Social emotions and social functioning in Chinese deaf and hard-of-hearing and hearing preschoolers

Li, Z.

Chapter 4

Validation of the Chinese Empathy Questionnaire for Pre-School Children

Li, Z., Li, B., Tsou, Y. T., & Rieffe, C. (2024). Validation of the Chinese Empathy Questionnaire for Pre-school Children. *Early Education and Development,* 1-16. https://doi.org/10.1080/10409289.2024.2372938

Abstract

Research Findings: Empathy plays an important role in children's early social-emotional development. However, there are not many instruments available for measuring the different aspects/levels of empathy in preschool children. Moreover, few studies examined the validity of the existing measures for empathy in Chinese children. The present study tested the Chinese version of the Empathy Questionnaire (EmQue) on a group of Chinese preschool children (N = 182) aged from 2 to 6 years. The Chinese EmQue is a parent report translated from Dutch that assesses the behavioral manifestations of three distinct levels of empathy: emotion contagion, attention to others' feelings, and prosocial behaviors among preschool children. Confirmatory factor analysis showed satisfactory goodness-of-fit indexes for a three-level structure with 19 loading items. The results of concurrent validity showed that the three levels of empathy were functioning differently in early social-emotional development. *Practice or Policy:* We could conclude that the 19-item Chinese EmQue is a valid, reliable instrument for measuring empathy in 2- to 6-year-old children in the Chinese context.

Key Words Empathy; Preschool age; Validation; Chinese children

INTRODUCTION

Empathy serves an important function in navigating daily social life, as it plays a key role in guiding interpersonal interactions, facilitating in-group behaviors, strengthening social bonds. Importantly, empathy, i.e., feeling and caring for another person's emotional response in a given situation, motivates people to care for one another and to focus on another person's needs, thereby temporarily overcoming one's own idiosyncratic needs and desires (De Waal, 2008; Decety & Jackson, 2006). The earliest signs of empathy can be already witnessed in infants, emphasizing that the capacity for empathy is innate and can be observed in children across different cultures (De Waal, 2008; Tong et al., 2012).

However, to date, research measuring empathy is mostly based on Western samples, and with participants of older ages (e.g., late childhood, adolescence, and adulthood), while this topic has rarely been discussed among young children from East-Asian countries. Moreover, current research conducted in East-Asian countries mostly operationalized empathy as a unidimensional construct, whereas empathy is perceived as a complex construct consisting of several distinct levels (i.e. emotional contagion, attention to others' emotions, understanding others' emotions, and prosocial behaviors; e.g., Hoffman, 1987; Rieffe et al, 2010).

It is important to identify these levels separately in young Chinese children, as these levels may mature differently through development (see explanation below), and the development of these distinct levels of empathy has been shown to relate to various

domains of social-emotional functioning, such as better developed social competence, and lower risks of developing externalizing behaviors (e.g., Neumann, et al., 2016; Noten et al., 2019), according to Western studies. Recent studies among Chinese preschool children showed similar results of empathy (as a unidimensional construct): higher levels were associated positively with children's sociability, emotion regulation, whereas negatively with their peer problems, internalizing or externalizing behavioral symptoms (e.g. Lin et al., 2023; Qian et al., 2021; Wu et al., 2018). Yet, how separate levels of empathy are related to social-emotional functioning in Chinese preschoolers remains unknown.

To fill these gaps, the present study aims to translate and validate a parent report that was specially designed for assessing three different levels of empathy in preschool children, the Empathy Questionnaire (EmQue, Rieffe et al., 2010; Dutch in origin), in a sample of Chinese preschool children aged 2 to 6 years. This questionnaire has already been translated and validated in four other languages, and the validation studies confirmed its factor structure and showed good psychometric properties for the separate (sub-)scales (Grazzani et al., 2017; Takamatsu et al., 2021; Lucas-Monlina et al., 2018; Da Silva, et al., 2022). The theoretical as well as practical meanings of validating the EmQue will be discussed below.

Empathy in Preschool Children

Hoffman (1987) proposed a multilevel model of how empathy develops during early childhood, where each of the levels builds upon the former. "Emotion contagion",

referred to also as "affective empathy" (Rieffe et al., 2010), is the first level, representing the extent to which an individual can be emotionally affected via perceiving others' emotional expressions or behaviors (e.g., Noten et al., 2019; Pouw et al., 2013). For instance, when one infant bursts out crying, other nearby infants may follow spontaneously. Emotion contagion was described by De Waal (2008) as an innate, hard-wired, core mechanism of empathy. It is supposed to be the trigger for "feeling what the other feels", to create an emotional bond between two individuals, stimulating interpersonal closeness, and thus feeding into social competence over time (Eisenberg et al., 2010). Yet, excessive levels of emotion contagion may lead to more internalizing behaviors, such as anxiety and depression (Geng et al., 2012; Rieffe et al., 2010; Yan, et al., 2021), as empathizing with others' negative emotions can also be overwhelming and disturbing for young children.

The second level of empathy, "attention to others' feelings", emerges at around the second year. Hoffman (1987) argued that when toddlers have acquired a basic sense of self-awareness, they undergo the "separation-individuation" process through which they identify their core self as separated from others. The awareness that the arousal they experience when they empathize with others is but a reflection of the arousal in someone else, helps them to be less overwhelmed by the intensity caused by others' emotions and enables them to pay more attention to interpersonal interactions in social life (Decety, 2011). Besides, children need to have a basic level of self-control/emotion regulation to temporarily overrule their instinctive, idiosyncratic desires to focus on others' internal states and to serve others' needs (Rodríguez et al., 2020). The change

of perspective to locate the source of emotional arousal is such a developmental milestone: it marks the shift of attention from the self and inner feelings to the social world, allowing further social learning and socialization processes to take place.

The shift in attention from self-focused to other-focused is an important step towards acting prosocially to relieve other's distress (i.e., "prosocial behaviors"), which is the third level of empathy. To initiate prosocial behaviors, first, children need to attend to the person in distress rather than being immersed in their own emotional arousal. Second, understanding another person's emotional states and needs is crucial for initiating prosocial behaviors. As argued by Hoffman (1987), although toddlers have the motivation to subside the impact of negative emotions on others, they often confuse their own needs with other people's needs, and thus, for example, offering candies to a crying adult. As children grow older, along with enhanced emotion regulation and perspective-taking abilities, children's prosocial actions increase in both quality and quantity (Flook et al., 2019, Li et al., 2020; Tsou et al., 2021).

Empathy Measurement in China

To the best of our knowledge, there is currently no parent questionnaire designed to measure the multi-component nature of empathy in Chinese preschool-aged children. Although some widely used self-report questionnaires have been validated to assess empathy in Chinese school-aged children, including the "Interpersonal Reactivity Index" (IRI, Siu & Shek, 2005), the "Jefferson Scale of Empathy" (JSE, Yi et al., 2020; Huang et al., 2019), the "Basic Empathy Scale" (BES, Geng et al., 2012),

and the "Empathy Quotient" (EQ, Guan, et al., 2012; Zhao et al., 2018), the inclusion of a scale focused on understanding others' emotions (perspective taking) makes them unsuitable for toddlers and preschoolers, who cannot yet acknowledge others' perspectives due to the different stages in their Theory-of-Mind development (Broekhof et al., 2015).

To date, parent reports used in Chinese preschoolers measured empathy as a one-dimensional construct (e.g., Lin et al., 2023; Qian et al., 2021; Wu et al., 2018). In other studies, observational tasks were employed (e.g., Huang et al., 2017; Roberts & Strayer, 1996). Both methodologies cannot capture the nuanced developmental characteristics of distinct empathy components in preschoolers. Drawing on prior research conducted with Western children, it has been observed that emotion contagion tends to decrease over time in early childhood, while attention to another's emotions and prosocial behaviors tends to increase with age (e.g., Rieffe et al., 2010; Tsou et al., 2021). By validating the EmQue, a parent report that gauges the three essential components of empathy in toddlers and preschoolers, we provide a valuable instrument for understanding the intricate development of empathy among Chinese preschool-aged children.

Furthermore, our understanding of empathy in preschool years has stemmed primarily from observations of Western children. Yet, cross-cultural studies on empathy in adults have revealed distinctions in the experience and expression of empathy between Western and East-Asian individuals: East-Asian individuals experience higher levels of personal distress when witnessing someone in distress, as compared to their

Western counterparts (Atkins et al., 2016). This variance can be attributed to the cultural emphasis on group harmony in many East-Asian societies, where expressions of negative emotions are viewed as disruptive or socially inappropriate – thus East-Asian children might have fewer opportunities to practice coping with negative emotions, in responding to others' negative emotions or in regulating their own (Matsumoto et al., 2008). Consequently, East-Asian individuals can be more alerted to and be more heavily affected by others' expressions of negative emotions (Markus & Kitayama, 1991).

In fact, at the social level, individuals in collectivistic cultures (common in East-Asian societies) might exhibit a higher frequency/intensity of attention to interpersonal relationships, which increases distress levels; whereas those in individualistic cultures (common in Western societies) would focus more on their feelings and needs, lowering the risk of developing emotional disorders (Hofstede, 1980; Markus & Kitayama, 1991). Validating the EmQue in Chinese preschoolers may provide a useful tool for conducting cross-cultural investigations, and contribute to new insights on the cultural variance of empathy development.

Present Study

The present study aimed to validate a Chinese version of the Empathy Questionnaire (EmQue; Rieffe et al., 2010, Appendix Table A) in a sample of Chinese preschool children. The EmQue is a Dutch-originated parent report comprising three scales: "Emotion Contagion", "Attention to Others' Feelings", and "Prosocial Behaviors" that

correspond to the first three levels of Hoffman's (1987) model. It has been used in cross-sectional and longitudinal research on social-emotional functioning and development, and in preschoolers who are autistic or deaf and hard-of-hearing (Li et al., 2023; Netten et al., 2015). Thus far, the EmQue has been translated and validated in four other languages in Italian, Japanese, Spanish, and Portuguese preschoolers (respectively, Grazzani et al., 2016; Takamatsu et al., 2021; Lucas-Monlina et al., 2018; Da Silva, et al., 2022). The studies demonstrated robust psychometric properties of the three levels of empathy across different languages and cultures, although slightly different numbers of items were removed from the original version (20 items) to form a shorter version of the EmQue (the 13-item version was validated in Italian, Japanese, and Spanish; the 15-item version was validated in Portuguese; see Appendix Table A for an overview). Validating a Chinese EmQue would further our understanding of the nuances in empathy development and facilitate cross-cultural investigation.

First, we assessed the construct validity by examining the model fit of the 3-factor structure in the confirmatory factor analysis (CFA) (i.e., Emotion Contagion, Attention to Others' Feelings, Prosocial Behavior). Additionally, we also tested the fit of the short version EmQue (13 items) previously validated in different languages (i.e., Italian,, Japanese, and Spanish; Grazzani et al., 2017; Takamatsu et al., 2021; Lucas-Molina et al., 2018), and compared it to our final model derived from the CFA. Second, we examined the internal consistencies of the three scales. Third, we investigated concurrent validity by examining the correlations between the EmQue scales and the variables of social-emotional functioning (internalizing behaviors, externalizing

behaviors, and social competence). Based on prior studies, we expected Emotion Contagion to correlate positively to internalizing behaviors and social competence (Noten et al., 2019; Pouw, et al, 2013; Yan et al., 2021), while negatively to externalizing behaviors (Tampke, et al., 2019; Yan et al., 2021). We expected Attention to Others' Feelings and Prosocial Behaviors to correlate positively to social competence, whereas negatively to externalizing behaviors (Bandstra et al., 2011; Da Silva et al., 2022). Prosocial Behaviors were expected to also correlate negatively to internalizing behaviors (Donohue et al., 2020; Huber et al., 2019; Salerni & Caprin, 2022).

METHOD

Participants

This study is part of a larger-scaled longitudinal research project on the moral development of preschoolers in China. 182 preschoolers (age range: from 24 to 72 months, mean age = 52.3 months, SD_{age} = 12.2 months) along with their caregivers participated in the research. This validation study is part of a larger project containing two validation studies, and to date, one study has been published using data from this sample. Further information on the demographic characteristics of our sample is provided in that publication (Li et al., 2023).

Participants were recruited from two public kindergartens in Jiangsu province, China. Each kindergarten had three classes from three different grades, thus in total six classes. Both kindergartens and their six classes agreed to participate in this research. First, we explained to the teachers the inclusion criterion of participants: participants

should not have any clinical diagnoses, such as autism spectrum disorder or attention-deficit hyperactivity disorder. Furthermore, we explained in detail the aim of our study, the execution plan, the privacy policy, and the voluntary nature of participation. Before the data collection, caregivers received information about the project and an informed consent form on which they could indicate if they approved for their child to participate. After our researchers verified the informed consent, the questionnaires were distributed by the teachers to the children included in the study, delivered to the caregivers, and brought back upon completion. Formal approval of this study was acquired from both the kindergartens and the Ethical Committee of Leiden University.

Measures

Sum scores for all scales involve the mean score based on the total number of items per scale. Internal consistencies of the scales were adequate to good (Table 1).

Table 1. The Descriptive Statistics, Internal consistencies of the study variables

Factors	Number of items	Mean (SD)	Cronbach alpha	Inter-item correlation	Composite reliability
EmQue Scales					
Emotional Contagion	6	1.49 (.41)	.70	.28	.72
Attention to Others' Feelings	7	2.34 (.36)	.67	.23	.63
Prosocial Behaviors	6	2.14 (.46)	.76	.35	.83
Social-Emotional Indexes					
Internalizing Behaviors	25	1.46 (.27)	.73	.27	.94
Externalizing Behaviors	18	1.31 (.25)	.87	.31	.83
Social Competence	10	2.37 (.38)	.69	.22	.76

The Empathy Questionnaire (EmQue, Appendix Table A). EmQue is a parental report

that aims at assessing empathetic behaviors in young children. The original EmQue was designed by Rieffe et al., (2010) in Dutch language consisting of 20 items. According to the validation of Rieffe et al., (2010), one item (item 19, Appendix Table A) was removed, leading to a 19-item Dutch version of EmQue. The 20-item version was also validated in Italian (Grazzani et al., 2017), Japanese (Takamatsu et al., 2021), and Spanish (Lucas-Monlina et al., 2018) and that have formed a new, 13-item version. It was also validated in Portuguese recently suggesting a new 15-item version (Da Silva, et al., 2022). Notably, all of the existing versions of EmQue had the same 3-factor structure, measuring Emotional Contagion, Attention to Others' Feelings, and Prosocial Behaviors for empathy. In the scales, respondents were asked to rate the degree to which each item represents their child's behaviors over the past two months on a 3-point scale (0 = never, 1 = sometimes, 2 = often). Higher scores indicate a higher tendency for their children to perform the behaviors described by the items. Respondents were recommended to answer to all items, even if the items did not apply to them.

Internalizing and Externalizing Behaviors. The Early Childhood Inventory 4th edition parent checklist (ECI-4; Sprafkin et al., 2002) was deployed to evaluate the severity of internalizing and externalizing behavioral problems. This questionnaire consists of 9 scales and 108 items, screening for 15 distinct emotional/behavioral disorders of preschoolers. The caregivers were asked to rate how each item represented their child's behaviors in the last two months on a 4-point scale (0 = never, 1 = sometimes, 2 = often, 3 = very often), with higher scores indicating more severe behavioral problems.

To embody the purpose of this study, we merged scales within the domain of

internalizing and externalizing symptoms to measure the two broadband categories of behavioral problems (Ketelaar et al., 2017; Li et al., 2020). For the assessment of "Internalizing Behaviors", we combined four scales: "Major Depressive Disorder" (10 items, e.g., "Always feeling tired for no particular reason", "Being too shy in front of their peers"), "Separation Anxiety" scale (8 items, e.g., "Having nightmares about being separated from parents", "Cannot fall asleep without parents staying around"), "Social Phobia" (3 items, e.g., "Wet the bed at night"), and "Generalized Anxiety" (4 items, e.g., "Cannot get rid of worrying thoughts"). A higher score indicates more internalizing symptoms.

For the assessment of "Externalizing Behaviors", we combined the two scales "Conduct Disorder" (10 items, e.g., "steal from others", "physically hurting others"), and "Oppositional Defiant Disorder" (8 items, e.g., "arguing with adults", "blame others when making mistakes"). A higher score indicates more externalizing symptoms.

Social Competence. The Chinese version of the Strengths and Difficulties Questionnaire (Chinese SDQ; Lai et al., 2010) was used to assess social competence. The questionnaire consists of five scales (each scale has five items), evaluating five aspects of children's social-emotional functioning. The caregivers were asked to rate to what extent their child manifested each behavior/symptom in the past two months on a 3-point scale (0 = not true, 1= somewhat true, 2 = certainly true). Following Ketelaar et al. (2017), we merged two scales to assess children's social competence: the "Peer Relation" scale (e.g., "Rather solitary, tends to play alone", "Picked on or bullied by other children"), and the "Prosocial Behavior" scale (e.g., "Kind to younger children",

"Helpful if someone is hurt, upset or feeling ill"). Negatively formulated items were reversedly coded. A higher score indicates a higher level of social competence.

Translation and Task Procedures

As there were no Chinese versions of the EmQue and the ECI-4, a back-translation procedure was performed to translate the two questionnaires (Brislin et al., 1973). First, the EmQue and ECI-4 were translated from English to Chinese by two researchers in our lab who are fluent in English/Chinese. Next, other researchers translated the two questionnaires from Chinese back to English. The back-translations were checked for language consistency as compared to the original questionnaires. Disagreement was resolved through discussion. Finally, the research team checked and approved the translations.

Statistical analyses

The construct validity of the Chinese EmQue was evaluated via confirmatory factor analysis[1] (CFA) (Jackson et al., 2009). We started with the 20-item Dutch EmQue (Rieffe, et al., 2010). Several models were sequentially tested, on the three-factor construct (Emotional Contagion, Attention to Others' Feelings, Prosocial Behaviors). Given the ordinal nature of three-point scales, we used the Weighted Least-Squares

[1] An exploratory factor analysis (EFA) was initially conducted in a pilot trial, and the 3-factor model was shown as appropriate according to the eigenvalue and scree-plot. Given that our aim was to confirm the 3-factor model in a Chinese sample, and that conducting EFA and CFA on the same sample could increase the risk of overfitting (Fokkema & Greiff, 2017; Whittaker & Worthington, 2016), in this study we only conducted the CFA as the formal analysis.

Means and Variance adjusted (WLSMV), which was considered a more adaptable estimator (Bandalos, 2014) to examine construct validity and measurement invariance. A set of absolute and relative fit indices were used to evaluate if our model fit was sufficient: the normed chi-square ($\chi 2$) < 3.0 and $\chi 2/df$ < 2.0 (Brown, 2014), the comparative fit index (CFI) > .90, the Root-Mean-Square Error of Approximation (RMSEA) < .08, the standardized root mean square residual (SRMR) < .08 (Hu & Bentler, 1999), on condition with the null model RMSEA > .158 (Kenny et al., 2015). Besides, the modification indices (MI) > 10 were used as the criterion to evaluate if any item was loaded on a non-target scale (Whittaker, 2012).

For internal consistencies, we assessed Cronbach's alphas, inter-item correlations, and composite reliabilities. A Cronbach's alpha value above .70 was deemed adequate (Ponterotto & Ruckdeschel, 2007). Inter-item correlations ranged [.30, .50] indicate consistency of the items within one scale. A scale could be too homogenous when this value was higher than .50. For composite reliabilities, a value higher than .70 was deemed sufficient; however, a value higher than .95 may indicate high homogeneity of items within a scale (Bacon et al., 1995).

Moreover, we assessed the concurrent validity of the three EmQue scales by checking their correlations with three social-emotional indices: Externalizing Behaviors, Internalizing Behaviors, and Social Competence. Given the ordinal nature of our data, partial Spearman's rank-order correlations were preferred over Pearson's correlations, and the age factor was controlled for. Also, when the EmQue scales were inter-correlated, they were controlled for too. Bonferroni corrections were used as our

adjustment method in the multiple testing.

Additionally, we assessed measurement invariance of the EmQue scales across gender and age groups (as 54 months was the median age, we compared two groups: < 54 months vs. ≥ 54 months) through multigroup CFAs. Following Brown's (2014) standard procedure, three measurement invariance hypotheses: configural, metric, and scalar, were tested sequentially. Equivalence of item factor loadings and intercepts were examined across the target groups. Partial invariance was achieved by freeing invariant items on the condition that full metric or scalar invariance was not achieved (Byrne et al., 1989). The criteria for the invariance were: the change value of CFI (ΔCFI) < .01; the change value of RMSEA (ΔRMSEA) < .015; and the change value of SRMR (ΔSRMR) < .030 (Chen, 2007). Items can be freed based on the univariate MI and the Lagrange multiplier test. Group differences were checked if > 50% of the items of one scale were invariant (Martín-Puga et al., 2022).

The CFAs and measurement invariance tests were carried out using the *Lavaan* package (Rosseel, 2012) on the platform of R version 4.0.5 (R Core Team, 2022). The calculations of Cronbach's alphas, inter-item correlations, and partial Spearman's correlations were performed in SPSS 24.0 (IBM, 2016). Composite reliabilities were computed by using the values of factor loadings acquired in the CFAs.

Missing Data Analysis

Out of the 186 responses we received, four had missing values larger than 10% and thus were removed from further analyses. The remaining responses had less than 0.5%

missing values and were included in the final analyses. The MCAR test suggested the missing pattern was random (Little, 2013). Listwise deletion was used to remove the missing values.

RESULTS

Construct validity

Table 2 shows the fit indices of the models tested successively. Model 1 was the original 20-item Dutch model (Rieffe et al., 2010), the starting point of the analysis. This baseline model showed an acceptable model fit. Yet, item 19 had a low factor loading of .18 (< .30) on its target scale (Emotion Contagion) and thus was removed from the analysis, leading to Model 2. The fit indices of Model 2 revealed an improvement: χ^2 = 167.37, df =107, CFI = .926, TLI = .910, RMSEA = .53, while the null model RMSEA = .173 (>.158). Furthermore, we examined the factor loadings, MI, SEPC, and inter-factor correlations, and no significant unfitness was observed. Hence, Model 2 was accepted as the final model. Additionally, we tested also the 13-item version, as suggested by the Italian, Japanese, and Spanish versions (Grazzani et al., 2017; Takamatsu et al., 2021; Lucas-Molina et al., 2018). The results showed an adequate model fit (Table 2).

Table 2. Fit indices for the model testing of the EmQue

Model	χ2	df	RMSEA [90% CI]	Null RMSEA	CFI	TLI	SRMR
Model1	248.82	167	.052 [.038, .065]	.161	.921	.910	.083
Model2	224.81	149	.053 [.038, .067]	.173	.926	.915	.083
13-item	79.83	62	.040 [.013, .063]	.205	.970	.962	.070

Notes: RMSEA [90% CI] = root-mean-square error of approximation [lower, upper limits, 90% Confidence Interval]; Null RMSEA = the baseline model RMSEA; CFI = comparative fit index; TFI = Tucker-Lewis Index; SRMR = standardized root mean square residual.
Model 1 = original design of the questionnaire which consists of 20 items (our starting point);
Model 2 = the final model, item 19 (from Emotional Contagion) was deleted based on Model 1;
Model 3 = the Italian model of 13 items tested additionally as a comparison to the final 19-item model.

Measurement Invariance Across Gender

Based on our final model, first, the configural model showed an adequate fit (see Table 3). The metric invariance examination revealed a significant change in the fit indices (ΔCFI= −.030; ΔRMSEA = .022; ΔSRMR = .007). Partial metric invariance was thus tested by freeing the equivalence constraint of item 14 (from Prosocial Behavior) and item 16 (from Emotional Contagion) and was found tenable (ΔCFI = -.009; ΔRMSEA = .010; ΔSRMR = .004), so the item loadings except for items 14 and 16 were invariant across gender groups. Later, testing scalar invariance revealed an insignificant change of the fit indices (ΔCFI = .004; ΔRMSEA = -.005; ΔSRMR = .002), confirming the scalar invariance hypothesis of equivalent intercepts. Given ≥ 50% of the items of each scale being invariant, the EmQue scores were comparable across genders. Independent t-tests revealed no gender differences in the three EmQue factors, ts < 1.42, ps > .157.

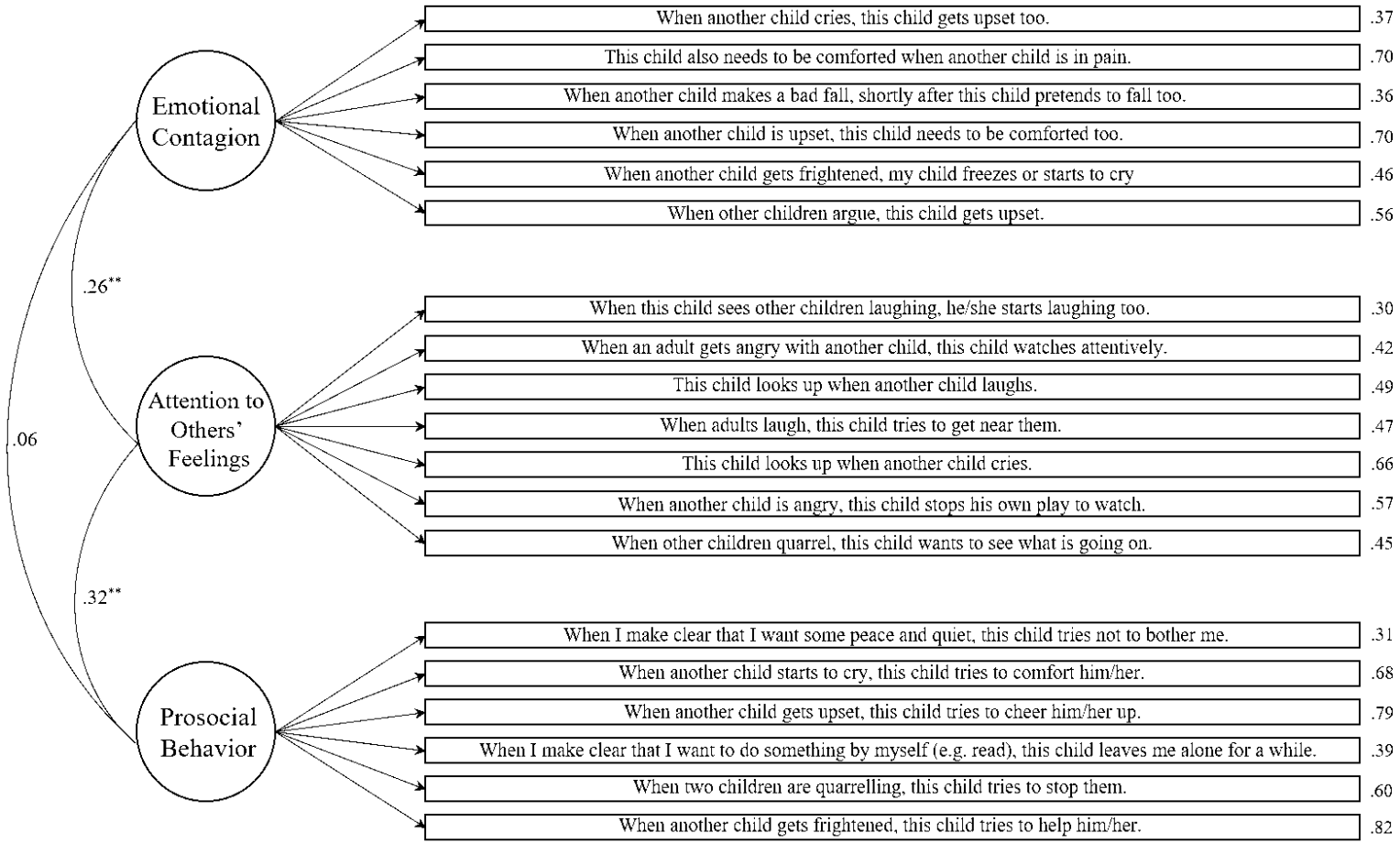

Figure 1. CFA of the three-factor structure of the Chinese EmQue

Table 3. Fit Indices of the Invariance Examinations Across Gender, Age Groups (half of participants <54 months; the other half ⩾54 month).

	χ2	df	CFI	RMSEA [90% CI]	SRMR	ΔCFI	ΔRMSEA	ΔSRMR
Gender								
Configural	303.97	298	.994	.015 [0, .044]	.091			
Metric	351.46	314	.964	.037 [0, .056]	.098	-.030	.022	.007
Partial Metric	328.89	312	.985	.025 [0, .048]	.095	-.009	.010	.004
Scalar	339.26	328	.989	.020 [0, .045]	.097	.004	-.005	.002
Age group								
Configural	343.50	298	.956	.041 [.012, .060]	.096			
Metric	389.03	314	.928	.052 [.032, .068]	.102	-.028	.011	.006
Partial Metric	367.52	311	.946	.045 [.021, .063]	.099	-.010	-.007	.003
Scalar	384.03	327	.945	.044 [.021, .061]	.101	-.001	-.001	.002

Note: CFI: comparative fit index; RMSEA: root mean square error of approximation; CI: confidence interval; SRMR: standardized root mean square residual. N=182. *p<0.05

[a] An equality constraint on the factor loadings of Item 14 and item 16 was freed from the model;

[b] Equality constraints on the factor loadings of Items 12, item 16 and 20 were freed from the model.

Measurement Invariance Across Age

We first divided our sample (by medium age) into two equal groups: one < 54 months, and the other ≥ 54 months. The configural model showed a good fit (see Table 3). The test of the metric invariance revealed a significant change in the fit indices (ΔCFI = −.028; ΔRMSEA = .011; ΔSRMR = .006). Thus, partial metric invariance was tested by freeing the equivalence constraint on items 12 and 20 (both from Attention to Others' Feelings) and item 16 (from Emotional Contagion), then a nonsignificant change was observed (ΔCFI = -.010; ΔRMSEA = -.007; ΔSRMR = .003), indicating that the factor loadings besides items 12, 16 and 20 were invariant across age groups. Next, testing scalar invariance revealed a nonsignificant change (ΔCFI = -.001; ΔRMSEA =-.001; ΔSRMR = .002), indicating equivalent intercepts across age groups. Independent t-

tests showed an age difference for Prosocial Behavior: Prosocial $_{elder}$ = 2.28 > Prosocial $_{younger}$ = 2.00, t(179) = 4.807, p < .001, but not for Emotional Contagion or Attention to Others' Feelings, ts < -.50, p > .453.

Internal Consistency

Internal consistencies were from acceptable to good. For the three EmQue scales, Cronbach Alpha's varied from 0.67 to 0.76; inter-item correlations varied from 0.23 to 0.35; composite reliabilities varied from 0.63 to 0.83, respectively (Table 1).

Table 4. Spearman's correlations between the MEQ scales and the concurrent indices

	Concurrent Indices: Social-Emotional Functioning		
	Internalizing Behaviors	Externalizing Behaviors	Social Competence
Emotion Contagion	.20**	.01	-.39***
Attention to Others'Feelings	.29***	.09	.22**
Prosocial Behaviors	-.25**	-.17***	.39***

Note. The significance level is adjusted with Bonferroni correction to $p \leq \alpha/3 = .017$. * p (one-tailed) ≤ .017. ** p (one-tailed) ≤ .010. *** p (one-tailed) ≤ .001.

Concurrent Validity

Table 4 presents the Spearman correlation coefficients of the EmQue scales with the indices of social-emotional functioning, age, and gender. After controlling for age and Attention to Others' Feelings, higher levels of Emotion Contagion were correlated with more Internalizing Behaviors and lower Social Competence; more Prosocial Behaviors were correlated with fewer Internalizing Behaviors and Externalizing Behaviors, and higher Social Competence. When controlling for age, Emotional Contagion, and Prosocial Behaviors, Attention to Others' Feelings were positively correlated to Internalizing Behaviors and Social Competence.

DISCUSSION

The present validation study provided support for a 19-item Chinese version of EmQue, thus replicating the results of the original Dutch EmQue (Rieffe, et al., 2010). The results of psychometric properties suggested a good construct validity, distinguishing between the three levels of empathy, i.e., Emotion Contagion, Attention to Others' Feelings, and Prosocial Behaviors in a sample of Chinese preschool children. Consistent with the Dutch EmQue, one item was removed from the 20-item version due to a low factor loading on its intended scale (item 19, "When another child cries, this child looks away"). The internal consistency within each scale was adequate. Also, partial measurement invariance was achieved when comparing across gender or age groups, meaning that the majority of the items of each scale were equivalent for young children of different ages and genders. Hence, the 19-item version of EmQue is reliable for measuring (multi-levels of) empathy in young boys and girls aged from 2 to 6 years old.

In addition, we assessed the 13-item model used in the Italian, Japanese, and Spanish versions of EmQue (Grazzani et al., 2017; Takamatsu et al., 2021; Lucas-Monlina et al., 2018). A comparable, yet not significantly higher, level of fit indices than the 19-item model was revealed. Thus in our case, we considered that the 19-item and 13-item models fit equally adequate in terms of the construct validity, while the 19-item version has the advantage of its breadth of valid items, and therefore is regarded as more informative, comprehensive, and capable of offering a good view for the topic.

Regarding concurrent validity, our hypotheses were partly confirmed by the results: as expected, the scale of Emotion Contagion was related to more internalizing behaviors; the scale of Attention to Others' Feelings was related to better social competence; and the scale of Prosocial Behaviors was related to fewer internalizing, externalizing behaviors, and better social competence. Yet, unexpected relations were also found for these scales: Emotion Contagion was related to lower social competence and Attention to Others' Feelings was related to more internalizing behaviors. Besides, the expected negative associations for Contagion and Attention to Others' Feelings with externalizing behaviors were not confirmed. Below we discuss the outcomes in greater detail.

As expected, emotion contagion was associated with more internalizing behaviors in Chinese preschool children (Rieffe et al., 2010). However, unexpectedly, it was associated also with lower levels of social competence. This finding may be explained by the cultural norm in East-Asian societies, which stresses that the expression of negative emotions is unwelcome (Tsai & Levenson, 1997). For Chinese children,, displaying personal distress and negative emotions in social settings may induce discomfort in others and disrupt interpersonal interactions, thus children manifesting too much emotion contagion can be perceived as being socially less competent and emotionally immature (Ip et al, 2021; Markus & Kitayama, 1991). Being overwhelmed by one's own arousal further hinders children from reacting empathically and socially adaptively toward others, which in the long run might compromise their development of social competence (Hamaidi et al., 2021; MacCormack et al., 2019;

Spinrad et al., 2006). In contrast, many Western cultures allow for more open expression of feelings, including negative emotions (Tsai & Levenson, 1997). In such a cultural context, expressing personal distress in an empathy-provoking situation may be viewed as showing an understanding of another person's plight, thereby enhancing social relationships.

Attention to others' feelings was related to better social competence, as expected. As suggested by Hoffman (1987), children switching their focus of attention from the self to others facilitates children's socialization. However, unexpectedly, attention to others' feelings was also associated with more internalizing behaviors. Possibly, cultural variance should be taken into account: for Western children, attending to others' emotions might yield social benefits by enhancing the understanding of someone else's inner mental state. Yet, this may have relatively little impact on their own inner feelings, given the prevalent independent self-image in Western cultures. In contrast, in an interdependent cultural context, such as in the Chinese context, where individuals prioritize collective values and group harmony, children may be more alerted by others' distress, which could bring social benefits (e.g., strengthens social bonds), but simultaneously impose costs to the self (Tsai et al., 2006). Prior research suggested that elevated attention to others' negative emotions may increase emotional arousal in preschool children (Tsotsi et al., 2021). Attending to another person's distress thus could induce high emotional arousal in Chinese children, potentially leading to anxiety and depressive mood (Atkins et al., 2016).

Whereas emotion contagion and attention to another's emotion showed a negative association with externalizing behaviors in Western children, such associations were not observed in Chinese preschoolers. This discrepancy might stem from the different coping mechanisms between Western and East-Asian children: even when they experience similar levels of emotional arousal in empathy-provoking situations, the resulting action tendencies can manifest in distinct ways. In the absence of good emotion regulation and a matured form of emotional understanding, Western preschool children tend to externalize their distress and frustration by venting out negative emotions, whereas Chinese preschool children internalize negative feelings and express their distress and frustration through manifestations of anxiety and depression (Atkins et al., 2016). This inclination is probably shaped by the internalization of cultural values of collectivism, as prioritizing group harmony is emphasized over personal expressions (Hofstede, 1980).

Overall, our study supports that the tri-dimensional construct of empathy exists cross-culturally, yet our findings also indicated nuanced distinctions in the specific functions of each empathy level, which may be shaped by the cultural contexts in which children are socialized. This underscores the need for validated instruments to measure empathy in children from diverse cultural backgrounds, which would enable researchers to further validate these assumptions made in this study.

Limitations and Future Directions

Validating a parent report to assess empathy in Chinese preschoolers is an important first step to studying empathy development in cross-cultural contexts. Yet limitations should be addressed. First, it is noteworthy that our sample size was relatively small, which could cause inaccuracy in our estimations of the psychometric properties of EmQue. For example, although our data distribution was close to symmetric (skewness was between -1 to 1), the standard errors can still be biased - approximately 10% larger than the standard error estimates of large samples (Bandalos, 2014; DiStefano & Morgan, 2014). The smaller sample size also prevented us from splitting the sample and respectively conducting an EFA and a CFA. Hence we highly recommend future validation studies to use larger-size (N > 500) samples to minimize the biases in estimations. With larger sample sizes, future studies could also conduct exploratory analyses and cross-validate the data to further understand how empathy can be assessed in preschool children. Second, our results of concurrent validity of empathy may indicate some degrees of cultural variances, yet we did not directly measure cultural identity or cultural cognition of participants in this study, thus cannot examine the latent culture effects on children's perception of empathy. Third, our participants were recruited from two public kindergartens in the Jiangsu province. Given that China is a large country with a diverse multiethnic, multi-subcultural composition, caution is warranted when applying the EmQue to assessing Chinese children of different demographic backgrounds. Nevertheless, we would like to mention that the social-economic characteristics of participants (e.g., parents' education level, net household income) was comparable to that of the general population in China (Akimov et al.,

2021). Fourth, another limitation of this study is that only parent questionnaires were included, which could create a response bias and contain the possibility of shared method variance. Future studies could apply a multi-method and multi-informant design, for example using tasks or playground observations, to further validate this questionnaire. Lastly, the correlational nature of this study does not allow for interpretations on the causality of concurrent relations we found. Longitudinal research may have an advantage in deepening our understanding of the developmental trajectories of empathy in young Chinese children.

Researchers and practitioners can assess preschool children's dispositional proneness to experience empathy by using this parent report. Notably, our findings revealed that greater empathic responses such as feeling and attending to others' emotions could induce more personal emotional arousal in Chinese preschoolers, which could further lead to lower social competence and increased internalizing behaviors. To better support these children, parents, teachers, and practitioners are recommended to provide these children with opportunities to learn how to cope with and communicate negative emotions. Yet, for the results to be informative to parents and teachers, a standardized norm needs to be established among Chinese children for this EmQue.

CONCLUSION

The outcomes of this validation study suggest that the 19-item Chinese version of the EmQue is suitable for assessing "the multilevel empathy" in Chinese preschool children utilizing a parent report. Our results on the construct validity indicated that as early as

the preschool years, empathy already develops into a multi-level construct consisting of distinct emotional, attentional, and social facets, which entail different social-emotional functioning. We hope that this questionnaire provides a useful tool for assessing empathy in Chinese preschool children, and could promote cross-cultural research to look into whether and how the development and the function of empathy may vary in different cultural contexts.

REFERENCES

Akimov, A. V., Gemuea, K. A., & Semenova, N. K. (2021). The Seventh Population Census in the PRC: Results and prospects of the country's demographic development. *Herald of the Russian Academy of Sciences, 91,* 724–735. https://doi.org/10.1134/S1019331621060083

Atkins, D., Uskul, A. K., & Cooper, N. R. (2016). Culture shapes empathic responses to physical and social pain. *Emotion, 16(5), 587–601.* https://doi.org/10.1037/emo0000162

Bacon, D. R., Sauer, P. L., & Young, M. (1995). Composite Reliability in Structural Equations Modeling. *Educational and Psychological Measurement, 55(3),* 394-406. https://doi.org/10.1177/0013164495055003003

Bandalos, D. L. (2014). Relative performance of categorical diagonally weighted least squares and robust maximum likelihood estimation. *Structural Equation Modeling: A Multidisciplinary Journal, 21,* 102–116. https://doi.org/10.1080/10705511.2014.859510

Bandstra, N. F., Chambers, C. T., McGrath, P. J., & Moore, C. (2011). The behavioural expression of empathy to others' pain versus others' sadness in young children. *Pain, 152(5),* 1074-1082. https://doi.org/10.1016/j.pain.2011.01.024

Brislin, R. W., Lonner, W. J., & Thorndike, R. M. (1973). *Cross-cultural research methods* (pp. 47-82). New York: John Wiley & Sons. Retrieved from URL: https://journals.sagepub.com/doi/10.1177/136346157501200101

Broekhof, E., Ketelaar, L., Stockmann, L., Van Zijp, A., Bos, M. G. N., & Rieffe, C. (2015). The understanding of intentions, desires and beliefs in young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45,* 2035-2045.

https://doi.org/10.1007/s10803-015-2363-3

Brown, T. A. (2014). *Confirmatory factor analysis for applied research*. New York: Guilford Publications. Retrieved from URL: https://www.guilford.com/books/Confirmatory-Factor-Analysis-for-Applied-Research/Timothy-Brown/9781462515363

Byrne, B. M., Muthèn, B., & Shavelson, R. J. (1989). Testing the equivalence of factor covariance and mean structure. *The Issue of Partial Measurement Invariance, 105(3),* 456–466. https://doi.org/10.1037/0033-2909.105.3.456

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modelling: A Multidisciplinary Journal, 14(3),* 464–504. https://doi.org/10.1080/10705510701301834

Da Silva, B. M. S., Rieffe, C., & Veiga, G. (2022). Empathy in preschool Portuguese children: Validation of the Empathy Questionnaire (EmQue). *Social Development, 31,* 933–944. https://doi.org/10.1111/sode.12588

De Waal, F. B. M. (2008). Putting the altruism back into altruism: the evolution of empathy. *Annual Review of Psychology, 59,* 279-300. https://doi.org/10.1146/annurev.psych.59.103006.093625

Decety, J. (2011). Dissecting the neural mechanisms mediating empathy. *Emotion Review, 3(1),* 92-108. https://doi.org/10.1177/1754073910374662

Decety, J., & Jackson, P. L. (2006). A social-neuroscience perspective on empathy. *Current Directions in Psychological Science, 15(2),* 54-58. https://doi.org/10.1111/j.0963-7214.2006.00406.x

DiStefano, C., & Morgan, G. B. (2014). A Comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Structural Equation Modeling: A Multidisciplinary Journal, 21,* 425–438. https://doi.org/10.1080/10705511.2014.915373

Donohue, M. R., Tillman, R., & Luby, J. (2020). Reparative prosocial behavior difficulties across childhood predict poorer social functioning and depression in adolescence. *Journal of Abnormal Child Psychology, 48,* 1077–1088. https://doi.org/10.1007/s10802-020-00646-3

Eisenberg, N., Eggum, N. D., & Giunta, L. D. (2010). Empathy-related responding: associations with prosocial behaviors, aggression, and intergroup relations. *Social Issues and Policy Review, 4(1),* 143-180. https://doi.org/10.1111/j.1751-2409.2010.01020.x

Flook, L., Zahn-Waxler, C., & Davidson, R. J. (2019). Developmental differences in prosocial behavior between preschool and late elementary school. *Frontiers in Psychology, 10,* 876. https://doi.org/10.3389/fpsyg.2019.00876

Fokkema, M., & Greiff, S. (2017). How performing PCA and CFA on the same data equals trouble. *European Journal of Psychological Assessment, 33(6),* 399–402. https://doi.org/10.1027/1015-5759/a000460

Geng, Y.G., Xia, D., & Qin, B. B. (2012). The Basic Empathy Scale: A Chinese Validation of a Measure of Empathy in Adolescents. *Child Psychiatry & Human Development, 43,* 499–510. https://doi.org/10.1007/s10578-011-0278-6

Grazzani, I., Ornaghi, V., Pepe, A., Brazzelli, E., & Rieffe, C. (2017). The Italian version of the Empathy Questionnaire for 18- to 36-months-old children: psychometric properties and

measurement invariance across gender of the EmQue-I13. *European Journal of Developmental Psychology, 14,* 118-126. https://doi.org/10.1080/17405629.2016.1140640

Guan, R. Y., Jin, L. L., & Qian, M. Y. (2012). Validation of the Empathy Quotient – short form among Chinese healthcare professionals. *Social Behavior and Personality: an international journal, 40(1),* 75-84. https://doi.org/10.2224/sbp.2012.40.1.75

Hamaidi, D. A., Mattar, J. W., & Arouri, Y. M. (2021). Emotion Regulation and Its Relationship to Social Competence among Kindergarten Children in Jordan. *European Journal of Contemporary Education, 10(1),* 66-76. Retrieved from URL: https://eric.ed.gov/?id=EJ1294654

Hoffman, R. R. (1987). The problem of extracting the knowledge of experts from the perspective of experimental psychology. *AI Magazine, 8(2),* 54-67. https://doi.org/10.1609/aimag.v8i2.583

Hofstede, G. (2003). What is culture? A reply to Baskerville. *Accounting, Organizations and Society, 28,* 811-813. https://doi.org/10.1016/S0361-3682(03)00018-7

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

Huang, H. Q., Su, Y. J., & Jin, J. (2017). Empathy-Related Responding in Chinese Toddlers: Factorial Structure and Cognitive Contributors. *Infant and Child Development, 26 (3),* e1983. https://doi.org/10.1002/icd.1983

Huang, L., Thai, J., Zhong, Y., Peng, H., Koran, J., & Zhao, X. D. (2019). The Positive

Association Between Empathy and Self-Esteem in Chinese Medical Students: A Multi-Institutional Study. *Frontiers in Psychology, 10,* 1921. https://doi.org/10.3389/fpsyg.2019.01921

Huber, L., Plötner, M., & Schmitz, J. (2019). Social competence and psychopathology in early childhood: a systematic review. *European Child & Adolescent Psychiatry, 28,* 443–459. https://doi.org/10.1007/s00787-018-1152-x

IBM Corp. (2016). IBM SPSS statistics for windows (version 24.0).

Ip, K. I., Alison, M. L., Mayumi, K., Hidemi, H., Kazama, M. Wang, L., Olson, S. L., Kessler, D., & Tardif, T. (2021). Emotion expression and regulation in three cultures: Chinese, Japanese, and American preschoolers' reactions to disappointment. *Journal of Experimental Child Psychology,* 201. https://doi.org/10.1016/j.jecp.2020.104972

Jackson, D. L., Gillaspy Jr, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: an overview and some recommendations. *Psychological Methods, 14(1),* 6-23. https://doi.org/10.1037/a0014694

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research, 44,* 486-507. https://doi.org/10.1177/0049124114543236

Ketelaar, L., Wiefferink, C. H., Frijns, J. H. M., & Rieffe, C. (2017). Children with cochlear implants and their parents: Relations between parenting style and children's social-emotional functioning. *Ear & Hearing, 38,* 321-331. https://doi.org/10.1097/AUD.0000000000000387

Lai, K. Y. C., Luk, E. S. L., Leung, P. W. L., Wong, A. S. Y., Law, L., & Ho, K. (2010). Validation of the Chinese version of the Strengths and Difficulties Questionnaire in Hong Kong. *Social Psychiatry and Psychiatric Epidemiology, 45,* 1179–1186. https://doi.org/10.1007/s00127-009-0152-z

Li, B., Blijd-Hoogewys, E. M. A., Stockmann, L., Vergari, I., & Rieffe, C. (2023). Toward feeling, understanding, and caring: The development of empathy in young autistic children. *Autism, 27(5),* 1204–1218. https://doi.org/10.1177/13623613221117955

Li, B., Bos, M.G.N., Stockmann L., & Rieffe, C. (2020). Emotion functioning and the development of internalizing and externalizing problems in young boys with ASD. *Autism, 24,* 200-210. https://doi.org/10.1177/1362361319874644

Li, Z., Li, B., Tsou, Y. T., Oosterveld, P., & Rieffe, C. (2023). Moral emotions in early childhood: Validation of the Chinese moral emotions questionnaire. *Social Development, 32(2),* 527–545. https://doi.org/10.1111/sode.12645

Lin, Z.M., Zhou, Z.Q., Zhu, L.I., & Wu, W.G. (2023). Parenting styles, empathy and aggressive behavior in preschool children: an examination of mediating mechanisms. *Frontiers in Psychology.* https://doi.org/10.3389/fpsyg.2023.1243623

Little, T. D. (2013). (Eds.). *Longitudinal structural equation modeling.* New York: Guildford Press. Retrieved from URL: https://www.guilford.com/books/Longitudinal-Structural-Equation-Modeling/Todd-Little/9781462510160

Lucas-Molina, B, Sarmento, R., Quintanilla, L. & Giménez-Dasí, M. (2018). The Spanish version of the Empathy Questionnaire (EmQue): Evidence for longitudinal measurement

invariance and relationship with emotional regulation. *Early Education and Development, 29(4),* 467–476. https://doi.org/10.1080/10409289.2018.1427929

MacCormack, J. K., Castro, V. L., Halberstadt, A. G., & Rogers, M. L. (2019). Mothers' interoceptive knowledge predicts children's emotion regulation and social skills in middle childhood. *Social Development, 29(2),* 578-599. https://doi.org/10.1111/sode.12418.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98,* 224-253. https://doi.org/10.1037/0033-295X.98.2.224

Martín-Puga, M. E., Justicia-Galiano, M. J., Gómez-Pérez, M. M., & Pelegrina, S. (2022). Psychometric properties, factor structure, and gender and educational level invariance of the Abbreviated Math Anxiety Scale (AMAS) in Spanish children and adolescents. *Assessment, 29(3),* 425-440. https://doi.org/10.1177/1073191120980064

Matsumoto, D., Yoo, S. H., & Nakagawa, S. (2008). Culture, emotion regulation, and adjustment. *Journal of Personality and Social Psychology, 94,* 925–937. https://doi.org/10.1037/0022-3514.94.6.92

Netten, A. P., Rieffe, C., Theunissen, S. C. P. M., Soede, W., Dirks, E., Briaire, J. J., & Frijns, J. H. M. (2015). Low empathy in deaf and hard of hearing (pre)adolescents compared to normal hearing controls. *PLOS ONE, 10(4),* e0124102. https://doi.org/10.1371/journal.pone.0124102

Neumann, D. L., Chan, R. C. K., Wang, y., & Boyle, G. J. (2016). Cognitive and affective components of empathy and their relationship with personality dimensions in a Chinese

sample. *Asian Journal of Social Psychology, 19,* 244–253. https://doi.org/10.1111/ajsp.12138

Noten, M. M. P. G., Van der Heijden, K. B., Huijbregts, S. C. J., Van Goozen, S. H. M., & Swaab, H. S. (2019). Indicators of affective empathy, cognitive empathy, and social attention during emotional clips in relation to aggression in 3-year-olds. *Journal of Experimental Child Psychology, 185,* 35-50. https://doi.org/10.1016/j.jecp.2019.04.012

Ponterotto, J. G., & Ruckdeschel, D. E. (2007). An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures. *Perceptual and Motor Skills, 105,* 997–1014. https://doi.org/10.2466/pms.105.3.997-1014

Pouw, L. B. C., Rieffe, C., Oosterveld, P., Huskens, B. Stockmann, L. (2013). Reactive/proactive aggression and affective/cognitive empathy in children with ASD. *Research in Developmental Disabilities, 34(4),* 1256-1266. https://doi.org/10.1016/j.ridd.2012.12.022

Qian, G. Y., Du, Y., Liu, C. Y., Han, R. Q., & Luo., L. (2021). Are second-born children more empathetic than only Children? A study on empathy development of Chinese preschoolers. *Early Education and Development, 33(3),* https://doi.org/407-417. 10.1080/10409289.2021.1970428

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Retrieved from URL: https://www.R-project.org/

Rieffe, C., Ketelaar, L., & Wiefferink, C. H. (2010). Assessing empathy in young children:

Construction and validation of an Empathy Questionnaire (EmQue). *Personality and Individual Differences, 49(5),* 362-367. https://doi.org/10.1016/j.paid.2010.03.046

Roberts, W., & Strayer, J. (1996). Empathy, Emotional Expressiveness, and Prosocial Behavior. *Child Development, 67(2),* 449-470. Retrieved from URL: https://www.jstor.org/stable/1131826

Rodríguez, F. M. M., Clares, R. R., & Muñoz, M. R. G. (2020). Influence of resilience, everyday stress, self-efficacy, self-esteem, emotional intelligence, and empathy on attitudes toward sexual and gender diversity rights. *International Journal of Environmental Research and Public Health, 17(17),* 6219. https://doi.org/10.3390/ijerph17176219

Rosseel, Y. (2012). Lavaan: An R package for structural equation modelling. *Journal of Statistical Software, 48(2),* 1–36. https://doi.org/10.18637/jss.v048.i02

Salerni, N., & Caprin, C. (2022). Prosocial behavior in preschoolers: effects of early socialization experiences with peers. *Frontiers in Psychology, 13,* 840080. https://doi.org/10.3389/fpsyg.2022.840080

Siu, A. M. H., & Shek, D. T. L. (2005). Validation of the Interpersonal Reactivity Index in a Chinese context. *Research on Social Work Practice, 15(2), 118–126.* https://doi.org/10.1177/1049731504270382

Spinrad, T. L., Eisenberg, N., Cumberland, A., Fabes, R. A., Valiente, C., Shepard, S. A., Reiser, M., Losoya, S. H., & Guthrie, I. K. (2006). Relation of emotion-related regulation to children's social competence: A longitudinal study. *Emotion, 6(3),* 498–510. https://doi.org/10.1037/1528-3542.6.3.498

Sprafkin, J., Volpe, R. J., & Gadow, K. D. (2002). A DSM-IV – Referenced Screening Instrument for Preschool Children : The Early Childhood Inventory-4. *Journal of the American Academy of Child & Adolescent Psychiatry, 41(5),* 604–612. https://doi.org/10.1097/00004583-200205000-00018

Takamatsu, R., Tsou, Y. T., Kusumi, T., Rieffe, C. (2021). The Japanese Empathy Questionnaire (EmQue) for preschool children: Psychometric properties and measurement invariance across gender. *International Journal of Behavioral Development, 45(5),* 453–462. https://doi.org/10.1177/01650254211022852

Tampke, E. C., Blossom, J. B., & Fite, P. J. (2019). The role of sleep quality in associations between peer victimization and internalizing symptoms. *Journal of Psychopathology and Behavioral Assessment, 41,* 25–35. https://doi.org/10.1007/s10862-018-9700-8

Tong, L., Shinohara, R., Sugisawa, Y., Tanaka, E., Yato, Y., Yamakawa, N., & Anme, T. (2012). Japan Children's Study Group. Early development of empathy in toddlers: Effects of daily parent–child interaction and home-rearing environment. *Journal of Applied Social Psychology, 42(10),* 2457–2478. https://doi.org/10.1111/j.1559-1816.2012.00949.x

Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology, 90(2),* 288–307. https://doi.org/10.1037/0022-3514.90.2.288

Tsai, J. L., & Levenson, R. (1997). Cultural influences on emotional responding. *Journal of Cross-Cultural Psychology, 28 (5),* 600-625. https://doi.org/10.1177/0022022197285006

Tsotsi, S., Borelli, J. L., Backer, M., Veragoo, N., Abdulla, N., Tan, K. H., Chong, Y. S., Chen, H., Meaney, M. J., Broekman, B., & Rifkin-Graboi, A. (2021). Preschoolers' emotion reactivity and regulation: Links with maternal psychological distress and child behavior problems. *Development and Psychopathology, 35(3),* 1079–1091. https://doi.org/10.1017/S0954579421000936

Tsou, Y. T., Li, B., Eichengreen, A., Frijns, J. H. M., & Rieffe, C. (2021). Emotions in deaf and hard-of-hearing and typically hearing children. *Journal of Deaf Studies and Deaf Education, 26(4),* 469-482. https://doi.org/10.1093/deafed/enab022

Whittaker, T. A. (2012). Using the modification index and standardized expected parameter change for model modification. *The Journal of Experimental Education, 80(1),* 26-44. https://doi.org/10.1080/00220973.2010.531299

Whittaker, T. A., & Worthington, R. L. (2016). Item response theory in scale development research. *The Counseling Psychologist, 44(2),* 216–225. https://doi.org/10.1177/0011000015626273

Whittaker, T. A., & Worthington, R. L. (2016). Item response theory in scale development research: A critical analysis. *The Counseling Psychologist, 44(2),* 216-225. https://doi.org/10.1177/0011000015626273

Wu, Z . L., Mak, M., Hu, B. Y., & He, J. B. (2018). A validation of the social skills domain of the social skills improvement system-rating scales (SSIS-RS) with Chinese preschoolers. *Psychology in the Schools, 56(1),* 126-147. https://doi.org/10.1002/pits.22193

Yan, Z. Q., Zeng, X., Su, J. L., & Zhang, X. X. (2021). The dark side of empathy: Meta-analysis evidence of the relationship between empathy and depression. *PsyCh Journal, 10(5),* 794-804. https://doi.org/10.1002/pchj.482

Yi, K. Q., Kang, M. Y., Li, D. J., Wang, Z. P., Bai, J. S., Xu, H. M., & Ma, S. L. (2020). A multi-institutional and cross-sectional study on empathy in Chinese medical students: differences by student cadre or not, future career preference, and father's education status. *BMC Medical Education, 20,* 24. https://doi.org/10.1186/s12909-020-1935-x

Zhao, Q., Neumann, D. L., Cao, X. Y., Baron-Cohen, S., Sun, X., Cao, Y., Yan, C., Wang, Y., Shao, L., & Shum, D. H. K. (2018). Validation of the Empathy Quotient in Mainland China, *Journal of Personality Assessment, 100(3),* 333-342. https://doi.org/10.1080/00223891.2017.1324458