Child Psychiatry Hum Dev (2008) 39:111–122
DOI 10.1007/s10578-007-0075-4

ORIGINAL ARTICLE

A Measure of Cognitive and Affective Empathy in Children Using Parent Ratings

Mark R. Dadds · Kirsten Hunter · David J. Hawes · Aaron D. J. Frost · Shane Vassallo · Paul Bunn · Sabine Merz · Yasmeen El Masry

Received: 20 February 2007 / Accepted: 26 June 2007 / Published online: 21 August 2007

Abstract The construct of "empathy" embodies a number of characteristics necessary for psychological health in children. Surprisingly, most research has been based solely on children and adolescent report and observational measures despite evidence that multi-informant assessment is fundamental to the accurate measurement of such constructs. We present research documenting the development and validation of a brief parent-report measure of child empathy targeted at the formative years for the development of empathic skills, through to adolescence. The Griffith Empathy Measure, adapted from the Bryant Index of Empathy, showed convergence with child ratings, and good reliability and validity across gender and age. Consistent with theoretical accounts of empathy, it was found to include affective and cognitive components that showed divergent associations with other aspects of child functioning.

Keywords Empathy · Child health · Measurement · Parent-report

Introduction

Empathy, and a lack of it, is an important construct in explanations of the most appealing and appalling aspects of human behavior. The ability to understand the emotional states of others emerges early in life long before the child is able to report on such abilities. The successful early detection and remediation of problems in the development of empathy depends on one's ability to accurately measure the construct and a hallmark of research

M. R. Dadds (✉) · D. J. Hawes · S. Vassallo · S. Merz · Y. E. Masry
Department of Psychology, The University of New South Wales, Kensington Campus, Sydney, NSW 2052, Australia
e-mail: m.dadds@unsw.edu.au

K. Hunter · A. D. J. Frost · P. Bunn
Griffith University, Brisbane, QLD, Australia

with young children is the use of multi-informant assessments. Surprising, then, is the unavailability of a parent-report scale of child empathy. A recent review [1] of the role of empathy in the development of aggression in children concluded that decades of research were hampered by measurement issues. They recommended a first step toward progress should be the development of a standardized measure in a large representative norming sample. The aim of the present study was to develop and evaluate such a measure.

Definitions of empathy vary in their reference to three related components: the sharing of another's emotional state; the explicit understanding of another's emotional state; and the prosocial behaviors that follow [e.g., 2]. At the heart of these is the historical differentiation of affective and cognitive constructs [e.g., 3; see also 4 for a recent review). The latter refers to the ability to intellectually take the role or perspective of another person [5] involving the ability to decode and label emotions and their situational cues. The affective component refers to an affective response more appropriate to [6] or congruent with [7] someone else's situation than to one's own situation.

Research has investigated how cognitive and affective empathy interact [8] and the temporal sequencing of these traits in childhood offers a longitudinal view of their interrelations. Empathic responses grow in complexity throughout both childhood and adolescence. The first to develop (0–12 months) is contagious emotional arousal that is largely involuntary and automatic. During this stage, children respond to another's distress by appearing stressed and seeking comfort for themselves. With development, the child's actions are more clearly designed to help the victim. As cognitive capacity develops, so does the capacity for empathy with increasingly subtle and diverse emotions. For example, a more nebulous appreciation of pain, characteristic of more primitive empathy, may be perceived as disappointment, longing, grief, and so on [9]. Hoffman [10] describes this as setting the stage for empathy for another's life condition, in which the individual combines immediate affective response with a general representation of the plight of the other outside the immediate context (e.g., an appreciation of poverty or oppression).

Although a considerable amount of effort has been devoted to assessing the development of empathy, valid measures are still lacking. Miller and Eisenberg [11] identified four methods traditionally used to assess child affective empathy. These include picture/story methods, whereby individuals respond using a self-report technique to hypothetical stories; experimental induction procedures designed to elicit affective empathic responses; self-report questionnaires; and facial affect/gestural reactions to others emotions as depicted in films or picture/story stimuli. The two most common self-report forms are Litvack-Miller et al. [12] adaptation of the Interpersonal Reactivity Index for children, and the Bryant's Index of Empathy [13]. Both measures show moderate reliability and validity, with the Bryant having been used extensively in longitudinal studies of childhood empathy.

There are substantial problems, however, with using self-reports of empathy in children [14]. Prior to about 8 years, children lack the cognitive and/or verbal abilities to report on internal states. For older children, their reports of affective empathy and their scores on picture-story indices still do not converge with their prosocial behavior [7] and are heavily affected by demand characteristics [15] or other factors such as gender of the experimenter [16]. Laboratory-based stimuli and coding systems have advantages but are expensive and not suited to larger community studies, and thus the lack of parent report measures is most surprising. A small number of studies have incorporated teacher report [e.g., 17], however, we were unable to locate any parent-report measure. The aim of this study is to present several years of data on the use of a parent-report measure of children's empathy. As scientists working in early intervention, we were particularly interested in a brief measure that could be used in large scale community and clinical studies with young children.

Method

Participants

Participants were the parents of $N = 2612$ children between the ages of 4 and 16 (M = 7.71, SD = 3.06; 3–6 years = 41.1%, 7–11 years = 39.6%, 12–16 years = 19.3%; males = 52.8%) recruited from primary and secondary schools in Brisbane and Sydney, Australia. Sample sizes for specific analyses according to the inclusion of mothers, fathers, and children and are specified for each analysis below. Overall, the sample was largely Caucasian with minorities of Asian, Indigenous, Semitic, and Pacific-Islander families. English was the first language spoken by 83% of the families who participated in the study. Biological families (biological mother and father both living with the child) accounted for 60.2% of families, with 19% being single parent families, 6.2% blended families (step parent) and 1.9% of children living with grandparents or guardians. Mothers' and fathers' education levels were recorded as the highest education level obtained; 12% junior certificate, 37% senior certificate, 7% trade or apprenticeship, 21% tertiary level; 2% no schooling. Fathers' education levels were; 10% junior certificate, 22% senior certificate, 22% trade or apprenticeship, 12% tertiary level. One percent had recorded no schooling, with 1% currently undertaking study. The range of family income were as follows; under $20,000, 23%; $20,001–$30,000, 16%; $30,001–$40,000, 11%; $40,001–$50,000, 7%; and income >$50,000, 18.5%. Specific sample details are given for each sub-study below.

The samples were selected and recruited from 1999 to 2006 as part of several studies assessing the development, prevention and treatment of mental health problems in children and adolescents conducted by the first author. Schools vary from primary government to secondary private schools and closely reflect the general population of Australia's large eastern seaboard cities. The only inclusion/exclusion criteria were parental and child consent to participate, competence in spoken English, absence of an established diagnosis of autism or other severe developmental delay. Recruitment occurred via public talks to parents followed by information and consent forms being sent home. Parents who agreed to participate were asked to return the completed measures in sealed envelopes. Children and adolescents completed the measures in class. Data on participation rates were not available for all schools as many required total control over the distribution and collection process. Of those that were available, rates varied from a low of 16% to a high of 73%. Given the important role of the schools in recruitment, these recruitment rates reflect a combined index of school and parental involvement. Previous analyses of these recruitment rates have shown no relationship between rate and indices of child adjustment or family socio-economic status [18].

Not all participants completed all measures and some of the more intense measures, such as verbal IQ, are cumbersome in such large community samples and had to be collected on smaller samples. Extensive analyses showed that these participants did not differ from the total sample on the socio-demographics variables of maternal and paternal education level, number of children in the family, family income, or ethnicity, or on child adjustment measures (the Strengths and Difficulties Questionnaire (SDQ) see below). The one exception to this is the final validation analyses we report using the Interpersonal Response Test [19] in which a clinic-referred group is used and full details are reported in that section.

Measures

The Griffith Empathy Measure (GEM) was adapted from Bryant's Index of Empathy for Children and Adolescents [13], and is a 23-item measure in which the respondent answers

each item on a nine-point Likert scale from strongly disagree (−4) to strongly agree (+4). The GEM adopted the original nine-point Likert scale used in the original adult version [20] rather than the yes/no format designed for use with children in Bryant's version. Questions were reworded in third person format, for example in Bryant's item "I get upset when I see an animal being hurt" was changed to "My child gets upset with he/she sees an animal being hurt", and "It makes my child sad to see another child who can't find anyone to play with". The measure is available from the first named author.

The Strengths and Difficulties Questionnaire [SDQ: 21] is a 25-item rating scale that includes child- and parent report versions. The SDQ has shown increasing popularity of usage due to its combination of brevity, broad measurement domain, and strong psychometric properties. It can be scored as a total difficulties score or into five subscales: Hyperactivity, Conduct Problems, Emotional Symptoms, Peer Problems, and Prosocial Behavior. The SDQ was completed by the primary caregiver, predominantly the mother. The SDQ converges well with other checklist measures and independent diagnoses of child disorders [21–23]. A subset of children who were randomly selected from the larger sample and did not differ in demographic profile, were given the *Peabody Picture Vocabulary Test—Revised* [24] and age-standardized scores calculated.

Independent Observations of Child Behavior

A subset of boys, in groups of three, were invited to play with a pet mouse in a specially set-up room in their school. Two 16-week old mice were used and were alternated after each group of children to reduce stress on the animals. The three activities with the animal included: Activity 1, Free-play using a 'Runabout Ball'^{TM} (5 min). Activity 2, Training the mouse to run a maze (3 min); and Activity 3, Feeding the mouse (3 min). See [25] for a full description. The children's behavior was scored and rated by trained observers from videotapes on three dimensions (Nurturing: caring, empathic, gentle behavior; Cruelty: careless and/or aggressive behavior with potential to distress animal; and Engagement: active verbal and/or nonverbal involvement with the animal). Two observers were used for 36% of observations in order to check inter-rater reliability (r = .77–.85).

The Interpersonal Response Task (IRT) [19] is a computer-based measure assessing reward-dominant style of response in children. Children play a ball-throwing computer game against two computer-controlled players; the task is to decide which of the two players to throw the ball to. Children are told that they will receive 'money' (score) for throwing the ball to a particular player, and that each player will show them how they are feeling. In Block 0, each computerized child returns 'money' to the subject when they receive the ball, and each presents a 'happy' face. In Block 1, one player quickly runs out of money, though continues to exhibit a 'happy' face even when the ball is not thrown to him. In Block 2, the child who has run out of money shows an increasingly distressed face whenever the ball is not thrown to him. The aim is to examine the child's behavior in terms of their sensitivity and reactivity to the distress of the 'sad' player, in the presence of rewarding stimuli from the other player. Measures taken include the number of times the participant throws the ball to the 'sad' player as opposed to the 'happy' player, the mean and minimum level of emotion the participant allows the 'sad' child to reach, as well as the reaction time i.e., speed with which the child throws the ball to each player. Thus, the IRT measures the extent to which the participant child will ignore a distressed child in order to obtain monetary reward, and the speed with which the child makes a decision to do this. Two outcomes are reported: the maximum distress the participant child allows the com-

puterized child to reach before choosing to forego his or her own reward, and the reaction time associated with the child's choices.

Results

First we confirmed that the GEM converged with self-ratings of empathy on the original scale, the Bryant Index for Empathy [13]. Both maternal and self-ratings were available for a sub-sample of $n = 49$ adolescents aged 11–17 years who did not significantly differ on any demographic or measurement variable from the total sample described above. Correlations between the GEM and BEI total scores were $r = .412$, $P < .01$ and did not differ significantly for males or females.

The 23 items of the GEM were analyzed using Principal Components Analysis with Oblimin rotation with Kaiser normalization in SPSS, with Velicer's MAP and Parallel Analyses [26] used to determine the number of latent factors. Rather than using the common but misleading rules-of-thumb such as scree plots and eigenvalues >1, these procedures are statistically based. In parallel analysis, the focus is on the number of components that account for more variance than the components derived from random data. The eigenvalues derived from the actual data are then compared to the eigenvalues derived from the random data. In the MAP test, the focus is on the relative amounts of systematic and unsystematic variance remaining in a correlation matrix after extractions of increasing numbers of components. Components are retained as long as the variance in the correlation matrix represents systematic variance. Components are no longer retained when there is proportionately more unsystematic variance than systematic variance.

Both tests indicated the measure included two non-random dimensions; a cognitive and an affective factor, accounting for a subset of 22.32 and 15.03% of variance respectively. The two factors were uncorrelated, $r = .068$. The loadings are shown in Table 1. In order to check this two-factor solution across genders and age groups, confirmatory factor analysis (CFA) was used in AMOS 6.0 with maximum likelihood estimation, and Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA) indices to check model fit. The former compares the theoretical model to a null model and is considered acceptable at values of $\sim .9$ and above, while the RMSEA is sensitive to parsimony of model specification and is considered acceptable at $\sim .06$ and below. The two factor model showed a reasonable fit to the data: CFI = .90, RMSEA = .05. To check model fit across age group and genders, a multi-group CFA was run in AMOS in which the sample was split into three age groups and by gender. This did not result in any substantial change in model fit, both when regression weights were free to vary across sub-samples, and under the restriction of equal weights across sub-samples (CFI = .88, RMSEA = .03 and CFI = .89, RMSEA = .04 respectively). These analyses show that the GEM can be used with a wide range of ages and both genders, as a single scale using the 23 items (alpha = .81), or alternatively, scored into a cognitive empathy subscale (alpha = .62, 6 items), and an Affective subscale (alpha = .83, 9 items) after omitting items that load on both subscales. The reliability for the cognitive scale is lower than optimal, however, it is comparable to that found for the original Bryant scale which showed good convergence with criterion measures. Convergence between maternal and child ratings on the cognitive and affective scales were significant, $r = .401$, $P < .01$, and $r = .381$, $P < .01$, respectively.

Table 2 shows means and SDs for the GEM broken down by age group. There were main effects for age category on Total scores, $F(2, 2335) = 7.84$, $P < .001$, and cognitive empathy, $F(2, 2335) = 6.49$, $P < .001$, in which each age group was associated with increased empathy

Table 1 Structure matrix showing the full item and two-factor solution to the Griffith Empathy Measure

	Component	
	Affective	Cognitive
My child becomes sad when other children are sad.	.702	
gets upset seeing another child being punished for being naughty.	.591	
seems to react to the moods of people around them.	.640	
gets upset when another person is acting upset.	.796	
cries or gets upset when seeing another child cry.	.743	
gets sad when watching sad movies or TV.	.469	
becomes nervous when other children around them are nervous.	.499	
acts happy when another person is acting happy.	.349	
can continue to feel okay even if people around are upset.	.473	
can't understand why other people get upset.		.751
rarely understands why other people cry.		.761
would eat the last cookie, even when they know someone else wants it.		.431
reacts badly when they see people kiss and hug in public.		.434
doesn't understand why other people cry out of happiness.		.592
doesn't seem to notice when I get sad.		.537
gets sad to see a child with no one to play with.	.527	.486
treats cats and dogs like they have feelings	.392	.457
feels sorry for another child who is upset.	.558	.493
likes to watch people open presents, even if not one for him/her.	.250	.465
gets upset when seeing another child being hurt.	.669	.353
laughs when seeing another child laugh.	.405	.336
gets upset when seeing an animal being hurt.	.278	.345
feels sad for people who are physically disabled.	.445	.334

Blank loadings are <.2

scores, but there was no effect for affective empathy. Gender was significant on Total, $F(1, 2336) = 28.40$, $P < .001$, cognitive, $F(1, 2336) = 16.31$, $P < .001$, and affective empathy, $F(1, 2336) = 5.23$, $P < .001$, with females showing higher scores on each.

Both mother and father ratings on the GEM were available for a subset of $n = 155$ families (44% female) of non-clinic children aged 5–12 years (M = 7.61, SD = 3.0). There was reasonable agreement between mothers and fathers for Total scores (males $r = .63$, females $r = .69$) and Affective scores (males $r = .47$, females $r = .41$), and Cognitive scores (males $r = .52$, females $r = .47$). Differences in mother's and father's ratings were evident for Total and Cognitive scores. That is, while parents rated their child as having similar levels of affective empathy (mother M = 8.70, SD = 9.31, father M = 7.56, SD = 8.79), fathers reported significantly lower levels of Total (mother M = 34.07, SD = 17.63, father M = 28.5, SD = 17.8; t(155) = 4.93, P < 0.1) and Cognitive empathy than mothers (father M = 6.04, SD = 7.80, mother M = 7.87, SD = 7.21; $t(155) = 5.83$, $P < .01$). These levels of convergence and mean differences were consistent across all age groups and gender of child.

Figure 1 shows bivariate correlations between empathy ratings and mothers' ratings of the boys and girls adjustment on the five subscales of the SDQ. A clear pattern is evident for both genders in which higher total and cognitive empathy are each associated with lower behavioral and emotional problems, and higher prosocial behavior. Affective

Table 2 Means and SDs for the GEM total score and the cognitive and affective subscales split by age group and gender

		Gender of child							
		Male				Female			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
4–6 years	Total	28.63	19.67	−52.00	92.00	36.76	17.12	−16.00	92.00
	Cognitive	7.59	7.19	−15.00	24.00	9.42	7.09	−17.00	24.00
	Affective	5.96	10.45	−28.00	34.00	8.55	10.69	−27.00	36.00
7–10 years	Total	30.28	21.89	−54.00	89.00	39.78	21.53	−41.00	87.00
	Cognitive	8.15	7.87	−16.00	24.00	10.48	7.81	−16.00	24.00
	Affective	6.00	11.17	−29.00	34.00	8.99	11.89	−30.00	36.00
11+ years	Total	37.02	22.18	−30.00	80.00	42.19	18.44	.00	78.00
	Cognitive	9.74	8.20	−16.00	24.00	12.03	9.35	−18.00	24.00
	Affective	7.90	12.16	−22.00	28.00	7.51	11.49	−19.00	32.00

empathy, on the other hand, shows the gender differences previously described in the literature [27]. Affective empathy shows a consistent positive association with all indices of behavioral and emotional problems in females. In males however, it is only associated with prosocial behavior such that higher levels of affective empathy are associated with higher prosocial.

Further evidence of discriminative validity of the cognitive and affective subscales was shown by correlations between GEM scores and verbal IQ measured in a sub-sample of $n = 70$ children (58% male, aged 5–8 years, M = 6.2, SD = 1.1). Overall the GEM did not correlate with VIQ, $r = .008$, however this was due to the expected different relationships found for cognitive empathy, $r = .30$, $P < .05$, and affective empathy, $r = -.15$ ns. We also checked whether maternal reports on the GEM were confounded with background demographic and health variables. Mothers' GEM scores were independent of her education level ($r = .028$), family income ($r = .017$), and maternal stress ($r = -.06$), anxiety ($r = -.045$), and depression ($r = -.06$) as measured on the well-validated DASS scales [28].

We next assessed stability and change in mother reported empathy scores. A sub-sample of 31 parents with non-clinic children aged 5–12 (grades 1, 3, 5 & 7) completed the GEM measure a week following the first administration. Total score demonstrated a strong positive correlation between administrations ($r = .91$), as did the two subscales (affective: $r = .93$, cognitive: $r = .89$). A further sub-sample of $n = 127$ parents of non-clinic children repeated the GEM 6 months later. The measure showed impressive stability given the time frame, $r = .69$, $P < .001$.

Finally, we checked for evidence of convergence of GEM scores with direct behavioral observations of child behavior. To do this we built upon previous research showing that children's openly display nurturing versus cruel behavior toward pets when observed in small groups, and that these behaviors show good convergence with parental and child reports of an aggressive versus empathic disposition toward subordinates [25, 29]. Participants were a sub-sample of boys from a state primary school in Brisbane, Australia. From a total school population of 654 children (334 boys, 320 girls) aged between 5 years and 13 years, 30 boys were selected to continue in the study based on their scores on the Cruelty to Animals Inventory [CAI: 25] in order to ensure that a broad range of cruel

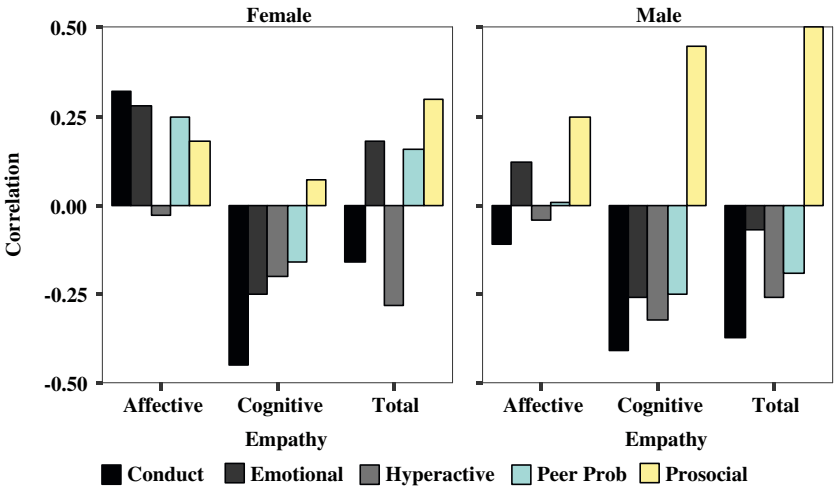

Fig. 1 Bivariate correlations between adjustment on the strengths and difficulties questionnaire subscales and total, cognitive, and affective empathy scores for males and females

versus empathic responding were represented in the final sample. Two children were absent during the session and so the final sample consisted of $n = 28$ (low CAI scores $n = 10$, medium CAI score $n = 10$, high CAI scores $n = 8$; see 25). Age ranged from 6 years to 12 years (M = 8.93, SD = 1.81). Correlations were not evaluated for statistical significance due to the small sample, but were all in the predicted direction and magnitude: Observed Pet Cruelty with GEM Total $r = -.31$, GEM affective $r = -.35$, GEM cognitive $r = -.12$. Observed Pet Nurturance with GEM Total $r = .25$, GEM affective $r = .34$, GEM cognitive $r = .05$. Clearly the convergence with both cruel and nurturing behavior is coming largely from the GEM affective empathy component.

In the second test of convergence with behavioral measures, a sub-sample of $n = 23$ clinically referred children were assessed with the GEM and the IRT. The latter is a computerized measure of (1) the extent to which the participant child will ignore a distressed child (on the computer screen) in order to obtain monetary reward, and (2) the speed with which the child makes a decision to do this. Thus two measures are reported: the maximum distress the participant child allows the computerized child to reach before choosing to forego his or her own reward, and the reaction time associated with the child's choices. The sample consisted of 19 males and 4 females ranging from 7 years to 12 years of age (M = 8.97, SD = 1.90). Participants were recruited through referrals to a Community Counseling Service in Sydney, Australia, and consisted of clinically referred children diagnosed with Oppositional Defiant Disorder ($n = 15$) and internalizing problem ($n = 8$) of Separation Anxiety Disorder, Specific Phobia, Social Phobia, Generalized Anxiety Disorder, and Adjustment Disorders. The sample was recruited according to the following criteria: (1) English as a primary language, (2) willingness to participate in the initial assessment, (3) absence of major developmental delay or psychotic illness, and (4) attending primary school. Of the 23 parents who were provided with consent forms, all agreed to participate in the study. Apart from the primary diagnoses, the children had high levels of comorbidity with a range of behavioral and emotional problems. These families were significantly lower than the total community sample on indices of parental education and income.

Correlations between mother reported Total, cognitive and affective empathy and Maximum Distress Allowed were $r = .38$, $r = .56$, and $r = .30$, such that lower GEM empathy scores were associated with the participant child allowing the computerized child to become increasingly distressed, in order to receive a monetary reward for themselves. Reaction times were highly correlated with total and affective empathy, $r = -.56$, and $r = -.57$, $P < .01$, respectively, but not cognitive empathy, $r = .15$, indicating that higher affective empathy scores were associated with taking longer to decide which child to throw the ball to (the distressed penniless child or the child with money).

Discussion

Previous research into the development of empathy has been limited by the absence of well-validated parent report measures. The current study adapted the Bryant Index of Empathy [13] to a parent-report format. The measurement properties of the scale were evaluated using a large community sample and smaller sub-samples of referred children. The evaluations provide support for the utility of the measure in capturing an overall dimension of empathy as well as two independent subscales of Affective and cognitive empathy. Specifically, we found that the scale, although based on items that were designed to capture the affective aspect of empathy, contained three distinct item sets: a subscale of affective items, a subscale of cognitive items, and a set that loading fairly equally on both. Consistent with theory and the alpha reliability data presented here, we would thus recommend that the measure could be used as a total score, or as separate affective and cognitive subscales.

The total score however, should be used advisedly, as we found clear evidence we found for divergent characteristics of the affective and cognitive components. Specifically, the affective component was unrelated to verbal IQ, and at high levels, tended to be associated with higher levels of risk of anxiety and depression type problems in girls. The cognitive component was associated with IQ, and at low levels, risk for externalizing problems in boys. There were also predictable developmental differences in the subscales. Cognitive empathy increased with age of the children and was positively associated with verbal IQ, whereas affective empathy was unrelated to age and IQ. In terms of independent observations of behavior, both affective and cognitive scores were associated with less cruel and more prosocial behavior toward other children and pets. Affective empathy scores however, were uniquely associated with increased reaction times when children were asked to respond to an "empathic dilemma". That is, children were significantly slower to make a decision about how to respond when given the choice of earning a reward versus being kind to a distressed child. Further, only affective empathy scores showed increases in response to a brief parent training intervention aimed at increasing harmonious parent–child relations. Finding differences in the affective and cognitive components are consistent with classic models of the nature of empathy [7, 30] and emphasize the importance of using measures that differentiate the constructs.

The psychometric properties of the empathy measure indicated acceptable internal consistency and stable measurement over a short test-retest interval. Of note was that the internal consistency for the cognitive factor was borderline acceptable. That is, the level obtained was consistent with the original child-report scale [13] but indicates this subscale contains considerable measurement error. Why have we not recommended against its use? As pointed out by Pedhazur and Schmelkin [31], no rule-of-thumb should apply to an acceptable level of alpha; rather the level should be evaluated against the full picture of the

performance of the scale. Given the theoretical integrity of measuring a 'cognitive' aspect of empathic behavior, the good 'fit' of the two factor affective/cognitive model across genders and ages using confirmatory factor analyses, and confirmation that the cognitive scale behaved the way we expected it should (viz., converge with verbal IQ, mental health, and independent observations of empathic behavior), we believe its use is warranted.

Inter-rater reliability was estimated using concurrent mother and father reports and the magnitude of the correlations was consistently at the high end of convergence for what is usually achieved for inter-parental ratings. This is particularly encouraging given that some of the items measuring both affective and cognitive aspects of empathy are require the parent to make an inference about the perspective of the child (e.g., my child doesn't understand why other people get upset). Compared to mothers, fathers reported that their children had less understanding of how other people feel and this applied across both genders and all age groups. Without replication and a theoretical model to contextualize this finding, it would be cavalier to make too much of it. Notwithstanding, it is possibly due to gender differences in the extent and quality of emotional talk that parent have with their children leading fathers to underestimate the level of understanding their children have about the emotions of other people.

Finding convergence between different formats of measurement, for example direct observations and self-report, has been rare in the psychometrics of empathy [1; cf. 32]. Encouraging then was the consistent evidence we found that maternal ratings of child empathy were associated with independently observed examples of children's behavior. In the IRT, children are put in a dilemma in which they can chose to be rewarded with coins or forfeit reward in order to reduce the distress of another child. Maternal ratings of empathy were associated with both the forfeiting of rewards and reaction times in deciding to do so. Further maternal ratings of empathy were associated with independent observations of how cruel versus nurturing children were when interacting with a classroom pet mouse.

The model we presented to try and capture the broader construct of empathy included cognitive, affective, and behavioral components. The measure we designed, however, was restricted to the first two intra-personal factors. The behavioral factor is controversial because outward enactment will be determined by a range of intra-personal and contextual variables that may only be loosely related to empathy. Critical to a comprehensive model, however, is the inter-connections of the components. Thus, higher levels of empathic understanding and shared affect should be associated with higher enactment of empathic behavior. The results of the current study support this basic tenet. Maternal reports of cognitive and affective empathy were consistently associated with higher enactment toward peers and subordinates. This was confirmed using independent observations of child behaviors.

Much research has tested the hypothesis that empathy problems are a characteristic or risk factor conduct problems in children, and results have not been consistent [1, 11, 32]. The likely reason for the strong differences noted in the current data is the use of parent report, both for good and bad reasons. On the positive side, parent reports are likely to overcome problems of reporting biases in the children themselves, however, on the negative side, parent reports of children referred for conduct problems are prone to contamination. That is, reports of negative qualities in these children are highly susceptible to the parent's own negative emotions, broader family problems, and acute referral factors. We tested whether the mother's reports of empathy were associated with her own social and psychological adjustment and no evidence was found that that the empathy scores were confounded with these variables that are likely to produce reporting error.

There are several limitations of the current study that should be considered. As noted above, the cognitive empathy scale contained a higher than desirable level of measurement

error and notwithstanding its good performance on other tests, may benefit from expansion in terms of the total number of items, but reduction in the diversity of emotion targets within it. Clearly the study has sampled a large group of community children and adolescents; however, we were somewhat mercenary in pooling samples and using a number of smaller sub-samples for the more intensive measures. Replication of these results in a large sample that were specifically chosen for the measure would add to generalizability of our findings. The number of fathers participating in the study were considerably lower than mothers, and father's data were not available for several of the more intensive measurements; fathers' rating converged nicely with mother's however, their overall mean scores showed some divergence. More work with fathers would allow researchers to more accurately characterize paternal involvement in the development and measurement of empathy.

Summary

This paper reports on the development of the Griffith Empathy Measure, a brief parent rating measure of empathic behavior in children and adolescents. Data were collected over several years with large samples of school children and adolescents and show it can be scored for total empathy or separated into a subset of largely orthogonal cognitive and affective components. Multi-informant and multi-method assessments of the reliability and validity of the measure indicate that it captures precise aspects of children's understanding of, and emotional resonance with, other's people's emotions.

References

1. Lovett BJ, Sheffield RA (2007) Affective empathy deficits in aggressive children and adolescents: a critical review. Clin Psychol Rev 27:1–13
2. Eisenberg N, Strayer J (eds) (1987) Empathy and its development. Cambridge University Press, New York
3. Borke H (1971) Interpersonal perception of young children: egocentrism or empathy? Dev Psychol 5:263–269
4. Blair RJR (2005) Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. Conscious Cogn 14:698–718
5. Gladstein GA (1983) Understanding empathy: integrating counseling, developmental, and social psychology perspectives. J Couns Psychol 30:467–482
6. Hoffman ML (1984) Interaction of affect and cognition in empathy. In Izard C, Kagan J, Zajonic RB (eds) Emotions, cognitions and behaviour. Cambridge University Press, Cambridge
7. Eisenberg N, Fabes, R (1990) Empathy: conceptualization, measurement, and relation to prosocial behaviour. Motiv Emotion 14:131–149
8. Duan C, Hill, CE (1996) The current state of empathy research. J Couns Psychol 43:261–274
9. Hart T (1999) The refinement of empathy. J Human Psychol 39(4):111–125
10. Hoffman ML (1990) Empathy and justice motivation. Motiv Emotion 14:151–172
11. Miller PA, Eisenberg, N (1988) The relation of empathy to aggressive and externalizing antisocial behaviour. Psychol Bull 103:324–344
12. Litvack-Miller W, McDougal D, Romney, DM (1997) The structure of empathy during middle childhood and its relationship to prosocial behavior. Genet Soc Gen Psychol Monogr 123:303–325
13. Bryant BK (1982) An index of empathy for children and adolescents. Child Dev 53:413–425
14. Eisenberg N, Miller, P (1987) The relation of empathy to prosocial and related behaviours. Psychol Bull 101:91–119
15. Eisenberg-Berg N, Hand, M (1979) The relationship of preschooler's reasoning about prosocial moral conflicts to prosocial behaviour. Child Dev 50:356–363
16. Eisenberg N, Lennon, R (1983) Gender differences in empathy and related capacities. Psychol Bull 94:100–131

17. Barnett MA, Howard JA, Melton, EM, Dino, GA (1982) Effect of inducing sadness about self and other on helping behaviour in high- and low-empathic children. Child Dev 53:920–923
18. Dadds MR, Fraser J, Frost A, Hawes, D (2005) Disentangling conduct problems and early psychopathic traits in children: a longitudinal community study. J Consult Clin Psychol 73:400–410
19. Dadds MR, Hawes D (2004) The interpersonal response test. Author: The University of New South Wales, Sydney Australia
20. Mehrabian A, Epstein N (1972) A measure of emotional empathy. J Pers 40:525–543
21. Goodman R (1997) The strengths and difficulties questionnaire: a research note. J Child Psychol Psychiatry 38:581–586
22. Goodman R, Scott S (1999) Comparing the strengths and difficulties questionnaire and the child behavior checklist: is small beautiful? J Abnorm Child Psychol 27:17–24
23. Hawes D, Dadds MR (2004) Psychometric properties of the strengths and difficulties questionnaire on a large Australian community sample. Aust N Z J Psychiatry 38:644–651
24. Dunn LM, Dunn LM (1981) Peabody picture vocabulary test-revised. American Guidance Service, Circle Pines, MN
25. Dadds MR, Whiting C, Bunn P, Fraser J, Charlson J (2004) Measurement of cruelty in children: the cruelty to animals inventory. J Abnorm Child Psychol 32:321–334
26. O'Connor BP (2000) SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. Behav Res Methods Instrum Comput 32:396–402
27. Zahn-Waxler C, Cole PM, Richardson DT, Friedman RJ (1994) Social problem solving in disruptive preschool children: reactions to hypothetical situations of conflict and distress. Merrill Palmer Q 40:98–119
28. Lovibond SH, Lovibond PF (1995) Manual for the depression anxiety stress scales. Psychology Foundation of Australia, Sydney
29. Dadds MR, Whiting C, Hawes D (2006) Associations among cruelty to animals, family conflict, and psychopathic traits in childhood. J Interpers Violence 21:411–429
30. Feshbach ND (1975) Empathy in children: some theoretical and empirical considerations. Couns Psychol 5:25–30
31. Pedhazur EJ, Schmelkin LP (1991) Measurement, design and analysis. Lawrence Erlbaum Associates, Hillsdale, NJ
32. Strayer J, Roberts W (2004) Empathy and observed anger and aggression in five-year-olds. Soc Dev 13:1–13