

## **CHAPTER 5: DESCRIPTION OF GENERAL METHODOLOGY**

In this chapter the methodology common to the different studies of the model and its instrument is described. As such, the characteristics of the participants, including demographical information, profiles of the measures used and specifics of the procedures implemented are detailed. Additional methodological issues specific to the individual studies are discussed within the relevant chapters.

### **Participants**

In line with the predominance of individuals of low socio-economic status in child protective services databases (e.g., Drake & Zuravin, 1998), the targeted participants for this study were parents dependent on social security pensions and of limited education (failure to complete high school). To further enhance homogeneity, participants were of the Australian majority cultural grouping. That is, targeted participants did not identify as having either aboriginal or non-English speaking backgrounds, and did not have known significant intellectual or physical disabilities. This is not to deny that such groups are vulnerable to child maltreatment, but each of these groups would warrant examination in its own right; such an objective was not feasible for the present study.

Participants in this study were drawn from the client populations of a number of agencies including DOCS, Family Support Services, neighbourhood centres, child care centres, and the Burnside Intensive Family Based Services (Campbelltown, Sydney). These agencies are located in a wide geographical area extending from Sydney to the NSW-

Queensland border and included both urban (Western Sydney) as well as coastal and isolated inland rural communities.

Three groups of participants were recruited: a group of known maltreating (abusive) parents (excluding paedophiles), a matched group of known 'good enough' parents and a group of parents who were not known to be abusive but were experiencing difficulties with their parenting and hence seeking assistance from community organisations. Consistent with the study by Rosenberg and Repucci (1983), this latter group of parents were labelled 'Distressed'. Abusive parents all had recent confirmed histories of child maltreatment (physical abuse or neglect or both) and their children were registered as at risk with DOCS.

Thus, the final sample comprised 103 parents in three risk groups: parents considered to be abusive ('Abusive' group;  $n = 50$ ); parents considered under stress (Distressed group;  $n = 32$ ); and parents not considered to be at risk of child abuse ('Control' group;  $n = 21$ ). Details of the demographics for parents and children relative to risk group membership and for the total sample are summarised in Table 1, below.

As shown in Table 1, significantly fewer fathers (15%) than mothers (88%) participated in the study ( $\chi^2_{(1,103)} = 51.738; p < .001$ ), however there was no difference in this gender distribution across risk groups ( $\chi^2_{(2,103)} = 2.309; p = .315$ ). The average ages of participating parents and target children for the total sample were 36.0 years ( $\pm 6.9$ ) and 8.6 years ( $\pm 3.2$ ), respectively. There were no differences in ages between risk groups for either parents ( $F_{(2, 100)} = .534, p = .588$ ) or children ( $F_{(2, 100)} = .543, p = .583$ ).

**Table 1: Demographics for each risk group and for total sample.**

Variable	Risk Group			Total
	Control	Distressed	Abusive	
Parent status (n/%)				
single	14/67	17/53	28/56	59/57
partner	7/33	15/47	22/44	44/43
Parent's gender <sup>a</sup> (n/%)				
males	5/24	5/16	5/10	15/15
females	16/76	27/84	45/90	88/85
Parent's age				
<i>M</i>	36.4	34.9	36.5	36.0
( <i>SD</i> )	(5.5)	(7.0)	(7.4)	(6.9)
Parent's education <sup>b</sup> (n/%)				
left before year 10	14/67	25/78	41/82	80/78
left before year 12	7/33	7/22	9/18	22/22
No. of Children <sup>c</sup>				
<i>M</i>	2.5	2.8	3.7	3.2
( <i>SD</i> )	(1.3)	(1.4)	(1.6)	(1.6)
Child's gender <sup>d</sup> (n/%)				
males	13/61	17/53	33/66	63/61
females	8/38	15/47	17/34	40/39
Child's age				
<i>M</i>	8.3	8.2	8.9	8.6
( <i>SD</i> )	(3.4)	(3.1)	(3.2)	(3.2)
Child's birth order <sup>e</sup>				
<i>M</i>	1.6	1.4	2.3	1.9
( <i>SD</i> )	(.9)	(1.0)	(1.2)	(1.2)
Stress level				
<i>M</i>	1.3	1.7	1.4	1.5
( <i>SD</i> )	(1.2)	(1.1)	(1.2)	(1.2)
Stress duration				
<i>M</i>	.9	1.3	1.0	1.1
( <i>SD</i> )	(.9)	(.9)	(1.0)	(.9)

<sup>a</sup>  $\chi^2_{(1,103)} = 51.738; p < .001$ ; <sup>b</sup>  $\chi^2_{(1,103)} = 3.796; p < .001$ ; <sup>c</sup>  $F_{(2, 100)} = 6.210, p = .003$ ; <sup>d</sup>  $\chi^2_{(1,103)} = 5.136; p = .023$ ; <sup>e</sup>  $F_{(2, 100)} = 6.827, p = .002$ .

The majority of parents in the total sample had left school before year 10 (79%), the remaining parents had completed year 10 but left before year 12 (21%). No parents in the sample had stayed at school beyond year 12. There was no difference in the pattern of

education levels achieved across risk groups ( $\chi^2_{(2,103)} = .991; p = .609$ ). In the total sample, 57% were single parents and 43% had partners, this pattern of parenting did not differ across risk groups ( $\chi^2_{(2,103)} = 1.015; p = .602$ ).

For the total sample, family sizes averaged 3.2 children ( $\pm 1.6$ ), and the average birth order of the target child (oldest child within the age-range of 12 months to 12 years) was 1.9 ( $\pm 1.2$ ). There were significant differences between groups for both average family size ( $F_{(2, 100)} = 6.210, p = .003$ ) and average birth order of target children ( $F_{(2, 100)} = 6.827, p = .002$ ). *Post hoc* tests (Scheffe') revealed that average family sizes for both the Control group ( $M = 2.5$ ) and Distressed group ( $M = 2.3$ ) did not differ ( $p = .772$ ), but both were smaller than that for the Abusive group ( $M = 3.7; p = .009$  and  $.039$ , respectively). Similarly, the average birth orders of the target children in both the Control group ( $M = 1.6$ ) and Distressed group ( $M = 1.4$ ) were comparable ( $p = .909$ ), but both were higher than that for the Abusive group ( $2.3; p = .049$  and  $.004$ , respectively).

Overall, the sample of target children comprised a significantly higher number of boys ( $n = 61\%$ ) than girls ( $39\%; \chi^2_{(1,103)} = 5.136, p = .023$ ). However, there was no significant difference in the distribution of genders across the three groups ( $\chi^2_{(2,103)} = .368; p = .505$ ).

The general level of current stress reported by participants was also noted. While some 32% of participants reported they experienced no current stress, 14% reported low levels of stress, 31% reported medium levels of stress, and 23% reported high levels of stress. The frequencies of reported stress levels did not differ across groups ( $\chi^2_{(6,103)} =$

2.548,  $p = .863$ ). In addition, 39% of participants rated the duration of their stress levels as only of short duration, 14% as medium duration and 48% had experienced stress for a long time. The reported stress duration did not differ across groups ( $\chi^2_{(4,103)} = 5.640, p = .228$ ).

In response to questions about prior diagnoses of psychiatric problems, 40% of participants reported a history of depression and 34% reported a history of anxiety-related illness. There were no differences in the frequencies of either depression ( $\chi^2_{(1,103)} = 1.978, p = .372$ ) or anxiety-related disorders ( $\chi^2_{(1,103)} = 1.598, p = .450$ ) across groups.

In summary, the Abusive group differed from the Distressed and Control groups on the basis of the number of children in the family and the birth order of the targeted child, but were otherwise equivalent across all variables. Thus, as these data make clear, the correspondence between groups was generally satisfactory.

### Measures

Two instruments were utilised in this study: the Parental Empathy Measure whose development is the focus of this thesis, and the Child Abuse Potential Inventory (Milner, 1986). Both of these instruments are described below.

#### *The Parental Empathy Measure*

The Parental Empathy Measure (PEM) has been described in considerable detail in Chapter 3: Evolution of the Parental Empathy Measure (PEM). A copy of the complete PEM, together with a description of scoring procedures, are included as Appendix G: Parental Empathy Measure (PEM) and Appendix H: PEM Scoring Guide, respectively. In general terms the PEM may be described as follows. The PEM consists of a semi-

structured interview divided into four sections. The first section of the PEM gathers relevant demographic information using an open-ended question format. Participant's gender, parenting status, number, ages and gender of children are charted within this section. This section also includes two screening questions aimed at identifying current or past involvement of child protection agencies. The second section of the PEM entitled 'Beliefs', comprises 13 forced-choice items that are designed to survey the participant's overall values and cognitive beliefs about children and parenting.

The third section of the PEM comprises a series of ten scenarios of common situations with children that may be challenging for parents. Each of these scenarios is accompanied by a series of questions that invite the participant to report how they would most likely respond *behaviourally* to the child in both normal circumstances and during a 'bad' day, to report how they would respond *emotionally* in both normal circumstances and during a 'bad' day; and finally, to *explain* why the child is behaving that way. These questions were designed to measure respectively, the participant's behavioural response both normally (Behaviour-unstressed), and under stress (Behaviour-stressed), the participant's emotional response both normally (Emotion-unstressed), and under stress (Emotion-stressed), and the nature of the attributions that the participant holds about children (Attribution).

The final section of the PEM comprises of a series of open-ended questions that were designed to assess a number of variables related to the participant's parenting, current situation and history. Most notably, the participant's ability to detect his or her children's signals was assessed by items within this section. Other variables assessed in this section

included the participant's own child abuse history; insight into parenting, use of drugs and alcohol and levels of stress; parental depression and anxiety; sense of success in parenting and openness to professional input. Additionally, a social desirability or 'Faking Good' scale is incorporated in both the scenarios section (Section Three) and the final OEQ section.

In all sections except Section 2 (Beliefs), items in the PEM generally took the form of open-ended questions. Exceptions to open-ended questions were Likert-style scales (10- and 5- point) that included questions that asked the participants to estimate: Their children's and their own self-esteem; the level of happiness/unhappiness of their childhood; the nature of their relationship with their children now, compared to previously; their disappointment/ satisfaction with parenting; their stress level currently compared to previously; and the level of support they enjoy in their parenting. Additionally, one question required the participant to describe the characteristics of the primary caregivers they experienced during their own childhood in terms of forced-choice responses to a series of dichotomous items (e.g., loving or unloving, critical or accepting, unaffectionate or affectionate).

The encoding and scoring of response protocols for the PEM have been described earlier (Chapter 3: Evolution of the Parental Empathy Measure (PEM) and Appendix H: PEM Scoring Guide). Data were obtained on each participant's parental empathy variable scores for all of their children within the target age range of 12 months to 12 years but, for the purposes of brevity the participants' responses only in regard to their oldest child within the age-range was included in the analyses.

*The Child Abuse Potential Inventory*

The Child Abuse Potential Inventory Form VI (CAPI; Milner, 1986) is a 160-item, self-administered questionnaire that is answered in a forced choice, agree/disagree format. The questionnaire contains a 77-item physical child abuse scale that has six factorially-determined subscales. Two factors (Distress and Unhappiness) are believed to describe emotional characteristics while the remaining three scales (Problems with Child, Problems with Self, Problems with Family, Problems from Others) are believed to describe interactional difficulties. A seventh factor, labelled Rigidity, is believed to describe attitudes towards children. In addition to the Abuse scale the CAPI includes three response distortion indexes (Faking-Good, Faking-Bad, and Random Response).

The CAPI is reportedly one of the best validated instruments available in child maltreatment research and clinical practice (Daro, 1988; Miller & Hauser, 1989). The CAPI has demonstrated strong internal consistency (KR-20) reliabilities for the Abuse scale, ranging from .92 to .96 for groups of abusive and comparison groups. Temporal stability for 1 week, 1 month and 3 months is reported to be .90, .83 and .75 respectively (Milner, 1986). Studies indicate that abuse classification rates are typically in the mid 80% to the low 90% range for physical child abusers on average, and 73% for mixed groups of physically abusive and neglectful parents (Milner, 1994). Construct validity has been extensively assessed (Milner, 1994), and evidence of predictive validity has also been reported (Milner, Gold, Ayoub, & Jacevitz, 1984).

The current study focussed primarily on three of the CAPI scales: the Abuse scale; the Lie scale, incorporating the faking good and faking bad indexes; and the Rigidity scale.



## Procedure

### *Ethics*

Approvals for the study was obtained from the University of New England's Deputy Vice-Chancellors Committee for Ethics in Human Experimentation prior to the commencement of the original case-studies (HE 970041, Appendix F: Ethics Approval Notice for Main Study).

### *Recruitment of Agencies*

Following the preliminary series of case studies, a process of enlisting the cooperation of referral agencies was undertaken. Negotiations with senior management of DOCS had already been successfully completed. Subsequent negotiations were undertaken at the level of Area Managers, followed by negotiations with District Managers and finally contact with individual child protection workers. Additionally, psychologists from DOCS were enlisted as interviewers using the PEM and the CAPI.

In regard to the community agencies, phone contact was initially made with the coordinators or managers. With permission of the coordinator or manager, a package of information including information sheets for workers (Appendix K: Information Sheet for Workers), a 'plain English' information sheet (Appendix L: Plain English Statement) and a 'consent to make contact' form (Appendix M: Consent to Make Contact) for distribution to potential participants was submitted for review by the agency. Following the receipt and perusal of these documents, the relevant agency was visited by the author and a meeting held with all case-workers to provide further explanation of the recruitment process. The ethics requirements that participants needed to be voluntary and that refusal

to participate would carry no penalties for the clients of the agency were stressed at all stages of the recruitment process.

#### *Recruitment of Participants*

In those agencies who agreed to participate in the study, including the offices of DOCS, suitable parents were approached by their case-workers, provided with the information sheets inviting their participation, and informed of their rights to refuse. Parents who indicated their agreement to participate in the study were then asked to complete a consent form which authorised the release of the parent's name, address, and telephone number and convenient times to call (Appendix M: Consent to Make Contact). Following the receipt of the 'consent to make contact', in the majority of the cases the author rang the parent and once again detailed the conditions of participation, including the participant's right to refuse to participate without penalty or to withdraw at any point of the interview process. The limits of confidentiality were also specified in the information sheet, by the case-worker, and by the author in the phone contact. If the parent was still agreeable to participating in the study, a date and time was organised for the interview to be conducted. In some instances, however, the agency worker scheduled the parent's participation at a pre-arranged date and time.

Participants known to be at risk of child abuse and neglect were recruited from the government child protection agency, DOCS, and also from the Burnside Intensive Family Based Service (Campbelltown, Sydney) and from several Family Support Services. The Burnside Intensive Family Based Service and the Family Support Services have a primary focus of working in partnership with DOCS to provide intervention and support to

families identified as at risk of child abuse and neglect. All participants in the abusive group had recent confirmed histories of child maltreatment (physical abuse or neglect, or both) and their children were currently registered as at risk with DOCS.

In addition to clients referred by DOCS, Family Support Services typically offer support and parent training intervention to self-referred and other-referred families. As indicated above, referred participants from Family Support Services who were not known to have been or currently to be at risk of abusing or neglecting their children were categorised as Distressed. Other sources of referrals of Distressed participants were Women's Services such as Women's Housing Organisations and Women's Refuges. These services provide short to medium term accommodation and support to women and children who have experienced domestic violence or are in some other way acutely disadvantaged and in distressed circumstances. Participants referred from these services were also identified by their case-workers as having difficulties with their parenting. However, with one exception none of these participants were known to be at risk of abuse or neglect. The particular participant was concurrently referred by DOCS and was subsequently assigned to the abusive group.

Participants in the control or non-abusive group were recruited largely from child care organisations such as long day care centres and from neighbourhood and community services. These participants were not known to be abusive or neglectful of their children nor to have involvement with DOCS. The neighbourhood and community services typically offer social, educational and self-growth type activities to self- and other-referred

individuals. As a further check, early in the interview schedule participants were asked about any agency involvement in their family, including involvement with DOCS.

From the distressed and control group only one respondent indicated involvement with DOCS. Upon further inquiry, this involvement occurred as a result of allegations of sexual abuse upon the children from the other parent. Nevertheless, this participant's interview was eliminated from the analysis.

#### *Process of Selection of Participants*

From the various agencies originally 116 parents (100 women, 17 men) of children aged between 12 months and 12 years were recruited. To maintain independence of data sets, none of the participants were members of the same immediate family. Of the 116, four were rejected on the basis that they were foster parents rather than natural parents, one potential participant was rejected on the basis that she did not meet the socio-economic criteria and one was rejected on the basis that she identified as an Aboriginal. A further two potential participants were rejected on the basis that their children had significant developmental disabilities and a third was rejected on the basis that new allegations of paedophilic activities were under investigation. Additionally, three interviews were culled on the basis of incomplete data, and one potential 'control' participant was rejected on the basis of previous DOCS involvement. Thus the final sample of participants comprised 103 participants; 50 in the Abusive group, 32 in the Distressed group and 21 in the Control group.

*Interview Procedure*

Although workers were fully informed of the purpose of the study, the information sheet for potential participants was less than fully frank in that the focus of the study was stated to be the experiences of parents rather than an assessment of parental empathy (see Appendix N: Generic Information Sheet). This strategy was considered necessary in order to minimise the potential of social desirability response biases.

Individual interviews were conducted either in the offices of the referring agency or in the participant's home. However, no interviews were conducted on DOCS' premises. With the participants' permission interviews were audiotaped to allow for a more comprehensive collection of the data.

Participants who had travelled to the agency for the interview were offered \$10 remuneration to cover their travel costs. In one instance a participant was offered \$20 reimbursement for her child care costs to allow the interview to be conducted in a minimally distractive environment.

Sixty-two percent of the interviews were conducted by the researcher. The remaining interviews were conducted by research assistants who were all qualified social workers or psychologists and had had at least 12 months experience in working with distressed and abusive families. Four of the psychologists were current staff members of the DOCS and undertook the interviews as part of DOCS' commitment to the research project. Interviews of clients from the Burnside Intensive Family Based Services were also undertaken by Burnside case-workers. All interviewers received detailed instructions on the administrative requirements of the instruments. In particular, appropriate prompts and

procedures to guard against influencing participants' responses in the administration of the PEM were specified. In instances wherein the client was interviewed by their case-worker, participants were provided with further information in that the original confidentiality issues were altered (see Appendix O: Case-Worker Information Sheet). Thus, potential participants were informed in advance that the information revealed in the interview may inform the case-planning offered by their case-workers.

Before the interview was formally commenced, all potential participants were again provided with the plain English information sheets (Appendix N: Generic Information Sheet or Appendix O: Case-Worker Information Sheet as applicable) and their rights of refusal and limits of confidentiality re-emphasised. A consent form which specified that the procedure (outlined above) had been undertaken and that the participant consented to the interview was then completed (see Appendix P: Consent to Participate Form).

Participants were interviewed individually with the PEM, then the CAPI was self-administered. Five participants were administered the CAPI verbally due to poor literacy skills. Interviews with the PEM included all of the participant's children who fitted within the required age category (12 months to 12 years). The interviewing process typically took from 1.5 to 3 hours to complete (average = 2 hours). Longer-than-average interviews were usually a consequence of participants having more than two children meeting the age criterion. In some circumstances, the interview process raised previously unresolved issues for the participants (e.g., previously undisclosed sexual abuse histories). At those times the interview process was halted to allow time for the participant and interviewer to achieve some initial resolution of the issues raised.

### Data Analyses

All audiotaped PEM interviews were transcribed verbatim by independent transcribers. In six cases, however, the audiotaping was too distorted to allow processing. In these six cases, encoding was conducted on the basis of the interviewer's written responses. To maintain confidentiality, all copies of transcribed interview records were assigned a particular number, this number was also recorded on the CAPI form, the consent and referral forms, the original written interview record, tapes and scoring sheets. The scoring sheets and transcribed interview were then separated from the referral forms, tapes, and consent forms. At the point of the scoring of the interview, therefore, the interviewer was blind to the identity of the participant, the source of the referral, and the categorisation of the participants' risk status.

Each PEM transcript was analysed separately and encoded in accordance with the proscribed scoring protocols. No decision was made to discard PEM interview data on the basis of high ratings of the faking good scales as appropriate cut-off scores had yet to be determined.

## **CHAPTER 6: STUDY OF THE RELIABILITY OF THE PARENTAL EMPATHY MEASURE**

One of the first objectives in the design and evaluation of a psychological instrument is the assessment of its reliability. Three common tests of reliability are inter-rater reliability, test-retest reliability, and internal consistency reliability (Aiken, 1997; Ferris & Norton, 1992; Groth-Marnat, 1984). The limited time-frame available for this study, in combination with the length of the interview schedule itself, prohibited an assessment of the PEM's test-retest reliability. Such an assessment is planned as funds become available. However, tests of the PEM's internal consistency and inter-rater reliability for the individual and collective parental empathy variables and the Faking Good scale were conducted and are reported in this chapter. In addition, four other PEM variable scales were assessed for their inter-rater and internal consistency reliabilities. These four variables: participants' childhood abuse history, drug and alcohol usage, cognitive beliefs about children and child-rearing, and insight into parenting, are later examined for their relationship with parenting behaviour (that is, Behavioural Responsiveness) and the other parental empathy variables (see Chapter 9: Mediation Potential of Parental Empathy, p. 161).

### **Inter-Rater Reliability**

An assessment of the PEM's inter-rater reliability was a clear priority for the present study for two reasons. First, the PEM is a semi-structured interview. Semi-structured interviews which incorporate open-ended questions are potentially more



vulnerable to subjective interpretation of responses than more rigid assessment tools such as paper and pencil surveys. Second, the PEM is proposed for use as a clinical tool in the assessment of child maltreatment. The nature of clinical practice in child maltreatment is such that inaccurate diagnoses, whether falsely negative or falsely positive, carry profoundly deleterious consequences for both children and families. Thus, to be useful, the PEM must demonstrate satisfactory reliability as an objective assessment measure.

Ten interview schedules were randomly selected from the three parent groups (3 from controls, 3 from distressed and 4 from abusive parents). A registered psychologist, experienced in working with families (both maltreating and distressed), served as the independent scorer. Basic instruction (3 hours duration) on the Parental Empathy Model and the scoring protocols was provided by the researcher. Inter-rater reliability (scorer variance) was assessed by correlating the total scores obtained by the researcher and the independent scorer for each of the variables in the ten interviews (Anastasi & Urbina, 1997). The observed agreement levels of scores, expressed in percentages, between the author and the independent scorer for the four empathy variables were 95% (Signals), 90% (Attributions), 85% (Emotion) to 98% (Behaviour). Inter-rater reliability for the Faking Good subscale was excellent at 100%.

Of the other supplementary four variables, Drug & Alcohol Use achieved the highest agreement across raters of 95%, followed by Childhood History (of abuse) with an inter-rater reliability of 90%. Insight achieved an inter-rater reliability of 88%. The variable Beliefs was not assessed for inter-rater reliability as its objective scoring procedure rendered this unnecessary.

Additionally, for the ten interview schedules, each rater independently achieved one hundred percent correct classification of abuse status. That is, both raters successfully allocated each participant to the abusive, distressed and non-abusive categories.

In summary, the PEM achieved highly satisfactory levels of inter-rater reliability in this initial assessment. However, further study with more raters is called for.

### **Internal Consistency**

Internal consistency refers to the homogeneity of the items in a scale (Anastasi & Urbina, 1997). As such, internal consistency measurements provide an estimate of the extent to which the scale items assess the same construct. Given the relatively small sample size, the estimates of internal reliability were calculated for abusive, distressed, and control participants in combination.

The internal consistency reliability for the variables of interest were assessed by Cronbach's alpha. The essential aspects of the results of these assessments are described below, and summarised in Table 2. (Further reliability information, including item-total correlations, is provided in Appendix R: Reliability Tables; variable code names and descriptions are explained in Appendix S: Variable Coding and Computation Tables. Means and *SDs* for each variable scale are summarised in Appendix Q: Group and Sample Means for Computed and Other Variables, according to risk-status group and overall sample.)

The variable Signals, with 14 items, had good reliability with a standardised item alpha of .83. Complete data were available for all 103 cases. Only two items (sg\_3oth &

sg\_560th), if deleted, increased alpha levels, but these increases were negligible at .005 and .006, respectively.

The variable Attributions, with 12 items, had good reliability with a standardised item alpha of .79. Complete data were available for 92 cases. Only one item (att\_sc4) if deleted would increase alpha levels and this increase was negligible at .001.

The variable Behaviour (unstressed), with 16 items, had very good reliability with a standardised item alpha of .84. Complete data was available for 60 cases. Only one item (b\_op15a), if deleted, increased the alpha level. This increase at .01 was a very minor level.

The variable, Behaviour (stressed), with 10 items, had good reliability with a standardised item alpha of .79. Complete data were available for 91 cases. Only one item (b\_sc1c), if deleted, increased the alpha level by a mere .004.

The variable Behaviour (overall), with 2 items (that is, behaviour unstressed and stressed) had good reliability with a standardised item alpha of .78. Data were available for all cases (103) for the reliability test as averages were used to calculate both variable subscales. As stated earlier, this procedure avoided the exclusion of cases when a small number of subscale items were missing.

The variable, Emotion (unstressed), with 13 items, had good reliability with a standardised item alpha of .74. Complete data were available for 91 cases. No individual item could be excluded to improve the alpha levels.

The variable Emotion (stressed), with 10 items, had good reliability with a standardised item alpha of .76. Complete data were available for 91 cases. Only one item (em\_sc9d), if deleted, would increase the alpha level, by only .001.

The variable, Emotion (overall) with 2 items (that is, emotion unstressed and stressed) had good reliability with a standardised item alpha of .73. Data were available for all cases for the reliability test as averages were used to calculate both variable subscales.

The total empathy score (PEM Total), aggregating all 6 variables had very good reliability with a standardised item alpha of .89. Data were available for all cases for the reliability test as averages were used to calculate variable subscales (items) if subscale items were missing. Only one item (Emotion - stressed), if deleted, would increase the alpha level by the negligible extent of .004.

Insight, with 10 items, had good reliability with a standardised item alpha of .86. Complete data were available for all cases. Only one item (ins\_23c), if deleted, would result in an increased alpha level by a very minor level of .007.

The variable Beliefs, with 13 items, had fair reliability with a standardised item alpha of .59. Complete data were available for all cases. Four items (belief1, belief7, belief10 & belief12), if deleted, would increase alpha levels. Those increases were at relatively minor levels of .016, .009, .008 and .026, respectively, in each instance. Notably, these items were those that are reverse scored, indicating a possible response bias.

Childhood Abuse, with 2 items, had good reliability with a standardised item alpha of .73. Complete data were available for all cases.

Alcohol & Drug Use, with 3 items, had relatively poor reliability with a standardised item alpha of .41. Complete data were available for all cases. One item (drug\_39d - problematic heroin use), if deleted, resulted in an increased alpha level. That

increase was at a minor level of .03. Notably, only three participants admitted to any heroin use, and only one of those rated their use as such to qualify as a serious concern.

Faking Good scores, with 30 items, had very good reliability with a standardised item alpha of .93. Complete data were available for 94 cases. Four items (fg\_op10, fg\_op23, fg\_op25 & fg\_op36), if deleted, would increase alpha levels. Those increases were at relatively minor levels of .002, .004, .005 and .004, respectively, in each instance.

A summary of the internal consistency reliability for each component of the PEM is presented in Table 2.

**Table 2: Reliability statistics for variable scales**

Variable	No. items	No. cases*	<i>M</i>	<i>SD</i>	Cronbach's Alphas	
					Alpha	Standardised item alpha
Attention to signals	14	103	29.7	14.7	.84	.83
Attributions	12	92	38.2	7.9	.80	.79
<u>Behaviour:</u>						
unstressed	16	60	52.6	10.0	.84	.84
stressed	10	91	26.4	6.6	.80	.79
(overall)	2	103	5.8	1.1	.77	.78
<u>Emotion:</u>						
unstressed	13	91	35.6	6.2	.74	.74
stressed	10	91	19.9	4.5	.75	.76
(overall)	2	103	4.6	.9	.73	.73
Total PEM	6	103	.0	4.8	.89	.89
Insight	10	103	9.9	5.5	.86	.86
Beliefs	13	103	8.7	2.4	.63	.59
Childhood Abuse	2	103	.0	1.8	.73	.73
Alcohol & Drug Use	3	103	.3	.7	.42	.41
Faking Good	30	94	9.2	8.0	.92	.93

\* Only cases with complete item sets were used to calculate scale reliabilities. Actual variable scores were calculated as averages of available items for cases with missing values.

The relationships between the individual PEM scales were examined for the expected levels of interactions between these variables according to the conceptual model constituting the the general construct of parental empathy (see Table 35, Appendix U). In general, moderate to strong positive correlations were found between the four empathy variables (.45 to .82). The other four supplementary scales demonstrated a varied pattern of correlations with the empathy variables. Beliefs showed moderate to strong positive correlations with the four empathy variables (.52 to .72) and Insight had weak to strong correlations with the four empathy variables (.34 to .74). Childhood Abuse was not correlated with Attention to Signals, but had weak negative correlations (-.21 to -.36) with the other three empathy variables. Alcohol & Drug Use was not correlated with any of the four empathy variables.

### Discussion

Overall, the assessment of inter-rater and internal consistency reliabilities provides strong initial support for the reliability of the PEM. In particular, the parental empathy variables performed well with very good inter-rater reliability and good to very good internal consistency reliabilities. Of note is that the total parental empathy score (PEM total) also demonstrated very good internal consistency and that identification of abuse status was 100% accurate for both raters. These results provide early support for the view that the PEM may be of diagnostic value in child maltreatment work. Further, the results provide support for the utility of the proposed model of parental empathy in child maltreatment risk assessment.

As an instrument designed to assess child maltreatment risk, the PEM's ability to detect positively biased responses is essential. The Faking Good scale of the PEM was found to have excellent inter-rater reliability and very good internal consistency. This finding provides initial affirmation that the faking good construct as assessed by the scale items is well-defined and cohesive. A further study with a large number of participants is necessary in order to identify appropriate cut-off points on the Faking Good scale that would justify the rejection of a participant's PEM profile.

Of the other four variables assessed for their reliability, Alcohol and Drug Use demonstrated the poorest internal consistency although inter-rater reliability was acceptable. One possible explanation for the poor internal reliability of this variable is that this result reflects a pattern of habitual usage of a preferred drug rather than poly-drug use. Additionally, the number of participants who reported problematic drug and alcohol usage was small. Given this possibility, the variable was retained for further assessment.

Insight achieved good levels of both inter-rater and internal consistency reliability, indicating that the subscale is assessing an homogeneous domain. This is an important finding given that previous papers on the role of insight in parenting have largely been of a qualitative nature and thus failed to define the variable in quantitatively testable terms (e.g., Fischhoff, Whitten, & Pettit, 1971; Webster-Stratton & Herbert, 1993).

The Childhood Abuse index achieved an adequate level of internal consistency reliability and a good inter-rater reliability score. Internal consistency reliability tends to be stronger the higher number of items included in the scale (Milner, 1986). As such, given that this variable comprised only two items the reliability scores achieved are encouraging.

In summary, the empathy variables were found to be measured in a reliable manner with high levels of internal consistency and inter-rater agreement. Of the supplementary variables, Insight and Beliefs were also found to have highly satisfactory levels of reliability. Thus, only the two variables, Drug & Alcohol Use and, to a lesser extent, Childhood Abuse, were found to have less than highly satisfactory levels of internal consistency and inter-rater agreement.

Unfortunately, no comment could be made in regard to the test-retest reliability of the PEM. An assessment of this feature of the PEM is a priority for the future development of the instrument. Another possible limitation of the current assessment of the PEM's reliability is the reliance on only two raters (one of whom was the author) for the assessment of the inter-rater reliability of the instrument. Additionally, this second rater was a psychologist experienced in working with families characterised by difficulties including child maltreatment. Whether a less experienced psychologist, social worker or professional case-worker would achieve a similar level of reliability is unknown. Further, the small sample size used in the current assessment limits the interpretation of the findings to the cautious statement that initial testing was encouraging. Clearly several assessments with large numbers of participants will need to be completed before any definitive comments are able to be made about the PEM's reliability.

Notwithstanding these considerations, the results found indicate that the PEM is a highly reliable measure of parental empathy.