Brief Report

The development of a scale to measure empathy in 8- and 9-year old children[1]

Alison F Garton[2] and Eyal Gringart
Edith Cowan University

ABSTRACT

Empathy has been suggested to facilitate effective collaborative problem solving in children. The current study adapted the Interpersonal Reactivity Index (IRI, Davis, 1980), a well-validated empathy measure for adults, for use with children aged 8 and 9 years. Four hundred and thirteen school children aged between 7;11 and 9;11 years completed the new measure, 'Feeling and Thinking' (F&T) aimed at measuring both the affective and cognitive components of empathy. Principal Components Analysis with Varimax rotation produced a clear and logical four factor solution that resembled but did not duplicate the IRI. Further scrutiny reduced the scale to a more parsimonious 12 item, two factor scale representing the two components of empathy. It is suggested that F&T is capable of measuring empathy in children. Further psychometric investigations of the F&T will enhance knowledge of the affective/cognitive distinction in empathy.

INTRODUCTION

In the study of children's collaborative problem solving, it has been speculated that an awareness of, or sensitivity to, one another may be beneficial to learning (Garton, 2004). Generally, collaborative problem solving between children of differing cognitive abilities leads to learning in the less competent child. This finding is reasonably robust and is consistent with Vygotsky's (1978) zone of proximal development in which benefits in learning accrue to children after working with a more competent peer or an adult. What 'causes' such learning has been debated vigorously for decades (see reviews in Garton, 1992, 2004). Much research has focused on aspects of the interaction that might facilitate problem solving outcomes, particularly the language used during collaboration (e.g., Garton & Pratt, 2001; Teasley, 1995). More recently, attention has been paid to characteristics of the children, independent of the problem solving interaction, that may be influential in how

---

[2] Contact
  Professor Alison Garton and Dr Eyal Gringart
  Centre for Psychological Research
  Edith Cowan University
  Joondalup WA 6027 Australia

  Email a.garton@ecu.edu.au
  Phone 08 6304 5110
  Fax 08 6304 5834

children work together and talk to one another. More specifically, an awareness of the other partner and his or her capacity to solve the problem presented may facilitate children with lesser cognitive ability to benefit from collaboration. Initially, it was proposed that children's ability to benefit from collaboration was due to an awareness that the partner, particularly a partner with greater cognitive competence, would be able to help. This presupposes that children can assess the cognitive capacity of their partner, evaluate it to be greater than their own, and use appropriate language such as questions to elicit this information. The help seeking literature (e.g., Nelson-Le Gall, 1985) suggests this is no small task.

Recently, the focus of research has shifted to study the extent to which children show social sensitivity to, awareness of, or empathy to one another, in the expectation that the measurement of such an ability can then be linked later to hypotheses about improved problem solving or learning in young children. Using a collaborative problem solving framework and experimental design, O'Connor (2000; see also Reeve, Garton, & O'Connor, 2002) focused on what was termed social sensitivity of 52 nine-year-old children towards others. This research examined the role of sociability through self-report responses to scenarios. The scenarios were developed to measure social sensitivity or empathy which it was hypothesised may be connected to children's competence at relating to others and to their predisposition to establish and maintain joint communication and sharing (O'Connor, 2000). The scenarios aimed to assess the extent to which children endorsed behaviours that were regarded as helping other children and showing awareness of the need to help others (Rogoff, 1998). The paper and pencil questionnaire used five different scenarios designed to represent situations at school familiar to the children. For each scenario, children were required to respond, on a 4-point scale of 'importance', to a situation where they were to provide assistance to a peer in need within a school context. Girls responded to scenarios involving female peers and boys were given identical stories with male peers. An example is 'One of your classmates is searching in the classroom for something he/she has lost. Your classmate tells you he/she is looking for his/her library book. How important is it for you to help him/her find the library book?'.

Sensitivity was measured before collaboration (viz., at pre-test) and the children classified into three interpersonal sensitivity profiles: high, medium or low sensitivity towards others. Using a proportional reasoning task, children were also classified according to their cognitive performance at pre-test. The language used during the problem solving interaction was coded into the categories used by Garton and Pratt (2001). The results of the study confirmed the usual finding that less cognitively capable children who worked with more cognitively capable peers generally showed pre- to post-test improvement compared with children working with partners of similar cognitive ability. Children who were classified as having high interpersonal sensitivity were found to be more likely to demonstrate improvement in their problem solving post-collaboration than children classified as having medium or low sensitivity. In addition, high sensitivity children were more likely to produce language that agreed with their partner and to exhibit problem solving gains. This finding indicates that the relationship between interpersonal sensitivity, language use and cognitive improvement is complex and children who report greater social awareness benefit, either directly or indirectly, from such interactions specifically through use of language that agrees with their partner (O'Connor, 2000).

In a study conducted by Garton and Harvey (submitted) and extending the above work, 8-year-old children were assigned to pairs for problem solving collaboration on the basis of their performance on a pre-test problem solving task and their social or interpersonal sensitivity classification. Four groups of pairs were compared: children with high reasoning ability and high interpersonal sensitivity were paired with either children of low reasoning ability and high interpersonal sensitivity or children with low reasoning ability and low interpersonal sensitivity, and children with high reasoning ability and low interpersonal sensitivity were paired with either children of low reasoning ability and high interpersonal sensitivity or children of low reasoning ability and low interpersonal sensitivity. The sensitivity measure was that devised by O'Connor (2000), with minor amendments, but the

five gender-appropriate scenarios were retained, and children classified based on the distribution of the problem solving and sensitivity scores.

The results showed that there was an improvement in problem solving performance after interaction for children classified initially as low problem solving/high sensitivity regardless of with whom they were paired. Children classified as low problem solving/low sensitivity also demonstrated post-test improvement but only when paired with high problem solving/high sensitivity children. As expected, children classified as high problem solving did not show significant improvements in their pre- to post-test levels of problem solving no matter with whom they were paired. It was concluded that social sensitivity did not have a direct effect on, but rather makes an indirect contribution to, a child's capacity to learn in the problem solving task. It was noted that a shortcoming of this study was the near-ceiling levels of performance on the sensitivity measure, with a negatively skewed distribution of scores. The mean score was 15.6 out of a possible total of 20, indicating that children demonstrated high levels of social awareness but that the measure was insufficiently discriminatory for the formation of contrasting experimental groups for comparison.

It is clear from this research that a valid and reliable measure is needed, both generally and specifically, to examine the role empathy might play in successful collaborative problem solving. There are two inter-related issues that dominate the existing research literature, namely the definition of empathy and its measurement. Each will be discussed below as part of the rationale for the development of the scale reported in this paper.

Definition of empathy

Many of the research papers that include empathy in children commence with a dictionary definition. Empathy is generally regarded as 'placing yourself in someone else's shoes' but what is missing is whether empathy refers to affective (emotional) or cognitive understanding or experience, or both. There is widespread agreement that both components are necessary: Empathy implies a certain perspective-taking ability and also prosocial behaviour, that is, sharing and giving help. Other synonymous phrases, used interchangeably in the literature, are social awareness and socia sensitivity. Both of these imply prosocial behaviour, an ability to recognise and understand the feelings, needs and perceptions of another person (Garton & Harvey, submitted).

In a review of the literature, Cotton (2001) captures the definitional issue before summarising the research findings and a comprehensive dictionary definition is offered:

> Empathy is typically defined as including: (1) the AFFECTIVE CAPACITY to share in another's feelings, and (2) the COGNITIVE ABILITY to understand another's feelings and perspective … the ABILITY TO COMMUNICATE one's empathetic (*sic*) feelings and understandings to another by verbal and/or non-verbal means. (Cotton, 2001, p.9, capitals in original)

Cotton then categorises empathy research with children into the following areas:
- Childrearing and parenting practices and how these influence children's prosocial behaviour;
- Empathy training in children to increase and enhance prosocial behaviour. This is further subdivided into the specific components of empathy being taught and their outcomes; and
- Classroom strategies and program design including co-operative learning and peer tutoring, types of learning linked to collaborative problem solving.

For the purposes of the rationale for the development of the current scale, it is noted that empathy is linked to learning in children, and it is claimed to be a "key attribute of a successful learner" (Cotton, 2001, p.9, capitalised in the original).

Measurement of empathy

Assuming both an affective and a cognitive component comprise empathy, then these two ought to be measurable. Previous attempts to develop a paper and pencil scale for use by young children have been few and mixed. Zhou, Valiente and Eisenberg (2003) reviewed the methods generally used in the research literature to measure empathy, namely self report with picture stories, self report questionnaires, and self report in simulated experimental situations. In addition, they note the use of other measures including facial and vocal indices and physiological measures. The self report questionnaire is the most favoured data collection tool in research where empathy is studied in relation to some other characteristic of children, such as aggression (Feshbach & Feshbach, 1969) and prosocial behaviour (Litvack-Miller, McDougall, & Romney, 1997). Empathy in children is also studied through the identification of poor or deficient parental childrearing practices which are seen as negatively related to children's empathy development (e.g., Kestenbaum, Farber, & Stroufe, 1989). The other measurement methods are favoured in intervention and educational programs and are also used as correlates in research studies (e.g., Eisenberg, Eisenbud, & Fabes, 1993).

The Interpersonal Reactivity Index (IRI, Davis, 1980) is a scale developed for use with adults. The IRI measures, through 28 items, four components of empathy: Perspective Taking, Empathic Concern, Personal Distress and Fantasy which reflect the cognitive and affective components of empathy. The IRI has good psychometric properties but is not suitable for use with children (M.H. Davis, personal communication to H. Dunbar, 23 February 2002). A modified IRI has been used in previous research with children (Litvack-Miller et al., 1997), specifically examining the relationship between components of empathy and altruism. These authors also lament the lack of a measurement instrument, noting that many researchers prefer picture/story methods and also that the relationship between empathy and things like prosocial behaviour is influenced by the nature of the measurement instrument.

Litvack-Miller et al. (1997) constructed a 22 item scale to be used by first grade children. Items were reworded from the original IRI and based on five week test-retest reliabilities which led to the exclusion of six items with low reliabilities (< 3.3). The scale was administered orally to the young children on both occasions. As interest in the present paper is solely on the scale, the other measures used and experimental hypotheses are not of concern. The 22 items remaining were subjected to oblique factor analysis with four factors emerging; factors similar to those found by Davis (1980). These were assigned the same labels as used by Davis as they "bore a sufficiently close resemblance …" (Litvack-Miller at al., 1997, p. 310) to Davis' factors. However, on scrutiny, the items do not overlap to a great extent, with most factors having only at most five relevant items loading on them, together with items found on other factors in Davis' original analysis. The authors conceded that they had psychometric concerns with the scale, specifically the reliabilities. We would argue that the validity of the scale, and of the factors in particular, is also questionable.

The present study sought to modify the IRI into a paper and pencil scale to be used by 8- and 9-year-old children; the ages of the children who participate in the collaborative problem solving research conducted by the first author. The affective and cognitive components of empathy were as far as possible retained as it is important theoretically to view empathy as having both and the scale to be developed preserved this distinction in the items included. This paper explores the psychometric properties of the scale so developed.

METHOD

Participants

The participants were 435 schoolchildren recruited from four primary schools in Perth, Western Australia. The sample was reduced to 420 with the exclusion of some very young and some much older children, and then further reduced to 413 by removal of multivariate outliers. The children ranged in age from 7 years 11 months to 9 years 11 months. There were 194 girls and 219 boys in the sample.

Materials

The IRI was modified into child-friendly language and expression by a team, who worked on items and tried them with children of similar ages to make sure they were understood. Pilot testing resulted in no further modification to the wording of the items so the data are incorporated into the main data set. Original item 16, 'After seeing a play or movie, I have felt as though I were one of the characters', an item in the Fantasy Scale, was eliminated at the development stage, as children did not appear to comprehend what this meant (it was probably the retrospective nature of the wording that caused difficulty; this item was retained in the Litvack-Miller et al. (1997) version but with the verb 'felt' in the present tense). The 5 point Likert-type scale was retained and items 3, 4, 7, 12, 13, 14, 15, 17 and 18 were reverse coded. The scale ranged from 'Not like me at all' through to 'Hardly ever like me', 'Occasionally like me', 'Fairly like me' and 'Very like me'.

The remaining 27 items were compiled into a booklet, entitled 'Feeling and Thinking' (F&T) to capture the fact that we wanted to 'find out how children your age know what others are thinking and feeling'. The booklet also asked for basic demographic information (gender and age) together with a single sheet of photographs representing the six emotions of sadness, anger, happiness, disgust, fear and surprise. A gender-specific set of pictures was used for girls and for boys, while the order of presentation of the emotions was the same, and for each photograph, children picked one of the six emotions presented. These data will not be discussed as children successfully identified all six emotions at high levels (61.4% correct for disgust through to 97.4% correct for happiness) implying children were able to understand the feelings tapped in the F&T instrument

Procedure

After agreement from the school Principals and relevant classroom teachers, parental consent was obtained for the children's participation. Completion of the scale was taken as assent from the children. Children completed the scale in classroom groups with the same male experimenter. Instructions were the same for all groups and assistance was provided when necessary. Pilot testing had requested children to read the items and indicate that they understood them. This identified the items that gave rise to the greatest difficulty so these were always dealt with in the detailed instructions given to subsequent groups.

RESULTS

An initial decision was made to exclude multivariate outliers and to remove the reverse coded items. The former was conducted through the computation of Mahalanobis distance for each case with $p < 0.001$ which led to the exclusion of the data of seven children. The latter decision was taken because of the number of queries made by children of the meaning of these items. Every classroom group raised questions about the meaning of at least one of these items. It was thus assumed that this would be reflected in subsequent difficulties with answering the items and unreliability in the scale. These nine items were distributed evenly across Davis' original four factors, so their elimination ought not to affect adversely the original four factor structure. These nine items were removed, leaving a scale of 18 items for analytic purposes; four items from the original Empathic Concern sub-scale, four from the Perspective Taking sub-scale, six from the Personal Distress sub-scale and four from the Fantasy sub-scale.

The mean total score was 60.22 with a standard deviation of 11.12 and range from 26 to 89 (out of a possible 90, with 18 items). Skewness was -.01 and kurtosis, -.15, both non-significant.

Factor analysis, using Varimax rotation and including items with a weight > 0.3, resulted in a four factor solution, accounting for 42.7% of the variance. Although Litvack-Miller et al. (1997) argue for the use of an oblique rotation due to high intercorrelations between items, an Oblimin rotation did not yield a substantially different factor structure. The factors are shown

in Table 1. Reliability of these four factors, as components of empathy, was 0.73, 0.55, 0.23 and 0.35 respectively. The intercorrelations between the four factors, based on the scores on the items, are shown in Table 2.

Table 1: Four factor structure solution of the' Feeling and Thinking' instrument

Item	Factor			
	1	2	3	4
I am quite a soft-hearted person	.66			
I want to help people who get treated badly	.61			
Emergency situations make me feel worried and upset	.56			
I often feel worried about people that are not as lucky as me, and feel sorry for them	.56			
I get very worried and upset when I see someone who needs help in an emergency	.56			
When I am angry or upset at someone, I usually try to imagine what he or she is thinking or feeling	.46			
I often get affected by things I see happen	.43			
When I am arguing with my friends about what we are going to do, I think carefully about what they are saying before I decide whose idea is best	.38			
When people around me are nervous or worried, I get a bit scared and worried too	.36			
When reading a good story, I imagine what it would be like if the story were true		.75		
When reading a book, I try to imagine what the people in the story are thinking		.68		
It is easy for me to pretend that I am the star of my favourite movie		.49		
I try to think about other people's feelings before I make mean comments on them		.47		
I am likely to lose control during an emergency			.61	
I day dream quite a lot about things that might happen to me			.61	
I sometimes try to understand my friends better by pretending I am them . . .				.68
Sometimes I feel helpless when people around me are upset				.51
I think people can have different opinions about the same thing				.46

Table 2: Intercorrelations between factors

	Factor I	Factor 2	Factor 3	Factor 4
Factor I		.424**	.218**	.407**
Factor 2			.166**	.297**
Factor 3				.102*

 * *p* <.05
 ** *p* <.01

   Further exploration of the data, driven by an interpretation of the factors and a perusal of the reliabilities, resulted in removal of the items loading on the second and third factors. The remaining items were then forced into a two factor solution of 12 items, six loading on Factor 1 and six on Factor 2. These accounted for 36.4% of the variance and three items originally loading on Factor 1 now loaded on Factor 2 (initial Factor 4). Factor 1 had a reliability of 0.69 and Factor 2's reliability improved to be 0.54. This final solution is shown in Table 3.

Table 3: Two factor structure solution of the' Feeling and Thinking' instrument

Item	Factor 1	Factor 2
Emergency situations make me feel worried and upset	.75	
I get very worried and upset when I see someone who needs help in an emergency . .	.58	
I want to help people who get treated badly	.57	
I often get affected by things I see happen	.56	
I often feel worried about people that are not as lucky as me, and feel sorry for them		
I am quite a soft-hearted person	.55	
I sometimes try to understand my friends better by pretending I am them	.53	
I think people can have different opinions about the same thing		.78
When people around me are nervous or worried, I get a bit scared and worried too . .		.52
When I am angry or upset at someone, I usually try to imagine what he or she is thinking or feeling		.45
Sometimes I feel helpless when people around me are upset		.42
When I am arguing with my friends about what we are going to do, I think carefully about what they are saying before I decide whose idea is best		.40 .38

Factor score variables, saved by the regression method, were inspected for gender differences and girls scored more highly than boys on Factor 1 ($F(1, 411) = 28.16$, $p > 0.001$) and on Factor 2 ($F(1, 411) = 7.19$, $p > 0.01$).

DISCUSSION

The scale demonstrated good psychometric properties. Screening for normality prior to analysis revealed a wide range of total scores on the scale with a slight negative skewness and kurtosis but the data were generally distributed normally.

The initial four factor solution was both similar and dissimilar to that obtained previously with adults (Davis, 1980) and children (Litvack-Miller et al., 1997). In all cases a four factor solution has been obtained but the composition, and hence nomenclature, of the factors varied. The first factor that emerged was a general affective factor, combining items that loaded on Davis' Personal Distress factor and Empathic Concern factor plus two items from the original Perspective Taking cognitive factor. Factor 2 is the Fantasy Factor, while Factor 3 contains two items that could be labelled 'Fatalistic'. Factor 4 is the cognitive Perspective Taking factor. The four factors are also unlike those found by Litvack-Miller at al. (1997). They too found a large first factor but labelled it Perspective Taking, even though it contained a number of items some of which were not in the original Davis' Perspective Taking factor.

Davis in his experimental work that succeeded the development of the IRI focused on three factors only – the cognitive Perspective Taking and the two affective ones of Empathic Concern and Personal Distress (e.g., Davis & Oathaut, 1987). In other words, he no longer included the Fantasy scale, a decision that could be supported by the factor structure that emerged with the scale modified for children. Forcing the factors interpreted as Affective and Cognitive (initial Factors 1 and 4 respectively) into a two factor solution resulted in a clear separation of these components in an Affective factor (1) and a Cognitive factor (2) and also reduced the scale from 27 to 12 items, making it more manageable for children. Furthermore, the two facets of affective empathy do not appear as separate sub-scales for children in this study. Four Personal Distress items were distributed across both factors, with two items loading on each. This could be because the children do not differentiate the two or perhaps the amended wording did not capture any subtle distinction that Davis was trying to make.

The gender differences on the scores on the two factors reveal that girls are more empathic than boys, scoring more highly on both the affective and the cognitive components. This confirms Livack-Miller et al.'s (1997) finding of girls scoring higher on all dimensions of empathy.

The measurement scale used was not linear. The second and third points on the 5-point scale included a temporal element ('hardly ever' and 'occasionally') which was lacking for the other three points. Thus it might be surmised that children might have some difficulty differentiating, for example, 'occasionally like me' from 'fairly like me' which are not immediately distinct from one another. The language used in the measurement scale warrants attention in future research.

For the purpose for which it is intended, namely the measurement of individual levels of empathy or social sensitivity, the scale certainly will be capable of differentiating children having various levels or degrees of empathy. These levels can either be psychometrically derived using percentile cut-offs or taken from the distribution of scores reported here. Future work will not only use the scale to measure children's empathy prior to collaborative problem solving but further psychometric investigation of the instrument will be undertaken to probe more precisely the affective/cognitive distinction in empathy and examine this in relation to age.

REFERENCES

Cotton, K. (2001). Developing empathy in children and youth. *School Improvement Research Series* Close Up # 13. Retrieved 12 December 2003 from http://www.nwre1.org/scpd/sirs/7/cu13.html

Davis, M.H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10,* 85.

Davis, M.H., & Oathaut, H.A. (1987). Maintenance of satisfaction in romantic relationships: Empathy and relational competence. *Journal of Personality and Social Psychology,* 53, 397-410.

Eisenberg, N., Eisenbud, L., & Fabes, RA. (1993). Behavioural and physiological correlates of children's reactions to others in distress. *Developmental Psychology,* 29, 655-663.

Feshbach, N.D., & Feshbach, S. (1969). The relationship between empathy and aggression in two age groups. *Developmental Psychology,* 1, *102-107.*

Garton, A.F. (1992). *Social interaction and the development of language and cognition.* Hove: LEA.

Garton, A.F. (2004). *Exploring cognitive development: The child as problem solver.* Oxford: Blackwell Publishing.

Garton, A.F., & Harvey, R. (submitted). Social sensitivity as a predictor of the benefits of collaborative problem solving in children.

Garton, A.F., & Pratt, C. (2001). Peer assistance in children's problem solving. *British Journal of Developmental Psychology,* 19, 307-318.

Kestenbaum, R, Farber, E.A., & Stroufe, LA. (1989). Individual differences in empathy among preschoolers: Relation to attachment history. In N. Eisenberg (Ed.), *Empathy and related emotional responses. New directions in child development series, no.* 44. San Francisco: Joosey-Bass.

Litvack-Miller, W., McDougall, D., & Romney, D.M. (1997). The structure of empathy during middle childhood and its relationship to prosocial behaviour. *Genetic, Social and General Psychology Monographs,* 123, 303-325.

Nelson-Le Gall, S.A. (1985). Necessary and unnecessary help seeking in children. *Journal of Genetic Psychology,* 148, 53-62.

O'Connor, J. (2000). *The relationship between interpersonal sensitivity and language use in children's collaborative problem solving.* Unpublished Honours thesis, University of Melbourne.

Reeve, RA., Garton, A.F., & O'Connor, J. (2002). *The role of interpersonal sensitivity in collaborative problem solving interactions.* Poster presented at the 17th Biennial meetings of ISSBD, Ottawa, Canada.

Rogoff, B. (1998). Cognition as a collaborative process. In D. Kuhn & RS. Siegler (Vol.. Eds.), W. Damon (Ed. In Chief) *Handbook of child psychology. Volume two: Cognition, perception and language* (5th ed., pp. 679-744). NY: John Wiley & Sons.

Teasley, S. (1995). The role of talk in children's peer collaboration. *Developmental Psychology,* 31, 207-220.

Vygotsky, L (1978). *Mind in society: The development of higher mental processes.* Cambridge, MA: Harvard University Press.

Zhou, Q., Valiente, C., & Eisenberg, N. (2003). Empathy and its measurement. In S.Lopez & C.R. Snyder (Eds.), *Positive psychological assessment: A handbook of models and measures* (pp. 269-284). Washington DC: American Psychological Association.