept invariance level; significant correlations at the total score and subscale level with the AQC and convergent validity with the additional external measures administered.

A three-factor structure was found to be the best model for both the AQC-P and AQC scores as the models met all the goodness of fit tests, whereas one- and two-factor structures failed to reach the acceptable limits of model fit. These findings are consistent with previous investigations in child (Rieffe et al., 2006) and adult (Taylor et al., 2003) samples using the AQC/TAS-20. A high level of partial measurement invariance was observed in both scales across boys and girls at the threshold, loading and intercept level after releasing constraints on the non-invariant items. As valid comparisons can still be drawn despite a small number of non-invariant items (Pokropek et al., 2019), it was concluded that alexithymia was measured comparably across the genders with these scales. Thus, this supports the use of the AQC/AQC-P in examining gender differences in predictors/outcomes of alexithymic traits.

Omega internal consistency values were all >.700 for the AQC-P total and subscale scores, though fell below .700 for the AQC EOT subscale.

At the total score and subscale level, the AQC and AQC-P scores were significantly correlated. Consistent with previous work in child psychopathology which suggests that self- and parent-reports capture overlapping yet distinct aspects of child behaviour, the correlations were small to moderate in magnitude (De Los Reyes & Kazdin, 2005). Child self-reports yielded higher alexithymia ratings than parent reports. While it was not possible to gauge which informant provided the ratings that best reflected a child's true level of alexithymia, child-reports appeared to be more sensitive to detecting alexithymia than parent-reports.

However, Fisher z-transformations revealed that the AQC-P and AQC scores had unique correlation patterns with the external measures administered. Compared to the AQC, the AQC-P correlated more negatively with empathic and prosocial behaviour. In contrast, the AQC was found to correlate more positively with depressive symptoms when compared to the AQC-P. Thus, while children can accurately report on their negative affect, they may have difficulties in reporting the external negative behaviours associated with alexithymia. In contrast, parents may accurately observe and rate their child's external negative behaviours, but may fail to detect their child's internal difficulties. Indeed, children rated themselves significantly higher on the DIF and DDF subscales when compared to their corresponding parent ratings. In comparison, no significant differences were observed between the child- and parent-rated EOT scores. Both instruments, while producing similar ratings, therefore appear to detect different degrees of cognitive and behavioural difficulties associated with alexithymia. Supporting this, post-hoc analyses (see ESM 8) revealed that the AQC scores of the top 10% scoring children (n = 25) correlated non-significantly with any of the additional measures. In contrast, the corresponding AQC-P scores still correlated significantly negatively with empathic and prosocial behaviour. Thus, the AQC-P appears to

gives a more accurate assessment of the child's associated difficulties when compared to the AQC. Despite this, further work is required to develop guidelines for combining the information from informants. For example, it is unclear whether higher scores based on single or multiple informants should be required to classify a child as at risk of alexithymia and, if the former, which informant's ratings best predict functional impairment. The most significant psychometric weakness identified was with the reliability of the AQC EOT subscale. While the corresponding subscale in the parent-reported version has omega >.70, the child-reported EOT yielded an omega value of .560. Concerns have previously been raised regarding this subscale, with previous studies showing poor factor loadings and unacceptable internal consistency (see Bagby et al., 2020 for review). It may therefore be beneficial for future research to identify the core items of the EOT, and revise or replace the items that have poor reliability.

Limitations

First, the sample was relatively small and recruited opportunistically. Bias may have been introduced as only families particularly motivated to take part in psychological experiments may have participated. Our preliminary results should thus be replicated in larger, more representative samples. Second, the age range of the sample was small and unequally distributed among the ages (i.e., 40% of the sample were ten-year olds, whereas 1% were thirteen-year olds). Consequently, child age measurement invariance could not be assessed. Future studies should assess the measurement invariance of the AQC and AQC-P's three-factor model in child populations with a more evenly distributed age range. The majority of parent reports were completed by the children's mothers. Collating psychometric assessment scores from both the child's mother and father has been recommended (Connell & Goodman, 2002). However, high interrater agreement between mother- and father- reports of child behaviour has been reported (Grietens et al., 2004), suggesting researchers can

adequately rely on one parent to give an accurate evaluation of their child's behaviour. Therefore, it is possible the data collected in the current study was not limited by the large proportion of mother informants. However, future investigations should investigate this by assessing the degree of cross-informant variance in mothers' and fathers' AQC-P scores. Concurrent validity of the AQC-P was not assessed as no additional parent-rated alexithymia was administered. Thus, it would be beneficial for future studies to assess the strength of the relationships between the AQC-P and other child-orientated observer-rated alexithymia scales (e.g., the CAM, Way et al., 2010). Lastly, future studies are required to translate the AQC-P into other languages, as the findings from the current study are only applicable to the English version.

Conclusions

Results support the factorial validity, reliability, convergent validity, gender invariance and cross-informant correlations of the AQC and AQC-P. This suggests that the recently developed AQC-P is a promising measure of parent-reported alexithymia that can be used alongside the AQC to provide a multi-informant measure of child alexithymia.

Auyeung, B., Wheelwright, S., Allison, C., Atkinson, M., Samarawickrema, N., & Baron-Cohen, S. (2009). The children's empathy quotient and systemizing quotient: Sex differences in typical development and in autism spectrum conditions. *Journal of Autism and Developmental Disorders, 39*, 1509. https://doi.org/10.1007/s10803-009-0772-x

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia scale-I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*, 23–32. https://doi.org/10.1016/0022-3999(94)90005-1

Bagby, R. M., Parker, J. D., & Taylor, G. J. (2020). Twenty-five years with the 20-item Toronto Alexithymia Scale. *Journal of Psychosomatic Research, 131,* 109940. https://doi.org/10.1016/j.jpsychores.2020.109940

Birleson, P. (1981). The validity of depressive disorder in childhood and the development of a self-rating scale: a research report. *Journal of Child Psychology and Psychiatry, 22,* 73-88. https://doi.org/10.1111/j.1469-7610.1981.tb00533

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: a Multidisciplinary Journal, 14,* 464-504. https://doi.org/10.1080/10705510701301834

Connell, A. M., & Goodman, S. H. (2002). The association between psychopathology in fathers versus mothers and children's internalizing and externalizing behavior problems: a meta-analysis. *Psychological Bulletin, 128*, 746-773. https://doi.org/10.1037/0033-2909.128.5.746

Costa, A. P., Steffgen, G., & Samson, A. C. (2017). Expressive incoherence and alexithymia in autism spectrum disorder. *Journal of Autism and Developmental Disorders, 47,* 1659–1672. https://doi.org/10.1007/s10803-017-3073-9.

Costa, A. P., Steffgen, G., & Vögele, C. (2019). The role of alexithymia in parent–child interaction and in the emotional ability of children with Autism Spectrum Disorder. *Autism Research, 12,* 458-468. https://doi.org/10.1002/aur.2061

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,* 483–510. https://doi.org/10.1037/0033-2909.131.4.483

Erni, T., Lötscher, K., & Modestin, J. (1997). Two-factor solution of the 20-ltem toronto alexithymia scale confirmed. *Psychopathology*, *30*, 335–340. https://doi.org/10.1159/000285079

Fukunishi, I., Yoshida, H., & Wogan, J. (1998). Development of the Alexithymia Scale for Children: A preliminary study. *Psychological Reports, 82*, 43–49. https://doi.org/10.2466/PR0.82.1.43-49

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry, 38,* 581-586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Grabe, H. J., Spitzer, C., & Freyberger, H. J. (2004). Alexithymia and personality in relation to dimensions of psychopathology. *American Journal of Psychiatry, 161*, 1299–1301. https://doi.org/10.1176/appi.ajp.161.7.1299

Grietens, H., Onghena, P., Prinzie, P., Gadeyne, E., Van Assche, V., Ghesquiere, 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. https://doi.org/10.1023/B:JOBA.0000013661.14995.59

Griffin, C., Lombardo, M. V., & Auyeung, B. (2016). Alexithymia in children with and

without autism spectrum disorders. *Autism Research*, *9*, 773–780.

https://doi.org/10.1002/aur.1569

Grynberg, D., Luminet, O., Corneille, O., Grèzes, J., & Berthoz, S. (2010). Alexithymia in

the interpersonal domain: A general deficit of empathy?. *Personality and Individual

Differences*, *49*, 845-850. https://doi.org/10.1016/j.paid.2010.07.013

Kline, R.B. (2010). *Principles and practice of structural equation modelling* . New York,

New York: Guilford Press.

Levant, R. F., Hall, R. J., Williams, C. M., & Hasan, N. T. (2009). Gender differences in

alexithymia. *Psychology of Men & Masculinity*, *10*, 190–204.

https://doi.org/10.1037/a0015652

Mattila, A. K., Salminen, J. K., Nummi, T., & Joukamaa, M. (2006). Age is strongly

associated with alexithymia in the general population. *Journal of Psychosomatic

Research*, *61*, 629–635. https://doi.org/10.1016/j.jpsychores.2006.04.013

Parker, J. D., Bagby, R. M., & Taylor, G. J. (1991). Alexithymia and depression: distinct or

overlapping constructs?. *Comprehensive Psychiatry*, *3*, 387-394.

https://doi.org/10.1053/comp.2001.23147

Parker, J. D., Eastabrook, J. M., Keefer, K. V., & Wood, L. M. (2010). Can alexithymia be

assessed in adolescents? Psychometric properties of the 20-item Toronto Alexithymia

Scale in younger, middle, and older adolescents. *Psychological Assessment, 22,* 798–

808. https://doi.org/10.1037/a0020256

Pokropek, A., Davidov, E., & Schmidt, P. (2019). A monte carlo simulation study to assess

the appropriateness of traditional and newer approaches to test for measurement

invariance. *Structural Equation Modeling*: *a Multidisciplinary Journal*, *26*, 724–744. https://doi.org/10.1080/10705511.2018.1561293

Revelle, W. (2018). psych: Procedures for Personality and Psychological Research (Version 1.8). Retrieved at: http://CRAN.R-project.org/package=psych

Ricciardi, L., Demartini, B., Fotopoulou, A., & Edwards, M. J. (2015). Alexithymia in neurological disease: a review. *The Journal of Neuropsychiatry and Clinical Neurosciences, 27*, 179-187. https://doi.org/10.1176/appi.neuropsych.14070169

Rieffe, C., Oosterveld, P., & Terwogt, M. M. (2006). An alexithymia questionnaire for children: Factorial and concurrent validation results. *Personality and Individual Differences*, *40*, 123–133. https://doi.org/10.1016/j.paid.2005.05.013

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software, 48,* 1–36.

Salminen, J. K., Saarijärvi, S., Äärelä, E., Toikka, T., & Kauhanen, J. (1999). Prevalence of alexithymia and its association with sociodemographic variables in the general population of Finland. *Journal of Psychosomatic Research, 46*, 75–82. https://doi.org/10.1016/s0022-3999(98)00053-1

Sifneos, P. E. (1973). The prevalence of 'alexithymic' characteristics in psychosomatic patients. *Psychotherapy and Psychosomatics, 22,* 255-262. https://doi.org/10.1159/000286529

Taylor, G. J., Bagby, R. M., & Parker, J. D. (1999). *Disorders of affect regulation: Alexithymia in medical and psychiatric illness.* Cambridge: Cambridge University Press.

Taylor, G. J., Bagby, R. M., & Parker, J. D. (2003). The 20-Item Toronto Alexithymia Scale: IV. Reliability and factorial validity in different languages and cultures. *Journal of*

Psychosomatic Research, *55*, 277-283. https://doi.org/10.1016/S0022-3999(02)006018

Way, I. F., Applegate, B., Cai, X., Franck, L. K., Black-Pond, C., Yelsma, P., Roberts, E.,

Hyter, Y., & Muliett, M. (2010). Children's Alexithymia Measure (CAM): a new

instrument for screening difficulties with emotional expression. *Journal of Child &*

Adolescent Trauma, 3, 303–318. http://dx.doi.org/10.1080/19361521.2011.609772

Open Science

Code: The R code and SPSS syntax for all analyses (ESM 1) is available at

http://dx.doi.org/10.23668/psycharchives.4406.

Electronic Supplemental Materials: All supplemental materials (ESM2 – ESM 8) are

available at http://dx.doi.org/10.23668/psycharchives.4407.

ESM 2: Analysis output

ESM 3: Additional descriptive analyses

ESM 4: Measurement invariance of AQC-P results

ESM 5: Partial measurement invariance of AQC-P results

ESM 6: Measurement invariance of AQC results

ESM 7: Partial measurement invariance of AQC results

ESM 8: Posthoc analyses

Data: Data and materials will be made be available upon request.

Preregistration of Studies and Analysis Plans: This study was not preregistered and

without an analysis plan.

Table 1.

Preliminary analysis.

Measure	Mean	SD	α	Range	Skewness	Kurtosis
AQC- P	32.78	6.11	.822	20 – 51	.634	.025
AQC	35.51	5.75	.737	22 – 51	.136	-.591

Note. AQC: Alexithymia Questionnaire for Children; AQC-P: Alexithymia Questionnaire for Children – Parent.

Table 2.

Confirmatory factor analyses of the one-, two- and three-factor solutions on the AQC-P and AQC.

		χ^2	Df	p	RMSEA	SRMR	CFI
AQC-P	Model 1: 3 Factor	141.76	167	.922	.052	.054	.902
	Model 2: 2 Factor	375.78	169	.000	.070	.075	.821
	Model 3: 1 Factor	523.93	170	.000	.091	.091	.693
AQC	Model 1: 3 Factor	235.53	167	.000	.041	.067	.963
	Model 2: 2 Factor	316.17	169	.000	.059	.079	.830
	Model 3: 1 Factor	341.63	170	.000	.064	.069	.801

Note. RMSEA: root-mean-square error of approximation; SRMR: standardized root-mean square residual; CFI: comparative fit index.

Table 3.

Correlations and Fisher z-transformations between the AQC, AQC-P and external measure scores.

Measure	Subscale	AQC	AQC-P	$Z_{observed}$
DSRS		.628***	.291***	4.87***
EQ-C		-.281***	-.490***	2.75**
SDQ	*Prosocial Behaviour*	-.239***	-.487***	3.21**
	Inattention-Hyperactivity	.204**	.382***	2.17*
	Emotional Symptoms	.193**	.296***	.690
	Conduct Problems	.200**	.372***	2.09*
	Peer Problems	.066	.227***	.610

Note. **p <01, ***p <.001. DSRS: Depression Self-Rating Scale, EQ-C: Empathy Quotient – Child, SDQ: Strengths and Difficulties Questionnaire.