

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]



**ASSESSING EMOTIONAL DISTRESS IN ABUSED CHILDREN DURING
VIDEOTAPED INVESTIGATIVE INTERVIEWS:
THE EFFECT OF RISK AND PROTECTIVE FACTORS**

by

Susan Rebecca Hall

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PSYCHOLOGY

**In Partial Fulfillment of the Requirements
For the Degree of**

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2 0 0 1

UMI Number: 3010246

UMI[®]

UMI Microform 3010246

Copyright 2001 by Bell & Howell Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company

300 North Zeeb Road

P.O. Box 1346

Ann Arbor, MI 48106-1346

THE UNIVERSITY OF ARIZONA ©
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Susan Rebecca Hall

entitled Assessing emotional distress in abused children
during videotaped investigative interviews:
The effect of risk and protective factors

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

Judith V. Becker

Judith V. Becker

5-26-00

Date

Bruce D. Sales

Bruce D. Sales

5-26-00

Date

Gary E. Schwartz

Gary E. Schwartz

5-26-00

Date

Date

Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Judith V. Becker
Dissertation Director Judith V. Becker

5/26/00
Date

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: _____

A handwritten signature in black ink, written over a horizontal line. The signature is cursive and appears to read "James R. Fawcett".

ACKNOWLEDGEMENTS

This research project could not have been conducted or completed without the assistance and support of some wonderful people and organizations, not all of whom can be mentioned here. First, I would like to thank the members of my dissertation and masters committees for their time, thoughtful comments, and guidance over the years: Judith Becker, Bruce Sales, Joel Dvoskin, Roger Levesque, Mary Koss, and Gary Schwartz. Special thanks to Bruce and Judith for your work with me and for teaching me what it means to be a mentor. Also, I want to express my appreciation to other members of the University of Arizona family who provided me with statistical consultation for this research program: John Allen, Lee Sechrest, Jim Coan, Patrick McKnight, and Tim Stickle.

I would also like to acknowledge the tremendous support and dedication of the research assistant members of my Child Abuse Investigation Lab: Fernanda Martinez, Martha Nunez, Alison Rivchun, Julie Seligson, Erin Mangurten, Douglas Preble, Benjamin Galarneau, Amy Marshall, Wilson Butz, Jason Contreras, Timothy Geiger, and Danielle Krause. In addition, I would like to extend a very special thank you to my favorite research assistant and loving, supportive partner and husband, Robert Scholz. Thank you also to my family and friends who have been there for me throughout this journey, especially Sarah Jewett, Daniel Krauss, and Lori Markson.

Next, I want to express my gratitude to the American Academy of Forensic Psychology, and the University of Arizona's Social and Behavioral Sciences Research Institute, for each awarding me a dissertation grant for this project. I would also like to thank the Arizona Children's Justice Task Force members, including Gene Siegel, and the American Psychology-Law Society for their early grant support of this research program.

Last, but not least, I would like to thank the professionals at the Center for granting me permission to conduct this research, as well as thank the children who participated in the investigative interviews sampled in the present study. This project is dedicated to you.

DEDICATION

This dissertation is dedicated to the children and families whose lives are affected by violence and abuse. May we continue to learn from their experiences and work together to build a future free of child maltreatment – one in which studies such as this are unnecessary.

TABLE OF CONTENTS

LIST OF TABLES	8
LIST OF FIGURES	9
ABSTRACT	10
INTRODUCTION	12
Negative Effects of Child Abuse	13
Mediating Variables	14
<u>Demographic Variables</u>	14
<u>Abuse-Specific Variables</u>	15
<u>Systemic Variables</u>	17
<u>Family dysfunction</u>	17
<u>Lack of social support</u>	18
<u>Secondary traumatization</u>	19
<u>Effects of disclosure and investigation, and policy changes</u>	19
<u>Effects of the legal system, and policy changes</u>	23
<i>Pre-trial stage</i>	24
<i>Trial stage</i>	25
<i>Post-trial stage</i>	30
Moderating Variables	32
<u>Resiliency and Protective Factors</u>	32
<u>Coping Factors</u>	33
Summary	34
The Present Study	36
<u>Hypotheses</u>	38
 METHOD	 45
Subjects	45
Coding Procedures	46
<u>Materials</u>	46
<u>VITAL Coding Procedures</u>	48
Statistical Analyses	50
<u>Interrater Reliability</u>	50
<u>VITAL Data Reduction</u>	51
<u>Hypotheses Testing</u>	56

RESULTS	58
Interrater Reliability	62
Concurrent Criterion Validity of Child and Interview VITAL Ratings	63
The Effect of Risk Factors on VITAL Ratings	76
<u>The Effect of Demographic Risk Factors</u>	76
<u>The Effect of Abuse-Specific Risk Factors</u>	81
<u>The Effect of Systemic Risk Factors</u>	86
The Effect of Protective Factors on VITAL Ratings	90
DISCUSSION	94
Interrater Reliability	94
Concurrent Criterion Validity of Child and Interview VITAL Ratings	94
The Effect of Risk and Protective Factors on VITAL Ratings	102
APPENDIX A: VITAL Record Review Form	109
APPENDIX B: Child VITAL Operational Definitions and Coding Procedures	110
APPENDIX C: Interviewer VITAL Operational Definitions and Coding Procedures	113
REFERENCES	136

LIST OF TABLES

TABLE 1.1, Proportion of time in original child VITAL categories.....	116
TABLE 1.2, Proportion of time in modified child VITAL categories.....	118
TABLE 1.3, Proportion of observed time in 2-category child VITAL	120
TABLE 1.4, Proportion of observed time in 5-category child VITAL	121
TABLE 2.1, Proportion of observed time in original interviewer VITAL categories.....	122
TABLE 2.2, Proportion of observed time in 4-category interviewer VITAL.....	125
TABLE 2.3, Proportion of observed time in interviewer question-type categories ..	126
TABLE 2.4, Proportion of observed time concerning the alleged abuser	127
TABLE 3.1, Child gender by alleged abuse type	128
TABLE 3.2, Child age groups and gender by case outcome and alleged abuse type.....	128
TABLE 3.3, Case outcome by child age groups	128
TABLE 4.1, Global interrater reliability	129
TABLE 4.2, Specific interrater reliability for child VITAL categories	130
TABLE 4.3, Specific interrater reliability for interviewer VITAL categories.....	133
TABLE 5.1, Criterion validity results.....	134
TABLE 5.2, Risk and protective factor results.....	135

LIST OF FIGURES

FIGURE 1, Mean 5-Category Child VITAL ratings	64
FIGURE 2, Mean Interviewer VITAL Question-Type ratings	65
FIGURE 3, The effect of age on the proportion of closed question-types	65
FIGURE 4, Time in rapport by 3-age groups of children	66
FIGURE 5, Child scared ratings by case outcome	67
FIGURE 6, The proportion of personal interviewer questions by case outcome.....	69
FIGURE 7, The proportion of alleged abuser interviewer questions by case outcome	69
FIGURE 8, Child angry ratings by type of alleged abuse and case outcome	70
FIGURE 9, Child internalizing ratings by personal abuse questions	72
FIGURE 10, Child internalizing ratings by personal abuse questions and case outcome	72
FIGURE 11, Child externalizing ratings by personal abuse questions.....	72
FIGURE 12, Child externalizing ratings by personal abuse questions and outcome.	73
FIGURE 13, Child internalizing ratings by alleged abuser questions and case outcome	73
FIGURE 14, Child externalizing ratings by alleged abuser questions	74
FIGURE 15, Child externalizing ratings by alleged abuser questions and outcome..	74
FIGURE 16, Child internalizing ratings by general questions.....	75
FIGURE 17, Child internalizing ratings by general questions and case outcome	75
FIGURE 18, Child externalizing ratings by general questions	75
FIGURE 19, Child externalizing ratings by general questions and outcome	76
FIGURE 20, The proportion of externalizing by child gender	77
FIGURE 21, The proportion of externalizing by child gender and age	77
FIGURE 22, The proportion of internalizing by child gender and age.....	78
FIGURE 23, The proportion of anger by child age.....	79
FIGURE 24, The proportion of anger by child age and gender	79
FIGURE 25, The proportion of positive by child age and gender.....	80
FIGURE 26, The proportion of internalize by frequency of alleged abuse.....	81
FIGURE 27, The proportion of externalize by frequency of alleged abuse.....	81
FIGURE 28, The proportion of externalize by duration of alleged abuse	82
FIGURE 29, The proportion of internalize by duration of alleged abuse.....	83
FIGURE 30, The proportion of uncomfortable by duration of alleged abuse.....	84
FIGURE 31, The proportion of scared by duration of alleged abuse.....	84
FIGURE 32, The proportion of internalizing by social support.....	88
FIGURE 33, The proportion of uncomfortable by social support	88
FIGURE 34, The proportion of externalizing by social support	89
FIGURE 35, The proportion of internalizing by non-offending parental support.....	91
FIGURE 36, The proportion of depressed by non-offending parental support	92
FIGURE 37, The proportion of uncomfortable by non-offending parental support .	92

ABSTRACT

Social science research shows that many maltreated children suffer adverse psychological consequences, but less evidence exists to identify which children may be more or less at risk for such effects. In addition, some negative effects may be compounded when abused children interact with the criminal justice system. While public policies and procedures have been modified to address these issues (e.g., courtroom accommodations), little research has assessed which children would benefit most from them. To fill these needs, the present study (a) tested the concurrent criterion validity of a new clinical-forensic assessment tool, the Videotaped Interview Trauma Assessment List (VITAL), a behavioral observation instrument that, when fully validated, can be used with other methods to assess emotional distress, risk/protective factors, and need for courtroom accommodations in allegedly abused children, and (b) assessed how risk and protective factors affected allegedly abused children's emotions and behaviors during actual investigative interviews.

The results from the present study lend support to the concurrent criterion validity of the VITAL, because it functionally measured the children's existing emotional state and assessed the impact of interviewers' questions/statements on children's emotions/behaviors. Results were consistent with prior research on investigative interviewing and risk/protective factors. Specifically, the investigative interviews were generally an uncomfortable experience for the children; however, the presence of one or more sources of social support resulted in statistically significantly lower internalizing ratings (i.e., sadness, anxiety, embarrassment, withdrawal). Similarly, non-offending

parental support tended to result in lower internalizing and depression ratings. While ratings of children's internalizing or externalizing emotions/behaviors were not affected by the majority of risk factors proposed in prior research (i.e., higher frequency of abuse, threat or use of force, closer relationship between the alleged abuser and the child, higher degree of intrusiveness of sexual abuse, less time in rapport, and family dysfunction), some trends were indicated (e.g., longest duration of alleged abuse tended to result in the most externalizing behavior; higher number of investigative interviews tended to result in higher internalizing and lower positive ratings). Future validity research on the VITAL is recommended because behavioral instruments are virtually absent in psycho-legal child maltreatment research and practice.

INTRODUCTION

There is general consensus in the psychological literature that many child victims of abuse suffer adverse consequences. However, evidence is less available to help identify which children may be more or less at risk for such negative effects. In addition, negative effects may be produced or compounded when abused children interact with the criminal justice and child protective systems. As a result, such children may be less able or willing to participate in the investigation and prosecution of their alleged offenders. While modified public policies and procedures are currently promoted to address these issues (e.g., non-leading investigative interviews; courtroom accommodations for children), little research has been done to assess which children would benefit most from them. To fill these needs, the purpose of this paper is twofold: (a) to determine how risk and protective factors, including those associated with secondary traumatization, affect allegedly abused children's emotions and behaviors during actual investigative interviews, and (b) to test the criterion validity of a new clinical-forensic assessment tool, the Videotaped Interview Trauma Assessment List (VITAL). The VITAL is a behavioral observation instrument that, when fully validated, can be used in conjunction with other methods to assess emotional distress, risk and protective factors in allegedly abused children, and evaluate their need for courtroom accommodations.

The first section describes the commonly reported negative effects of abuse on children, and the various mediating and moderating variables that may affect their emotional and psychological responses to the abuse. Special attention is given to the risk factors/mediating variables that have been associated with the criminal justice system's

response to such children, and to the policy changes intended to ameliorate these children's negative experiences. This section will conclude by describing the hypotheses of the present study. The next two sections will explain this study's methods and describe its results. The final section will discuss the results, including the public policy implications of this project and suggestions for future research.

Negative Effects of Child Abuse

Child abuse is recognized as a public health problem (Chadwick, 1999; Kaplan, Pelcovitz, & Labruna, 1999; Kolko, 1992) because it has wide-ranging negative effects on its victims and on the communities in which they live. Previous research has shown that all types of child maltreatment (i.e., sexual abuse, physical abuse, emotional abuse, neglect) are associated with many different negative effects on child victims/survivors (Becker et al., 1995; McGee, Wolfe & Wilson, 1997) that extend into adolescence and beyond (Briere, 1992; Kelley, Thornberry, & Smith, 1997). In a recent article (Hall, 1999), the empirical, clinical, and scientific support for six main short- and long-term problem areas associated with child maltreatment were reviewed. Specifically, across all types of abuse, some children have been found to suffer from problems with: (a) post-traumatic stress symptoms and PTSD, (b) dissociative symptoms and disorders, (c) altered social cognitions about themselves and the world, (d) altered internalization and externalization involving depression, anxiety, fear, and somatization, (e) impaired self-reference due to insecure attachments, and (f) disturbed patterns of relating with others, such as acting out sexually and aggressively, which often lead to legal and academic problems (Becker et al., 1995; Briere, 1992).

Of these abuse-related problem areas, only posttraumatic stress symptoms and unusual sexual behavior appear to differentiate sexually abused children from children in general clinical populations (Kendall-Tackett, Williams, & Finkelhor, 1993). Still, no single item or cluster of items can be used to definitively determine whether or not a child has been sexually abused (Friedrich, 1995).

In addition, not all abused children and adolescents develop emotional and behavioral problems. Whether an abused child experiences negative effects may depend on a series of mediating and moderating variables, and their interactions. However, very little is known about the role of such risk and protective factors in the development of symptomology in children (Feiring, Taska, & Lewis, 1998).

Mediating Variables

Several categories of risk factors have been identified that potentially heighten the likelihood of abused children suffering short- and long-term negative effects. These include demographic variables, abuse-specific variables, and systemic variables, including secondary traumatization.

Demographic Variables

Some reviews of the literature have found that greater negative short-term effects of abuse have been associated with the child's age and gender (Becker et al., 1995). Specifically, younger maltreated children may be more vulnerable to trauma than older children, and sexually abused girls at more risk for emotional distress than boys. However, other studies and reviews examining the risk factors of age and gender report mixed or

nonsignificant findings (Davenport, Browne, & Palmer, 1994; Foy, Madvig, Pynoos, & Camilleri, 1996; McClellan et al., 1995; Williams, 1993).

Demographic variables may have less effect on long-term abuse-related problems. For example, a common and debilitating long-term posttraumatic response is posttraumatic stress disorder (PTSD) (Carlson, 1997). Recent investigations have shown that children who have suffered severe or chronic maltreatment are at significant risk for developing PTSD (Famularo et al., 1994; Foy et al., 1996). However, abused children's age, gender, ethnicity, and intelligence have not been consistently associated with posttraumatic stress or other negative symptomology (Foy et al., 1996; Williams, 1993).

Continued research is needed, however, because some demographic variables that may affect outcome are rarely examined in research with abused children and adolescents. In a recent review of the literature, only one study was located that explicitly studied the effects of ethnicity on the long-term effects of abuse. Rao, DiClemente, and Ponton (1992) reviewed records of substantiated sexual abuse cases, and compared random, matched samples of Asian victims with African American, Caucasian, and Hispanic victims. The researchers found that Asian victims were more likely to express suicidality, but were less likely to express anger and sexual behavior problems.

Abuse-Specific Variables

Both short- and long-term negative effects have been associated with certain aspects of the abusive incident(s). Greater short-term effects of abuse have been linked to abuse that involved higher degrees of force, more intrusiveness, greater frequency, longer duration, larger numbers of perpetrators, shorter amount of time since the incident, and

closer relationships between the child and the perpetrator (Becker et al., 1995). In addition, some clinicians have posited that “distinctive” abuse events will cause a child to experience overwhelming fear, thus putting the child at risk for longer-lasting traumatization (Carlson, Furby, Armstrong, & Shales, 1997). Long-term effects of sexual abuse have been associated with similar characteristics of the abusive incident(s), including the use or threat of force, abuse that involved penetration, abuse by a father or stepfather, and abuse of longer duration (Becker et al., 1995).

While the vast majority of research on abuse risk factors has explored characteristics of abusive incidents as predictors of emotional distress in children, only a modest amount of variability among children appears to be attributable to these variables (Chaffin, Wherry, & Dykman, 1997). Thus, it is not surprising to find that studies continue to report conflicted results regarding abuse-specific factors. For example, the risk factors of greater frequency and duration of abuse have been disputed. Some commentators have stated that subjecting a child to multiple incidents of abuse over a period of time may negatively affect the nature and the duration of their reactions to the abuse (Nader, 1997). Yet, others have found a lack of empirical evidence that repeated exposure is more traumatic (Foy et al, 1996; Williams, 1993). Other studies measuring duration of abuse revealed mixed findings (Williams, 1993). In addition, a review of 25 recent studies on childhood PTSD revealed that the length of time since trauma exposure was negatively correlated with PTSD symptomology (Foy et al., 1996).

There are a number of reasons to explain the variability of abuse-specific risk factors. One reason may be due to measurement error. Abuse-specific factors are not

consistently measured or consistently included as variables in research. Another reason for conflicting results may be explained by the interaction of these variables with other risk factors and child characteristics. For example, the lives of many children may be affected by more than one type of abuse (McGee, Wolfe & Wilson, 1997). Recognition of the high co-morbidity rates in the types of abuse suffered by these children is important because co-morbidity has been found to increase the risk for psychological problems (Ackerman, Newton, McPherson, Jones, & Dykman, 1998). Thus, more research is needed on abuse-specific risk factors. Without additional empirical support, these variables should not be considered causal markers of emotional distress or trauma in abused children and adolescents.

Systemic Variables

Compared to demographic and abuse-specific variables, broader family and community systems may have a greater impact on abused children's mental health. This subsection will discuss three types of systemic variables that have been shown to increase abused children's risk for negative effects: family dysfunction, lack of social support, and secondary traumatization due to abuse disclosure, investigation and legal system interactions (e.g., testifying in court).

Family dysfunction

Various family dynamics have been associated with child maltreatment, including poor parent-child relationships, more frequent severe physical punishment, conflict between parents, and disrupted relationships with biological parents (Brown, Cohen, Johnson, & Salzinger, 1998). Sexual abuse has been found to occur with greater

frequency in two types of families (a) one characterized by dysfunctional interactions in which children are also emotionally abused, neglected and often physically abused, and (b) the other in which the biological father is no longer present, and where a stepfather or cohabitee has access to children in the family (Bagley & Thurston, 1996). Such family dysfunction has been associated with long-term negative effects in child abuse victims (Becker et al., 1995).

Lack of social support

In addition, it has been strongly suggested that the traumatic impact of abuse is mediated by the child's psychological reaction to the abuse (Foy, Madvig, Pynoos, & Camilleri, 1996; Kendall-Tackett, Williams, & Finkelhor, 1993; Wolfe, Sas, & Wekerle, 1994). A child's cognitive appraisals are often affected by his or her perceptions and experiences of social support from caretakers and others in the community (Galatzer-Levy, 1994; Kelly, Thornberry, & Smith, 1997). Social support has been defined as a socio-emotional variable that can be conceptualized as a form of social interaction or communication that engenders a feeling of well-being in the target person (Burlison, Albrecht, Goldsmith, & Sarason, 1994). For example, many abusers tell children that they caused or were at fault for the abuse, and many allegations are not believed by families or substantiated by investigators. Children who internalize blame and responsibility for their abuse have shown more negative outcomes than those who adopted more external attributions (Amaya-Jackson & March, 1995; Inderbitzen-Pisaruk, Shawchuck, & Hoier, 1992; McMillen & Zuravin, 1997; Morrow, 1991). Not being believed or supported by parents or caretakers may result in more emotional trauma (Davenport et al., 1994). In

addition, children who perceived that they did not receive just or fair treatment may suffer more negative effects than those who believed that they were treated more fairly (O'Donoghue & Elliott, 1991). Indeed, negative social cognitions and lack of social support have been identified as potential predictors of long-term negative effects in child victims/survivors of abuse (Becker et al., 1995).

Secondary traumatization

Finally, child victims of abuse who interact with the criminal justice system may experience additional negative effects or an increase in their existing abuse-related problems (Hall, 1995; Whitcomb, 1993). Referred to as secondary victimization (Lipovsky, 1994) or secondary traumatization (Whitcomb, 1993), such iatrogenic harm is generally caused by aspects of the abuse disclosure and investigation (e.g., multiple interviews) (Jones, 1991), and/or prosecution (e.g., multiple times testifying) (Goodman et al., 1992). The rise in the number of child sexual abuse reports makes the potential for such system-induced trauma an acute problem (Lipovsky, 1994).

Effects of disclosure and investigation, and policy changes. Mental health, child protective, law enforcement, and legal professionals possess the power to influence the effect that systemic intervention has on abused children (Davies & Seymour, 1997; Henry, 1997). Positive or negative effects can start as soon as the child discloses an alleged abusive experience to a trusted person or when that child is first asked to talk about an alleged abusive incident to a professional.

In general, the literature is divided on the effect that disclosing traumatic events has on children and adolescents. The majority of the research has focused on disclosure of

sexual abuse, and some researchers have found that children who disclose sexual abuse to receptive adults have decreased levels of distress and physical symptoms (Arata, 1998). If abuse disclosure is made to a professional, such disclosure may have positive effects on children if trust between the child and professional is established (Henry, 1997). Trust and rapport may also enable professionals (e.g., prosecutors, law enforcement officers, child protective service workers) to obtain more accurate abuse-related information from the children (Saywitz & Camparo, 1998).

However, research has also shown that disclosure may have negative effects and increase symptomology in children (Arata, 1998). Negative effects may occur if professionals are not educated about child development and child sexual abuse (Sattler, 1998), or are otherwise insensitive to the child and unprofessional. Thus, it is unclear whether greater distress is a precipitant of disclosure, a result of disclosure, or a result of intervention. Future research is needed in this area, especially on how disclosure of different types of abuse interacts with later interventions, and how children's functioning is affected.

After disclosure, subsequent system intervention, or the lack thereof, has the potential to adversely affect some allegedly abused children. Such interventions include the investigatory process, the decision whether or not to remove the child from the home, and the provision of services to the child and family. Indeed, the nature of intervention in some cases has the potential to be more distressing than the abuse itself (Berliner & Conte, 1995).

Some studies have found that younger children have reported more stress during the investigation stage than in the courtroom stage (Lipovsky, 1994 [ages not reported]; Lipovsky & Stern, 1997). In particular, investigative interviews are one aspect of the investigatory process that has been found to be a source of potential harm. Investigative interviews are conducted by a neutral member or members of the local multidisciplinary or interdisciplinary team (MDT or IDT), usually a trained law enforcement agent, child protective service worker and/or forensic interviewer, to investigate allegations of child abuse. Investigative interviews differ from therapeutic or psychosocial interviews, which are designed to evaluate a child's need for treatment and are usually conducted by mental health professionals. Investigative interviews have the potential to increase the risk of negative effects in abused children and adolescents in two ways.

First, the quality of investigative interview practices may negatively affect children and impair the reliability of the information provided to professionals. Specifically, the rapport phase of the investigative interview is considered crucial because poor rapport can have harmful effects on the child (Jones & McQuiston, 1986). For example, high anxiety levels can lead to lower cognitive functioning (Saywitz & Camparo, 1998) and thus lower the quality of information given in the interviews. Prior studies of interviewer-provided social support has generally demonstrated that interviews conducted in a friendly, supportive manner improves the accuracy of children's statements because it increases their resistance to misleading questions (Carter, Bottoms, & Levine, 1996; Davis & Bottoms, in press; Goodman, Bottoms, Rudy, & Schwartz-Kenny, 1991; but see Imhoff &

Baker-Ward, 1999). Yet, the determination of how stressful investigative interviews are for abused children is an empirical question yet to be answered (Lipovsky, 1994).

To improve interviewing practices with children, protocols and professional guidelines have been developed (e.g., AACAP, 1997; APSAC, 1990, 1998; Sattler, 1998). Guidelines and standards for conducting investigative interviews in child abuse cases may not only result in improving the likelihood of obtaining an accurate history from children (Vogeltanz & Drabman, 1995; White & Edelstein, 1991), but also may protect children by providing a less stressful interview environment (Westcott, Davies, & Spencer, in press). However, while a consensus in the literature has developed about the basic structure of child abuse investigative interviews (Sternberg, Lamb, Esplin, & Baradaran, 1999), many professionals in the field are not aware of or do not follow interviewing guidelines and recommended procedures (Oberlander, 1995).

Second, multiple investigative interviews have been found to be associated with an increased level of trauma in abused children (Bala, in press; Goodman, Quas, Bulkley, & Shapiro, in press; Henry, 1997). To reduce the number of unnecessary and improperly conducted interviews of children, many communities have implemented multidisciplinary teams, interagency councils, coordinated law enforcement and CPS abuse investigations, and child advocacy centers. Outcome data supports the efficacy of these procedural changes in reducing the number of interviews with children and adolescents, improving collaboration among professions, and strengthening legal cases against alleged abusers (for a review of the literature and research on multidisciplinary teams, see Hall, 1997; also see Wolfe, Sas, & Wilson, 1987).

Effects of the legal system, and policy changes. The other major area that poses potential secondary harm for abused children is the legal system. However, it is important to note at the outset that while involvement with the legal system can be stressful for some children, it may be therapeutic for others (Cecchetti-Whaley, 1992; Melton, 1992). Participation in their legal cases may enhance children's sense of control over the process (Small & Melton, 1994; Wolfe, Sas, & Wilson, 1987), and may play a role in their healing and recovery (Henry, 1997). Positive effects have been demonstrated in cases where the child was given some control over the process (Lipovsky, 1994), was asked about his/her feelings during the case (Murray, 1995), or was referred to counseling before or after court appearances (Davies & Seymour, 1997).

Certain aspects of the legal process may be traumatic or stressful for some abused youth. These aspects include: (a) longer length of and delays in the legal process (Davies & Seymour, 1997; Runyan et al, 1988); (b) lack of knowledge about the legal system and its processes and procedures (Saywitz, 1989); (c) negative treatment of the child and family by legal actors; and (d) participation of the child in non-criminal (Runyan et al, 1988) and non-trial proceedings (Lipovsky, 1994).

The majority of research on secondary traumatization has focused on the effects of criminal court involvement on abused children and adolescents. Children may encounter secondary trauma in at least three stages: (a) pre-trial anticipation, including waiting to be called into court to offer testimony, (b) during the trial, when giving face-to-face testimony, and (c) post-trial, when waiting for decisions (e.g., the verdict, decisions about placement).

Pre-trial stage

First, it should be recognized that negative effects from trial anticipation may not, in practice, affect many children. Few reported cases of child sexual abuse ultimately involve a child giving testimony, due to substantial attrition of cases and the resolution of cases by plea bargaining (Lipovsky, 1994). Notwithstanding, some children may not know about or be told about the status of their case, and thus wrongly assume they will testify. The prospect of future court testimony is usually not discussed during early stages of criminal justice system involvement, such as during the child's investigative interview(s). In addition, children may not understand why they might have to testify. In one study of child witnesses, significant concern was evoked by the children when anticipating the court experience, in part, because of their inaccurate views (Berliner & Conte, 1995). It has been hypothesized that accurate knowledge about the legal system provides children with a frame of reference in which they can make sense of their experiences, put their fears into perspective, and improve their performances as witnesses (Saywitz & Snyder, 1993).

However, the extent to which abused children's anticipatory distress is a contributory factor to mental health problems is unknown. Despite the lack of research in this area, improvements in court policies and procedures have been introduced to correct some of the perceived problems associated with anticipatory distress (Lipovsky, 1994). Some of these innovations that target children prior to actual court involvement include court education and preparation projects (Sas et al., 1993; Wolfe, Sas, & Wilson, 1987), the use of support persons (Davies & Seymour, 1997), and the use of improved, child

friendly waiting rooms (Davies & Seymour, 1997). Future research is needed to evaluate the effectiveness of and continued need for these policy changes.

Trial stage

Second, there is a small but growing body of literature studying the effects of criminal courtroom experience on children (Small & Melton, 1994). While some children find the experience empowering (Melton & Thompson, 1987), others find it stressful. In a literature review of child victim-witness studies, many (but not all) of the children found the process distressing (Lipovsky, 1994). Of child witnesses who find testifying stressful, many do not suffer short- or long-term negative effects. Fortunately, many child victim-witnesses' functioning tends to improve over time, although they may improve at a slower pace than children who did not testify (Lipovsky, 1994). Still, some children are severely traumatized by testifying in criminal court (Goodman, Levine, Melton & Ogden, 1991).

Retrospective studies indicate the most distress in abused children is produced by the following stressors: face-to-face confrontation with the defendant in the courtroom (Goodman, Levine, & Melton, 1992; Underwager & Wakefield, 1992), fear of the defendant's retaliation, and fear of not being believed (Lipovsky, 1994). For example, some children who found confrontation with the defendant to be the most stressful aspect of testifying, reported (Berliner & Conte, 1995, p. 381):

“...He made me very uncomfortable and scared. He was giving me dirty looks”...“I was going to see my father in there, that's what I was scared of. He didn't make it any better on me when I was there. When I was up on the stand he'd wink at me, going 'its OK dear' you know what fathers do. It made it real tough because why would he say that?”

To date, however, no prospective research has exclusively and directly assessed the emotional impact of children testifying in front of defendants (Small & Melton, 1994; Underwager & Wakefield, 1992).

In addition to the above factors, children's reactions to court testimony may be affected by their external sources of support. Support may come from family members or legal professionals. Maternal support has been identified as an important mediating variable in determining how children react to confrontation (Lipovsky, 1994). In addition, research on British court cases has shown that children who were given some control over the way they gave evidence were more cooperative and happier under examination and cross-examination (Davies & Seymour, 1997).

Legal professionals need to consider the emotional well-being of abused child witnesses not only for therapeutic/humanitarian reasons, but also for legally-relevant reasons. When children are highly negatively aroused or distressed, they may refuse to testify or may testify incompletely (Goodman, Levine, & Melton, 1992). Several studies have shown that children who were most intimidated by face-to-face confrontation with the defendant provided lower quality testimony (Goodman, Taub, Jones, England, Port, Rudy & Prado, 1992; Saywitz & Nathanson, 1993). This may be due to the fact that while moderate anxiety can enhance performance, high levels of stress and anxiety generally narrow attention, interfere with memory and reduce the accuracy of cognitive processing (Bugental, Blue, Cortez, Fleck, & Rodriguez, 1992; Saywitz, 1995).

In response to the perceived needs of child victim witnesses in a growing number of cases, national groups as well as federal and state legislatures recommended and

implemented innovative accommodation procedures for use during criminal trials. The goal of courtroom accommodations is to help reduce systemic traumatization of child witnesses (i.e., the state's interest in protecting the safety of children) while simultaneously preserving defendants' federal and state constitutional rights to confront the witnesses against them. Some of the most promising accommodations include videotaped investigative interviews and closed-circuit televised testimony (Goodman, Levine, Melton, & Ogden, 1991; Lipovsky, 1994; Virginia Conference, 1997).

However, some commentators have expressed concerns about the use of some courtroom accommodations, especially those that have the potential to affect the defendant's Sixth Amendment right to confront the witnesses against him or her. As discussed above, some reforms were enacted despite little empirical research on the psychological effects of child witnesses testifying in the presence of the defendant (Cecchetti-Whaley, 1992; Small & Melton, 1994; Underwager & Wakefield, 1992). Other commentators suggest practical reasons for placing less emphasis on the use of accommodations. Few children actually testify in criminal court cases (Berliner & Conte, 1995), and if they do, some innovative procedures (e.g., CCTV) are not often used in practice (Goodman et al., in press; Lipovsky, 1994; Smith, Elstein, Trost, & Bulkley, 1993). Accommodations may be costly, considered difficult to implement, or not consistent with state and local laws (Lipovsky, 1994).

Despite these obstacles, social science research indicates that the use of courtroom accommodations enhances the psychological well-being of child witnesses as well as the quality of children's testimony (e.g., Davies & Seymour, 1997; Davies & Noon, 1993;

Murray, 1995; Saywitz & Nathanson, 1993; Swim, Borgida & McCoy, 1993). Indeed, the use of such procedures is considered consistent with issues of justice (Lipovsky, 1994) and the UN Convention on the Rights of the Child (Davies & Seymour, 1997). At the same time, accommodations have not been found to tip the scales of justice against the defendant. For example, research has shown that televised alternative procedures do not bias jurors against the defendant, or increase the number of guilty pleas or conviction rates (Davies & Seymour, 1997). For a detailed analysis of these procedures and of the empirical research supporting their use, see Hall (1999).

In addition, not all children should or will qualify for these accommodations. For a child to qualify for an accommodation that potentially implicates the defendant's Sixth Amendment Confrontation right, there must be a showing that the child meets the three-part test set out by the United States Supreme Court. Specifically, *Maryland v. Craig* (1990, p. 856) requires that the following factors be satisfied before alternatives to direct, face-to-face confrontation can be permitted:

- 1) there must be a case specific particularized showing that the alternative procedure is necessary to protect the child witness;
- 2) there must be a showing that the child would be traumatized by the defendant's presence, and not just the courtroom generally;
- 3) the trial court must determine that the child's emotional distress is more than de minimus, which was described as more than mere nervousness or excitement or some reluctance to testify.

The Court emphasized the need for an individualized, case-by-case determination of the need for courtroom accommodations. While the actual determination of the need for accommodation and what constitutes “more than de minimus” “emotional distress” or “trauma” are legal decisions that should be left to the trier of fact (Small & Melton, 1994), mental health practitioners (MHPs) can assist the court in these cases.

MHPs can act as impartial evaluators (Small & Melton, 1994) and/or as expert investigators to generate specific facts and data on the emotional well-being and experiences of the child (Goodman, Levine, & Melton, 1992). Evaluators should obtain information from persons who have had an opportunity to observe the child (e.g., parents, teachers, therapists, other neutral parties) (Eth, 1988; Goodman et al., 1991; Myers, 1992). Evaluations should be conducted based upon theory and research, “without forcing the child to undergo the very harm the State wishes to avert before protection may be invoked” (Goodman et al., 1991, p. 17).

However, it should be noted that there is a lack of scientifically validated and normed clinical methods for not only identifying children who may be in special need of court procedures, but also for assessing PTSD and other trauma-related symptoms in children (Briere, Elliott, Harris, & Cotman, 1995). Most research on PTSD has concentrated on adults, notably combat veterans and rape victims. With children, current instruments rely on self-report of the child or parents through structured interviews or questionnaires (Miller & Veltkamp, 1995; Allen, 1994). However, self-report measures are vulnerable to false reporting (Allen, 1994), inconsistencies and inaccuracies (Steward & Steward, 1996), and interviewer bias (Nader, Stuber, & Pynoos, 1991). Furthermore,

no research separates out the single factor of the defendant's presence from all other factors in assessing the effects of courtroom testimony on a child (Underwager & Wakefield, 1992).

Thus, MHPs who conduct *Craig* evaluations should employ wide-ranging, systematic, and as detailed evaluations as possible in order to best measure all of the potential short- and long-term effects of abuse and their mediating and moderating variables, as discussed above. A multi-axial, multimodal approach to assessment is preferred in which self-report measures are evaluated together with information from different collateral sources, objective measures, and behavioral assessments (Ebert & Fairbank, 1996; Kendall-Tackett et al., 1993; McNally, 1996). Direct behavioral observation instruments are particularly needed in this area (Kratonchwill, 1996; Nader et al., 1991) to explore "more specific behavioral manifestations for psychological constructs such as stress, fear, and intimidation that children may exhibit in certain testimonial conditions" (McAuliff & Kovera, 2000, p. 13). When offering an opinion in a case, MHPs should note the limits of their methods and expertise, exercise caution in making predictions, and avoid the ultimate issue (e.g., that the child will experience "more than de minimus" trauma, and that accommodation is "necessary" to protect the child) (Melton, Petrila, Poythress, & Slobogin, 1997).

Post-trial stage

The final area that may cause children psychological distress is after their testimony and after the legal process has been completed. As with pre-testimonial procedures, post-testimonial interventions directed toward the child and family may be helpful in

ameliorating emotional distress in child abuse victims/survivors. Post-testimonial interventions may include providing the following services to children and families: education; stress management procedures; programs to improve parental attitudes about court; services to improve parental/maternal support; and debriefing (Davies & Seymour, 1997).

In sum, while the legal system may traumatize some abused children and adolescents, others will benefit from it. Variability in outcome depends on the individual characteristics of the child, his or her external sources of support, the child's level of anticipatory distress, his or her testimonial experience, the verdict, and the aftermath of the case (Davies & Seymour, 1997). Mental health professionals might be able to help identify the relevant conditions that diminish trauma for child witnesses, as well as those that increase satisfaction with the legal system (Small & Melton, 1994). Thus, research continues to be needed to assess the impact that systemic intervention has on abused children, and to evaluate the procedures designed to ameliorate any secondary traumatization.

Moderating Variables

Moderating factors help some children survive the abusive experience without the substantial psychological, emotional and behavioral problems suffered by other abused children. This relatively new area of research with child victims of abuse is examining two types of moderating factors: resiliency/protective and coping factors.

Resiliency and Protective Factors

The resiliency paradigm can be used to explain why positive health outcomes occur when negative outcomes are predicted by the presence of risk factors (Chandy, Blum, & Resnick, 1996). Accordingly, research programs search for *predisposing factors* to help determine why some people are more severely affected by adversity than others, and *protective factors* to help explain how some people survive the abuse despite risk factors and continue to normally develop and function in society. Research on protective factors identifies an “at risk” group and a “resilient” subgroup and looks for factors within the “at risk” group to see how some factors buffer others and prevent negative outcome.

A study of protective and risk factors in female adolescents with history of sexual abuse found that some were at greater risk for disordered eating, pregnancy, suicidal involvement, poor school performance, and substance use (Chandy, Blum, & Resnick, 1996). Risk factors included stressful school environment, maternal alcohol use, and high levels of concern about sexual coercion. Yet, the three factors were identified as protective: perception of health, a sense of spirituality, and accessible, available sources of help. A second study of protective factors examined a group of 4 year old children and their mothers, and later studied them again when the children were 13 years old (Seifer,

Sameroff, Baldwin, & Baldwin, 1992). The study found that positive changes in children's cognitive or socio-emotional functioning were related to the children's perceived competence, locus of control, life events, and social support. Positive effects were also due to the child's interaction with his or her mother and the mothers' parenting values, social support, depression, and expressed emotion.

Indeed, the family environment of the abused child may be a strong potential moderating variable (Cohen & Wills, 1985; Heller, Larrieu, D'Imperio, & Boris, 1999). Considerable evidence indicates that a supportive family environment and responsive caregiving may buffer against the negative impact of many stressful childhood events, including maltreatment (Heller et al., 1999; Ray & Jackson, 1997). Families who offer crisis support (i.e., family support specifically in response to sexual abuse) may be a valuable moderator of abuse consequences. In addition, extra-familial social support (i.e., structured school environment, involvement in a religious community or in extracurricular activities) have been related to resilience in abused children (Heller et al., 1999). The literature supports continued work in the area of resiliency and protective factors (van der Kolk, McFarlane, & Weisaeth, 1996).

Coping Factors

Coping strategies and responses may also moderate the negative effects of child sexual abuse. Coping can be conceptualized as a process. In other words, a person may use a number of different coping strategies that change as the demands of the situation unfold (Curle & Williams, 1996). While the use of a coping measure is intended to help the child get through a situation, the chosen strategy may not always be healthy or positive.

Some general strategies that youth may employ in response to stressful situations have been identified. These coping strategies may include avoidance, resistance/disclosure/support seeking, emotional release, and cognitive restructuring (Spacarrelli, 1994). One study attempted to identify sexually abused children's coping strategies using principal components analysis (Chaffin, Wherry & Dykman, 1997). Abuse-related symptoms measured by multiple informants (child, parent, and teacher reports) were found to be associated with the following coping styles: (a) avoidant coping, (b) internalized coping, (c) angry coping, and (d) active/social coping. The less potentially healthy coping styles were associated with more negative symptoms. Specifically, internalized coping was correlated with increased guilt and PTSD symptoms and the angry coping style was associated with a wide range of behavioral and emotional problems as rated by child's teacher. However, the more potentially healthy coping strategies did not provide for moderating effects. Avoidant coping was associated with fewer behavioral problems but with greater sexual anxieties, and active/social coping was not associated with measured benefits. Clearly, further research is needed in this area before conclusions about protective factors can be made.

Summary

Responding appropriately and effectively to the needs of child abuse victims in society, therefore, requires thorough assessments of abused children, and consideration of current public policies and procedures. As discussed in the previous sections, evaluations of allegedly abused children and adolescents should consider:

- factors unique to the child (e.g., age/developmental level, gender, ethnicity);

- characteristics of the abusive incident (e.g., use or threat of force, intrusiveness of abuse, longer duration of abuse);
- systemic variables (e.g., family functioning and conflict, level of social support, secondary traumatization from investigative interviews or court testimony), and the effect of public policies designed to improve child functioning in these systems; and
- the child and/or his or her family's protective and coping factors.

Given the complexity of these cases, reliable and valid clinical-forensic assessment tools are in demand to improve the therapeutic and forensic evaluation and treatment of allegedly abused youth, especially those who may need accommodations in the criminal justice system. While the field needs empirically-supported, multidimensional assessments for abused youth, it currently lacks an objective, behavioral measure of emotional distress in children and adolescents and a measure of risk factors associated with secondary traumatization. Also, few studies have examined children's emotional states during abuse investigations, even though the literature suggests that some children find it to be one of the most stressful stages in the criminal justice and child protective processes (Lipovsky, 1994; Lipovsky & Stern, 1997). Moreover, no measure directly addresses the U.S. Supreme Court's three-part test in *Maryland v. Craig* (1990) - that a child can benefit from a courtroom accommodation only if an individualized finding is made that s/he would suffer more than *de minimus* trauma or emotional distress from face-to-face confrontation with the defendant (Briere, Elliott, Harris, & Cotman, 1995; Hall, 1999).

The Present Study

The purpose of this research program is to address many of the empirical issues posed by child abuse victims who become involved in the criminal justice system that have been reviewed above. First, there is a need to examine the presence and effect of risk and protective factors in abused youth, including those that have been associated with secondary traumatization, in actual child abuse cases – and in particular, during investigative interviews. Second, the field needs an objective, scientifically based, behavior observation instrument to assess emotional distress and trauma in allegedly abused youth. Moreover, there is a need for more research on the *Craig* factors, including assessing emotional distress in abused youth, and the influence of the defendant on the child.

To address these needs, the present study contains two major objectives. First, the effect of mediating and moderating factors (including those related to secondary traumatization) on the emotions and behaviors of a sample of allegedly sexually and physically abused children during investigative interviews will be assessed.

Second, and at the same time, this study will test the concurrent criterion validity of the Videotaped Interview Trauma Assessment List (VITAL). The VITAL was designed to fill a need in the field of child psychology for a behavioral observation instrument that, when validated, can be used in conjunction with other methods to assess emotional distress and risk factors associated with secondary traumatization in allegedly abused youth. In addition, the VITAL was developed to assess the *Maryland v. Craig* three-part test (e.g., the influence of the defendant on the child's emotional state), so that it may be used as part of a complete, multimethod assessment for courtroom accommodation

determinations once it is validated. Because discussions of alleged abuse in the investigative interview setting is similar to that which the child might experience during a subsequent interview with a prosecutor or during testimony in court, the VITAL should be generalizable and ecologically valid for use in such settings.

Criterion validity assessment of the VITAL is an appropriate next step in the validation of the instrument, because basic evidence of the VITAL's content/construct validity has already been shown in a pilot study (Hall, 1998). Concurrent criterion procedures are used to assess the effectiveness of a test for a specific program or activity (Anastasi, 1988). The present study was designed to test the effectiveness of the VITAL in assessing the emotion/behaviors of children and question/statements of interviewers during investigative interviewing. Subsequent research can focus on predictive criterion validity and different type of construct validity to indicate the VITAL's effectiveness in predicting children's performance on related measures (e.g., concurrent validity; convergent and discriminant validation) and on the activities of interest (e.g., courtroom testimony).

Construct/content validity can be defined as the extent to which the assessment tool may be said to measure a theoretical construct or trait (Anastasi, 1988). Evidence of content validity comes from the original design of the observational codes (Suen & Ary, 1989), which must exhaust the behaviors under study (Herbert & Attridge, 1975). Direct observation and the gradual accumulation of data from a variety of sources is also required to establish construct validity (Anastasi, 1988; Kaminer et al., 1990, citing Cronbach & Meehl, 1955).

As previously described in the pilot study, the VITAL meets these criteria for construct validity because the VITAL was designed to encompass all dimensions of child trauma and PTSD as reported in the literature, and all types of questions that investigative interviews might ask. Specifically, categories of behavior were created such that the child's emotional/behavioral reactions and the interviewer's statements were recorded in only one category at a time (i.e., categories are mutually exclusive), and at every moment of observation, the behavior would fit one of the categories (i.e. categories are exhaustive) (Mudford, Hogg, & Roberts, 1997). After construction, the VITAL was approved by a focus group of local psychologists who specialize in child abuse. Data was gathered according to established recommendations for collecting continuous observations/behavioral data (Bakeman & Gottman, 1997; Hops, Davis, & Longoria, 1995; Topf, 1988). Findings from the pilot study about coding categories were used to refine and improve the instrument. Lastly, results from the present study were taken from a different sample than the pilot study, will provide further evidence about the validity of the categories in measuring the emotional/behavioral components of trauma/emotional distress in allegedly abused children and adolescents.

Hypotheses

To determine the concurrent criterion validity of the VITAL, the criterion measure should be tested. In the case of the VITAL, observer ratings on the child and interviewer VITAL forms comprise the criterion measure (Anastasi, 1988). Because the VITAL contains a number of categories of observer ratings about the allegedly abused child and the investigative interviewer, the current study involves testing a number of hypotheses to

help establish criterion validity. As a general rule, testing a number of hypotheses requires a large number of subjects; otherwise one may interpret findings as significant when findings instead reflect capitalization on chance. The present study, however, is not as interested in testing each separate hypothesis, as in discovering that many of them contribute to a picture that supports the utility and concurrent criterion validity of the assessment and coding procedure (J. Allen, personal communication, April 4, 2000). Moreover, the following hypotheses were based on previous research and determined *a priori*, thus lending additional support for this approach (see, e.g., Iacobucci & Wasserman, 1988).

Based on the research described above on investigative interviews, the first set of hypotheses posits that certain VITAL ratings will be distributed differently over the entire sample of allegedly abused children. First, because investigative interviews are generally considered stressful (e.g., Lipovsky, 1994), it is hypothesized that children will have a fairly high average VITAL child rating of tense/uncomfortable, and fairly low average ratings of positive emotions/behaviors. Second, previous research on investigative interviews has also shown that the most frequently asked question-types are specific and closed; those questions asked less frequently are open and leading (e.g., Westcott, Davies, & Horan, 1998). Yet, interviewers may use fewer open-ended questions and more specific, closed and leading questions with younger children, which is considered appropriate because of developmental differences in language and maturity (APSAC, 1990, 1998). For the same reasons, it is also hypothesized that time spent in the rapport

phase will be higher for younger children as compared to older children, especially adolescents.

Ideally, child VITAL ratings should also vary by whether the child's case was deemed founded or unfounded by the MDT at the family violence center. However, it is hypothesized that there will be only a slightly higher degree of certain child VITAL ratings for children in the founded group as compared to the unfounded group. Because a determination of founded or unfounded does not mean that the truth of the allegation is verified or true, it is likely that the unfounded group contained some abused children. In addition, two of the children in the unfounded group were also discovered to be witnesses to the alleged abuse of a sibling. Witnessing the abuse of another has been associated with emotional distress in children (Arroyo & Eth, 1995). Thus, slightly higher ratings on anger, scared and depressed are expected in the founded group; higher ratings on positive emotions/behaviors are expected in children from the unfounded group; and little difference is expected on the uncomfortable ratings for both groups.

Similarly, interviewer VITAL ratings should vary by the case outcome. It was expected that interviewers would spend similar proportions of time with allegedly abused children in rapport as well as in general issues during cases that were later deemed founded or unfounded. However, because interviewers may spend more time asking about specific abuse-related issues with children whose stories revealed more information about their personal experiences, it was hypothesized that higher proportions of time would be spent in personal questions in founded cases as compared to unfounded cases. For the

same reason, interviewers may spend higher proportions of time asking questions about the alleged abuser in founded as compared to unfounded cases.

Differences in child VITAL ratings may also be found according to the interaction between the type of abuse and the determination of a case as founded or unfounded. Because physically maltreated children may be at higher risk for reactive aggression than other abused children and non-maltreated children (Shields & Cicchetti, 1998), it is hypothesized that children from founded cases of physical abuse will have higher anger/aggressive or externalizing ratings than children from founded physical cases or unfounded cases.

The last hypothesis in this set concerns the interaction between case outcome and the child and interviewer VITAL ratings. Based on *Maryland v. Craig* rationale, it was hypothesized that children in the founded group would display higher amounts of emotional distress when asked personal questions about the alleged abuse, the alleged abuser, and court/criminal justice intervention, and lower amounts when asked about abuse in general, or about neutral/other issues.

Mediating factors involving demographic, abuse-specific, and systemic variables will be assessed in the second set of hypotheses. Based on the conflicting research regarding demographic variables, it is hypothesized that the child's age, gender and ethnicity will have minimal effect on the child's VITAL ratings. If a difference is found, it is likely to be similar to other studies which have found that abused girls score higher on internalizing variables than boys (e.g., Ackerman et al., 1998) and that boys of increasing age will score higher on externalizing variables (e.g., Livingston, Lawson, & Jones, 1993).

Regarding abuse-specific risk factors, many of these variables rest on disputed empirical evidence. As previously discussed, the risk factors of greater frequency and duration of abuse have been disputed. Similarly, recent research indicates that the severity or intrusiveness of the alleged abusive act does not consistently influence trauma (Davenport et al., 1994). In addition, evidence for the use of force is ambiguous with regard to its traumatizing effects because while it might be initially traumatic, it may serve to might alleviate guilt (Davenport et al., 1994). Finally, abuse by someone who the child knows seems more likely to cause lasting damage than abuse by a stranger (McClellan et al., 1995), but negative effects may instead be caused by the quality of the relationship as perceived by the child. It is hypothesized, therefore, that abuse-specific factors will have little effect on child VITAL ratings.

Regarding systemic risk factors, family dysfunction and lack of social support have been associated with negative effects in abused youth as described previously. Thus, it is hypothesized that children from unstable homes (e.g., difficult divorce; previous abuse) will have higher VITAL ratings than children from stable homes. It is also hypothesized that children with no or limited sources of social support (i.e., zero to one supportive person) will have higher VITAL ratings than children with more sources of social support (i.e., more than one source of social support).

The secondary traumatization risk factor with the most support in the research literature is multiple interviews. Therefore, it is hypothesized that an increase in the number of investigative interviews or interviews conducted by trained professionals will be associated with higher child VITAL ratings. Initial disclosure to a parent or other

untrained person was not included in the coding of multiple interviews, due to the previously discussed conflicting literature on the effects of disclosure. Instead, the nature of the child's disclosure to a non-trained person was considered when making a rating of parental supportiveness and/or social support.

The other systemic, secondary traumatization variable with some empirical support is the effect of rapport during investigative interviews. While the interviewer VITAL does not measure the nature of the rapport, it records the amount of time an interviewer spends in asking rapport-related questions or in making rapport-related comments. Thus, it is hypothesized that children will have lower VITAL ratings of uncomfortable when higher proportions of time are spent in rapport. Because the age of the child might be a confounding variable in the analysis of rapport, due to the developmental reasons for spending more time in rapport with younger children described above, the effect of age will be statistically controlled.

While improper interviewing techniques (e.g., leading questions with older children) may affect a child's experience of comfort during the interview, their effect on interviewed children will not be adequately ascertained in the present study. The analyses in the present study of the interviewer VITAL form simply measure the proportion of time spent during interviews in asking types of questions, such as leading questions about the abuse. The content of the question is not assessed, nor its impact on the child. Future studies with the VITAL will explore sequential analyses to connect interviewer questions with the child's emotional response over time. In addition, other studies that assess the use

of investigative interview guidelines could address this issue, and are planned by the author.

The final set of hypotheses concerns moderating variables. In the present study, the protective factors of social support and non-offending parental support were assessed. Based on the previously reported research, it is hypothesized that the presence of these two measures of support will be associated with lower VITAL ratings of emotional distress in the entire sample, with small differences between the founded and unfounded abuse groups. Because the present study could not include researcher-conducted interviews with the children, internal coping mechanisms could not be evaluated. Future prospective research with the VITAL should include such measures.

METHOD

Subjects

The subjects of the present study consisted of thirty videotaped child-interviewer pairs. Every videotape recorded a full investigative interview, in which an interviewer questioned a child regarding allegations that the child had been sexually or physically abused. Trained detectives, child protective service workers and/or forensic interviewers conducted the interviews at a family violence center or other MDT-affiliated facility located in the largest county in a southwestern state. The interview rooms had a comfortable atmosphere, and were equipped with two-way mirrors and virtually-invisible, ceiling-mounted microphones and video-recording equipment.

A random sample of videotaped interviews and their corresponding written files were gathered from detectives' closed files at this family violence center. Closed cases involving founded and unfounded allegations of sexual and physical abuse from January 1, 1995 to January 1, 1998 were randomly selected and obtained from detectives at the center. Determinations of whether the case was founded or unfounded (i.e., met local standards for use as evidence in criminal proceedings) were made by the MDT at the family violence center and noted in the written files. This time period was chosen to ensure that a sizable number of cases would have completed their investigation and the charts would have been closed by the start of the present study (see Rao, DiClemente, & Ponton, 1992). Due to the exploratory nature of the present study and to address confidentiality issues, closed cases were chosen instead of ongoing cases.

Consent and written permission to gather data were granted to the author from the family violence center and the city chief of police. A confidentiality agreement between the center and the research team was signed, and updated for new members of the research team. Human subjects approval for the research program was obtained from the University of Arizona. Procedures during data collection and analyses were followed to maintain the anonymity of the participants and protect the confidentiality of the data.

Coding Procedures

Materials

The Videotaped Trauma Assessment List (VITAL) consists of three parts: (a) a written record review form, (b) a child emotions/behaviors form (child VITAL), and (c) an interviewer questions/statements form (interviewer VITAL). Demographic and background information about the sample cases were gathered using the VITAL record review form (Appendix A). The author completed the record review forms after coding of the videotapes was completed to remain blind to case information. Research assistants were not authorized to view the written files. Without recording identifying information about the parties to a case, the record review form is used to record demographic information about the child, the alleged perpetrator(s), the nature of the alleged abuse experience(s), family functioning, social support, interview information, and case status.

The child and interviewer coding forms were designed for use primarily with videotape data, which allow detailed analysis over multiple viewings (see Coster, Tickle-Degnen, & Armenta, 1995). Less expensive and more accessible than computerized systems but more time-consuming, the VITAL child and interviewer forms allow for

paper-and-pencil recording of behavioral ratings. Child and interviewer behavior categories were operationally defined, and coding precedence rules were developed to clarify and improve the reliability of categories (see Appendix B for child VITAL operational definitions and coding procedures, and Appendix C for interviewer VITAL operational definitions and coding procedures). In addition, the categories were designed to be both mutually exhaustive, such that the subject can, by definition, be recorded in only one category at a time, and exhaustive, such that every moment of observation can be recorded in one of the categories.

The child VITAL is a 20-category code form in which one of 20 emotion/ behavior states is coded every 10 seconds for the duration of the interview. The emotion/behaviors were rationally clustered together on the coding sheet to facilitate recording, as follows: unobservable, neutral, anger/aggressive, disgust/contempt, sad/hurt, tense/anxious/fear, embarrassed/shame, whine/frustrated, distract, interest/affection, and happy/excited (see Appendix B).

The interviewer VITAL is a 30-category code form in which one of 30 interviewer question and/or statement categories is coded every 10 seconds for the duration of the interview. The codes include: none, other, connector, rapport, question/statement about abuse (general), question/statement about abuse (specific), and question/statement about legal system. The two abuse-related question/statement categories are coded in subcategories that describe the form of the question/statement (i.e., open, specific, closed, leading, and recapping question/statements). Finally, each content code (i.e., none category excluded) has two different categories: one category refers to

questions/statements referring to the alleged abuser, and the other category refers to questions/statements not referring to the alleged abuser (see Appendix C).

VITAL Coding Procedures

The coding of the Child and Interviewer VITAL was conducted in a private room in a university-owned house that could comfortably seat four coders and a supervisor who controlled the timing of the coding process. Because the method used to rate videotapes may affect interrater reliability, it is recommended that raters should code independently, and each tape should be coded by more than one person (Bakeman & Gottman, 1997; Orsak, 1993; Wood et al., 1996).

The VITAL pilot study followed recommended procedures by using coding teams of four psychology undergraduate students who independently rated the child and interviewer every 10-seconds for the entire interview. Two pairs of coders simultaneously coded each interview with one member of each pair coding interviewer behaviors and one member coding child behaviors. They were seated to ensure that no coder rating the same person could see the ratings made by the other coder. Coders were blind to case information, including whether the case had been deemed founded or unfounded.

The 10-second intervals were tracked using either the pre-recorded time stamp on the original videotape, or by the timing device on the investigator's VCR unit. The principal investigator controlled the coding by pausing the recorder every 10 seconds, giving each coder enough time to make a code before proceeding. The tape could be stopped or rewound if a coder found it useful in coding the behaviors. No discussion of coding decisions was allowed during the coding process.

The same coding procedures were followed in the present study, except that fewer raters were employed. Specifically, two undergraduate seniors majoring in psychology (one participated in the pilot study) and the author independently rated the videotapes. The smaller number of coders was appropriate given the positive interrater reliability findings of the pilot study, as suggested by Bakeman and Gottman (1997).

Coders received independent study credit to be observers for the present study. Each received at least 16 hours of initial training on the VITAL, including training with the FACS (Eckman & Friesen, 1978) and SPAFF (Gottman, McCoy, Coan, & Collier, 1996). During training, coders memorized code abbreviations and definitions of each behavior, received a copy of operational definitions for the child VITAL (Appendix B) and interviewer VITAL (Appendix C), received a copy of the FACS faces, and practiced using the VITAL with role play and training videotapes not used in the study (i.e., SPAFF training videotapes and a Children's Institute International & United Way, Inc. videotape (1986)). During training, coders discussed their ratings to resolve discrepancies. After meeting criteria (i.e., interobserver agreement at least .7), coders practiced using videotapes that were not included in the present study. Other methods to maintain reliability (Hops et al., 1995) were followed as coders continued to receive individual and group training and feedback about interobserver agreement throughout the coding process. The coding lasted approximately three months for the pilot study and over two semesters for the present study.

The author conducted the training and rated all of the tapes, as the master or standard coder (Bakeman & Gottman, 1997). By participating in the rating process, the

author's goal was to maintain reliability by minimizing observer drift (Hops et al., 1995), and avert the pitfalls of covert assessment, the tendency for a decline in interrater reliability when the investigator is absent (Topf, 1988).

Statistical Analyses

Interrater Reliability

The first step in the dissertation analyses was to determine that coders reliably coded the subjects on the child and interview VITAL forms. The present study used Cohen's kappa to calculate interrater reliability because the statistic reflects agreement among raters and controls for chance agreement (Bakeman & Gottman, 1997). Global kappas were computed for overall agreement between the three pairs of coders on all categories for each child and interviewer VITAL in the sample. In addition, kappas for each separate child and interviewer VITAL category with at least minimal base rates (i.e., frequency of at least 1% during the interviews) were computed using the formula developed by Kvalseth (1989) (see, e.g., Ambrosini, Metz, Prabucki, & Lee, 1989). Calculating kappas for specific categories is recommended when developing new categorical tests (Garvin, Kennedy, & Cissna, 1988). Because no difference was found among the average kappa scores for pairs of raters, one of the undergraduate coder's ratings was selected and used for subsequent analyses. The authors' ratings were not chosen in order to minimize any possibility of bias, even though the author remained blind to the outcome of the cases.

VITAL Data Reduction

Because VITAL child and interviewer categories were rated in 10-second intervals for each subject, the data for each subject needed to be expressed in a meaningful way that allowed for comparisons with ratings of other subjects in subsequent analyses. In the present study, a subject's ratings in each category were expressed and standardized by calculating the proportion of time that the subject was coded in each category over the total amount of time he or she spent in the interview ("proportion").

However, as seen in **Table 1.1** and **Table 2.1**, many of the categories suffered from low base rates. To improve the reliability and the interpretability of the VITAL ratings, categories were collapsed using rational methods (J. Allen, personal communication, March 2000). A first step involved excluding the child *unobservable* and *neutral* categories and the interviewer *none* category from analyses, such that proportions would consist of *observed* child emotion/behaviors and *observed* interviewer question/statements.

Next, regarding the child VITAL, child categories were collapsed according to the way the categories were organized on the coders' rating sheets (see **Appendix B**). For example, child VITAL categories of sad and pain/hurt were combined as depressed. However, base rates for many child VITAL categories remained too small to consider further analyses based on them (**Table 1.2**). Thus, these categories were collapsed using two different rational methods, both based on organizations used by other reliable and valid measures of abused children's emotions and behaviors, which informed the design of the VITAL (e.g., Child Behavior Checklist – Direct Observation Form (CBCL-DOF,

Achenbach, 1986); Trauma Symptom Checklist for Children (Briere, 1996)). Assumption tests were conducted on the two new sets of child VITAL categories (described next) to determine their appropriateness for parametric analyses following the general linear model. In other words, for analyses of variance, including ANOVAs, MANOVAs, and ANCOVAs, dependent variables should have normal distributions and equal variances (Bray & Maxwell, 1985; Iverson & Norpoth, 1987; Wildt & Ahtola, 1978). Descriptive statistics were used to summarize the results of the new categories.

For the child VITAL, two combinations of categories were found to be appropriate. First, envisioned during the VITAL's design, a more general combination of categories produced two categories of child emotion/behaviors: *externalizing* and *internalizing*. As found in other measures (e.g., CBCL-DOF) and in the research, internalizing abuse-related problems include anxiety, fear, depression, withdrawal, and avoidance (Briere, 1992; Dubowitz, Black, Harrington, & Verschoore, 1993). Thus, internalizing consisted of the original sad/hurt, tense/anxious/fear, embarrassed/shame, and withdraw codes, and externalizing included the original angry/belligerent, disgust/contempt, whine/cranky codes. Proportions of time spent in these two categories is reflected in **Table 1.3**.

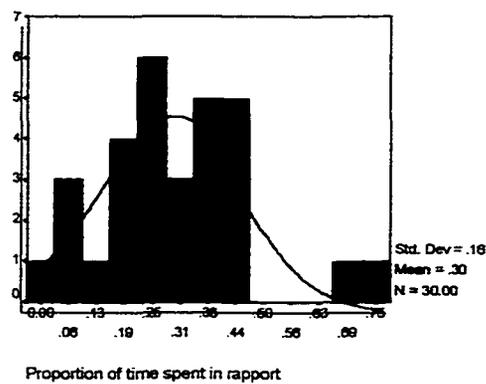
However, assumption tests revealed non-normal distributions for the externalizing (skew = 1.320, standard error = .427; kurtosis = 2.925, SE = .833; Shapiro-Wilk test of normality = .909, $p < .05$); and internalizing codes (skew = -.757, SE = .427; kurtosis = 2.089, SE = .833; Shapiro-Wilk = .946, $p = .201$). Removal of one extreme outlier (i.e., one score was at least three standard deviations from the mean) transformed each category

and corrected the problem. Thus, after transformation, both externalizing (skew = .309, standard error = .434; kurtosis = -.940, SE = .842; Shapiro-Wilk test of normality .954, $p = .330$) and internalizing (skew = .309, standard error = .434; kurtosis = -.691, SE = .845; Shapiro-Wilk = .957, $p = .320$) were normally distributed, with distributions that were slightly more peaked and to the right than a perfect normal distribution (i.e., where skew = 0; kurtosis = 0).

Another, more specific, combination of child VITAL categories was selected that would reflect a wider range of different emotions and behavior states in children. Specifically, five categories were rationally chosen: *positive* (interest/affection, happy/excited), *uncomfortable* (tense, whine/cranky), *angry* (angry/belligerent, disgust/contempt), *depressed* (sad/hurt, embarrassed/shame, withdraw), *scared* (anxious/fear). Proportions of time spent in these the five-factor VITAL categories is reflected in **Table 1.4**.

While the uncomfortable code had a normal distribution (skew = -.491, SE = .427; kurtosis = -.189, SE = .833; Shapiro-Wilk test of normality = .968, $p = .528$), the other four codes did not. A natural log transformation was found to be most appropriate and effective for these categories: positive (skew = -.268, SE = .550; kurtosis = -1.171, SE = 1.063; Shapiro-Wilk = .912, $p = .116$), angry (skew = .142, SE = .481; kurtosis = .492, SE = .935; Shapiro-Wilk = .968, $p = .622$), depressed (skew = -.472, SE = .427; kurtosis = .106, SE = .833; Shapiro-Wilk = .967, $p = .494$), scared (skew = -.786, SE = .427; kurtosis = .368, SE = .833; Shapiro-Wilk = .933, $p = .074$).

For the interviewer VITAL, categories were combined rationally in accordance with the hypotheses to be tested. First, proportion of time spent in rapport was obtained by collapsing the two rapport categories (i.e., one in which rapport included reference to the alleged abuser and the other that did not include such reference). Proportion of time spent in this category, with the normal curve superimposed, is reflected in **Table 2.2**:



Rapport was approximately normally distributed (skew = .485, SE = .427; kurtosis = .724, SE = .833; Shapiro-Wilk test of normality = .960, $p=.401$).

Similarly, the proportion of time spent in open, specific, closed, leading, and recap question type was obtained by collapsing ratings from both *general abuse* and *specific abuse* areas, and from questions that did and did not refer to the alleged abuser. Proportion of time spent in these five question-type categories is reflected in **Table 2.3**. Analyses of test assumptions indicated approximate normal distributions for the *closed* category (skew = .154, SE = .427; kurtosis = -.454, SE = .833; Shapiro-Wilk test of normality = .985, $p=.944$). A log transformation was selected for the *open* category (skew = .671, SE = .501; kurtosis = -.437, SE = .972; Shapiro-Wilk = .930, $p=.178$). Removal of

one extreme outlier was the appropriate transformation for the remaining categories: *specific* (skew = .438, SE = .434; kurtosis = .672, SE = .845; Shapiro-Wilk = .973, $p=.685$); *leading* (skew = .825, SE = .434; kurtosis = .280, SE = .845; Shapiro-Wilk = .926, $p=.049$); and *recap* (skew = .160, SE = .434; kurtosis = -.821, SE = .845; Shapiro-Wilk = .970, $p=.605$).

Third, in order to examine differences among general categories of questions, subcategories were rationally collapsed. For example, specific questions about abuse included open, specific, closed, leading, and recap questions that did and did not pertain to the alleged abuser in that category. However, low base rates did not allow for separate analyses of the question/statements about neutral/other issues or legal system issues categories. Thus, *other* category was rationally collapsed into the *general questions* category, and the *legal system* category was collapsed into the *specific abuse* category. Thus, four categories remained: connector; rapport; general and educational questions (*general*); and personal abuse-related questions (*personal*). Proportions of time spent in these four categories is reflected in **TABLE 2.2**. Because the connector category was not related to the hypotheses of the present study, it was not included in subsequent analyses. Test assumptions for *general* questions (after a log10 transformation, skew = -.028, SE = .448; kurtosis = -.113, SE = .872; Shapiro-Wilk test of normality = .985, $p=.946$) and *personal* questions (skew = -.202, SE = .427; kurtosis = .113, SE = .833; Shapiro-Wilk = .975, $p=.710$) revealed approximate normal distributions.

Finally, the original categories were separated into two groups (i.e., questions about the alleged abuser and questions not about the alleged abuser) in order to examine

hypotheses about the effect on the child of being asked about the alleged abuser.

Proportions of time spent in these two categories is displayed in **TABLE 2.4**. Because only the question/statements about the alleged abuser category (alleged abuser) was of interest in the present study, subsequent analyses focused solely on this category. Analysis of test assumptions indicated approximate normal distributions for the *alleged abuser* category (skew = .342, SE = .427; kurtosis = -.330, SE = .833; Shapiro-Wilk test of normality = .975, $p > .05$ ($p = .703$)).

Hypotheses Testing

The next phase of data analysis involved analyzing information from the case files with descriptive statistics. The case file information included: demographic information about the child (i.e., gender, age, ethnicity), demographic information about the child's alleged perpetrator (i.e., gender, age, ethnicity), information about the interviewer (i.e., gender, profession); risk factors (i.e., duration of abuse; frequency of abuse; use of force; intrusiveness of sexual abuse; relationship of perpetrator to child; multiple interviews; less rapport), and protective factors (i.e., family environment, non-offending parental support; social support).

The final phase of data analysis concerned testing the criterion validity of the VITAL. Specifically, the relationships between the dependent variables (DV) of child emotions/behaviors and interviewer questions measured by the VITAL, and the independent variables (IV) of risk and protective factors were assessed. When test assumptions were met, DVs were tested for normal distributions and transformed if necessary, and homogeneity tests were conducted on each hypothesized DV-IV pair), the

general linear model was used to test hypotheses. To reduce the number of statistical tests conducted in the present study, multivariate analyses (i.e., MANOVAs) were used whenever appropriate instead of a series of univariate analyses (e.g., ANOVAs). In addition, child VITAL ratings of internalizing and externalizing were examined against hypothesized IVs using MANOVAs before any specific 5-category child VITAL ratings were tested. Specific child categories were only subsequently examined if trends or statistically significant findings were revealed using the internalizing-externalizing model. However, when variances were not approximately equal, hypotheses were tested using non-parametric statistics (i.e., Mann-Whitney U test).

In addition, a within-groups design was used to further compare children whose cases were deemed founded (n=15) with cases deemed unfounded (i.e., not enough information or evidence presented to continue the investigation, or support filing of legal charges) (n=15) by the center. Because the dependent variable of case outcome was dichotomous (i.e., founded versus unfounded), the Mann-Whitney U test was used to test hypotheses between risk and protective factors, and case outcomes.

RESULTS

Thirty videotaped child-interviewer dyads participating in an investigative interview about alleged abuse were examined in the present study. Twenty (66.7%) cases involved allegations of sexual abuse and 10 (33.3%) cases involved allegations of physical abuse, 15 (50%) of which were deemed founded and 15 (50%) unfounded by the family violence center. Two (6.6%) cases also involved the child being witness to the sexual or physical abuse of a sibling. An unfounded decision was made when investigators determined that there was not enough information or evidence presented to continue the investigation, or to support the filing of legal charges against the alleged perpetrator(s) at that time.

Twenty-one of the children (70%) were female and 9 (30%) were male. Of the girls, 16 were allegedly sexually abused and 5 were allegedly physically abused. Of the boys, 4 were allegedly sexually abused and 5 were allegedly physically abused. (See **TABLE 3.1**, Child gender by abuse type, and **TABLE 3.2**, Child age and gender by case outcome and abuse type). At the time of the videotaped investigative interview, 10 children (33.3%) were pre-school-aged (3 to 5 years of age), 13 children (43.3%) were school-aged (6 to 9 years of age), and 7 children (23.3%) were pre-teen and teenagers (10 to 15 years of age). The mean age of the child subjects was 7.7 years old, with a standard deviation of 3.55 years. Approximately half of the children (17, 56.7%) could alternatively be categorized as younger (3 to 7 years of age) and half (13, 43.3%) as older (8 to 15 years of age). The majority of children (26, 86.7%) were of European American descent.

One child (3.3%) was African American, and 3 children (10%) were Hispanic/Latino/Mexican American.

The majority of alleged abusers were male. Twenty-seven (90%) were male and 3 (10%) were female. The gender of the child matched that of the interviewer in 10 cases (male child-male interviewer in 8 cases; female child-female interviewer in 2 cases; male child-female interviewer in 1 case; female child-male interviewer in 19 cases). At the time of the videotaped investigative interview, the mean age of the alleged perpetrators was 33.2 years, with a standard deviation of 11.2 years. They ranged in age from 15 to 59 years of age. Five (16.6%) were under the age of 21. Ten (33.3%) of the alleged perpetrators may be considered of highest risk (15 to 27 years of age), 10 (33.3%) of less risk (29 to 38 years old), and 10 (33.3%) of least risk (39 to 59 years of age), according to lifetime prevalence data of antisocial/criminal behavior for males (Grisso, 2000; Moffitt, in press/n.d.). Twenty-four (80.0%) of the alleged perpetrators were Caucasian, 4 (13.3%) were Hispanic/Latino/Mexican American, and 2 (6.7%) were African-American. In 28 cases, the ethnicity of the child matched that of the alleged perpetrator (the other 2 cases involved a Caucasian child allegedly abused by a Hispanic or an African American person). Ten (33.3%) of cases involved two alleged abusers and 20 (66.7%) involved one.

Of the 30 videotaped interviews, 24 (80%) were primarily conducted by female interviewers. Six (20%) interviews were primarily conducted by male interviewers. The gender of the interviewer matched that of the alleged perpetrator in 9 (30%) cases. Sixteen (53.3%) of the interviewers were Caucasian, and 14 (46.7%) Hispanic. The ethnicity of the interviewer matched that of the alleged abuser in 17 (56.7%) cases.

Detectives who were trained in interviewing children conducted the majority (27) of the interviews. However, one interview (3.3%) was conducted by a Ph.D. forensic interviewer, and 2 (6.7%) by a M.S.W. forensic interviewer.

An average of approximately 2 (mean = 1.86, SD = .69) interviews were conducted by trained professionals. In 8 cases (26.7%), only one interview was conducted with the child. Two interviews were conducted in 18 (60%) of cases. Only a few cases involved higher numbers of interviews. Two cases (6.7%) involved 3 interviews and one case (3.3%) involved 4 interviews.

Regarding other risk factors, the use of force or threats was reported in 10 (33.3%) cases, and not reported in 20 (66.7%). Abuse was found to be most frequent (more than 5 times) in 10 (33.3%) cases, less frequent (2 to 5 times) in 11 (36.7%) cases, and least frequent (1 time) in 9 (30%) cases. The duration of abuse was measured by the approximate number of days from when the abuse was reported to start, to when it reportedly ended. The mean number of days was 243.9 (standard deviation = 350.4), and range of 1 to 1095 days. About half of the cases involved alleged abuse of duration up to a month in length (14 (46.7%) cases involved a duration of 1 to 31 days). Within this month time period, 9 cases allegedly lasted up to a week (1 to 7 days) and 5 from two weeks to a month (14 to 31 days) in duration. Nine cases allegedly lasted from a month to a year (45 to 355 days), and 7 cases reportedly lasted from a year to 3 years (380 to 1095 days).

Overall, the sample of cases in the present study involved allegations against people who were known to the child (100%). One case involved an acquaintance of the child, and 6 (20%) an involved adult who lived outside the child's home. The rest of the

cases involved more closely-related people to the child (i.e., in-home sibling (2 cases), in-home adult other than the parent (2 cases), out-of-home parent (4 cases) and in-home primary caregiver (15 cases)). Thus, the majority of the cases (19 cases, 63.3%) involved alleged abusers who lived in the child's home, most of whom (15 cases, 50%) were the child's primary caregiver.

For the 20 sexual abuse cases, the degree of intrusiveness was measured on a 12-item scale (see Appendix A) (J. V. Becker, personal communication, 1995). Ten (50%) of the cases involved lower levels of intrusiveness: 1 (3.3%) emotional abuse due to witnessing alleged sexual abuse of another; 6 (20%) fondle over child's clothes; and 3 (10%) fondle under child's clothes. Ten (50%) cases involved more intrusive allegations of sexual abuse: 2 (6.7%) child masturbate abuser/simulated sex under clothes; 1 (3.3%) oral abuser-to-child; 1 (3.3%) digital or object penetration; and 6 (20%) vaginal/anal intercourse. Degree of intrusiveness for physical abuse could not be ascertained from the written files and was not included in the analysis of this risk factor.

Regarding systemic risk and protective factors, nine (30%) children came from stable homes and 21 (70%) can from unstable homes (i.e., those homes characterized by prior abuse and contested divorces. The non-offending parent was rated supportive in 21 (70%) of cases, and unsupportive in 9 (30%) cases. Fifteen (50%) children apparently had more than one source of social support in their lives. However, 11 (36.7%) children had only one source and 4 (13.3%) children apparently had no social support. Interviewer-provided social support was measured by the proportion of time spent in rapport. The

average proportion of time spent in rapport was .30 (standard deviation of .16), with a range from .0 to .73.

Interrater Reliability

The present study assessed the interrater reliability of the child and interviewer VITAL ratings between the three coders. The overall Cohen's kappa for the child VITAL averaged .539 over the three pairs of raters (rater 1 x rater2, $K = .506$; rater 1 x rater 3, $K = .588$; rater 2 x rater 3, $K = .524$). This indicates moderate agreement (range = .41 to .60) according to Landis and Koch (1977). The overall kappa for the interviewer VITAL averaged .834 over the three pairs of raters (rater 1 x rater2, $K = .808$; rater 1 x rater 3, $K = .881$; rater 2 x rater 3, $K = .814$). This indicates almost perfect agreement (range = .81 to 1.0) according to Landis and Koch (1977). **TABLE 4.1** provides a summary of the findings.

Next, the specific interrater reliability for each original child and interview VITAL category with at least a 1% base rate of occurrence was calculated. As seen in **TABLE 4.2**, the child VITAL categories revealed a wide range of kappas. When null sets ($K = 0$, because no rating was made in that category) were removed from the analyses, fair ($K = .21$ to $.40$) kappas were found for embarrassed ($K = .35$), whine ($K = .39$), cranky ($K = .36$), and interest ($K = .37$). The majority of categories received moderate kappas ($K = .41$ to $.60$): neutral ($K = .44$), anger ($K = .56$), disgust ($K = .41$), contempt ($K = .45$), sadness ($K = .49$), tense ($K = .50$), anxiety ($K = .46$), fear ($K = .49$), and withdraw ($K = .51$). Substantial kappas ($K = .61$ to $.80$) were found for pain/hurt ($K = .69$).

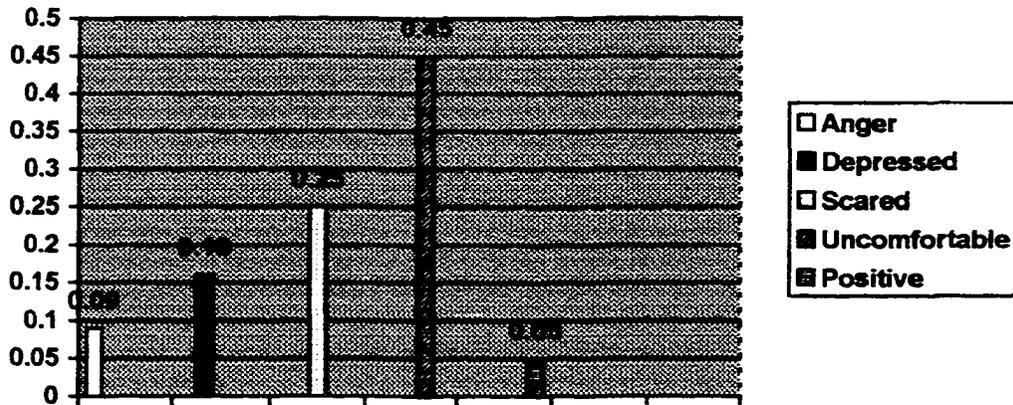
Finally, specific interrater reliability for the interviewer VITAL was calculated. As seen in **TABLE 4.3**, the interviewer VITAL category kappas revealed a strong chance-corrected agreement between raters. Specifically, when null sets ($K = 0$, because no rating was made in that category) were removed from the analyses, kappas ranged from moderate ($K = .41$ to $.60$) to almost perfect ($K = .81$ to 1.0) (Landis & Koch, 1977). Specifically, the categories that received moderate ratings included personal specific about alleged abuser ($K = .6$); personal closed ($K = .51$); personal closed about alleged abuser ($K = .48$); personal leading ($K = .5$); personal leading about alleged abuser ($K = .4$); personal recap ($K = .5$); and personal recap about alleged abuser ($K = .4$). A similar number of interviewer VITAL categories obtained substantial ($K = .61$ to $.80$) ratings: general specific ($K = .7$); general closed ($K = .63$); general recap ($K = .75$); personal open ($K = .74$); personal specific ($K = .66$); and legal system ($K = .62$). Finally, three categories were in the almost perfect range: none ($K = .98$); connector ($K = .9$); and rapport ($K = .9$).

Concurrent Criterion Validity of Child and Interviewer VITAL Ratings

The final phase of the data analyses involved testing the criterion validity of the VITAL, by assessing the hypothesized relationships between VITAL DVs and IVs. The first series of hypotheses concerned general base rate information about the behaviors of the children and interviewers that were measured by the VITAL during videotaped investigative interviews.

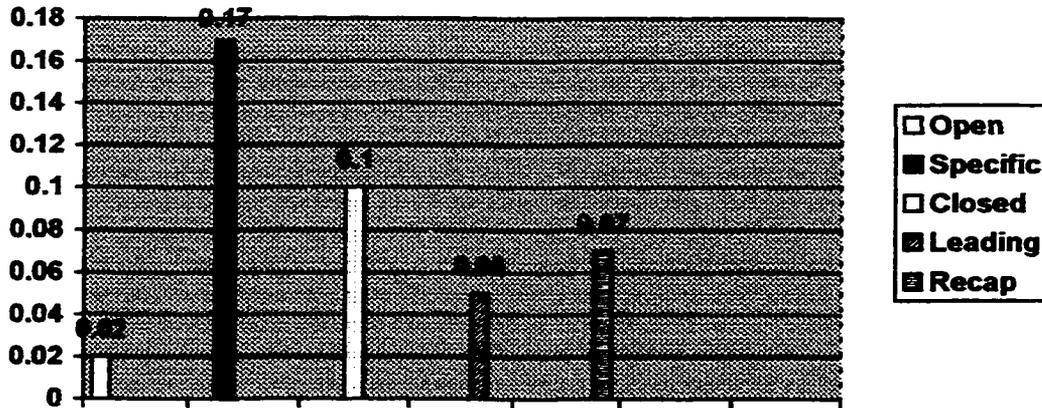
As hypothesized, VITAL child ratings across the entire sample of allegedly abused children varied, such that coders rated certain categories more often than they rated

others. Referring to the five-factor child VITAL categories, children were most often rated in the uncomfortable category (mean = .45, SD = .17). Also as hypothesized, children were least often rated as exhibiting positive emotions/behaviors (mean = .05, SD = .07). See **FIGURE 1**, Mean 5-Category Child VITAL ratings:

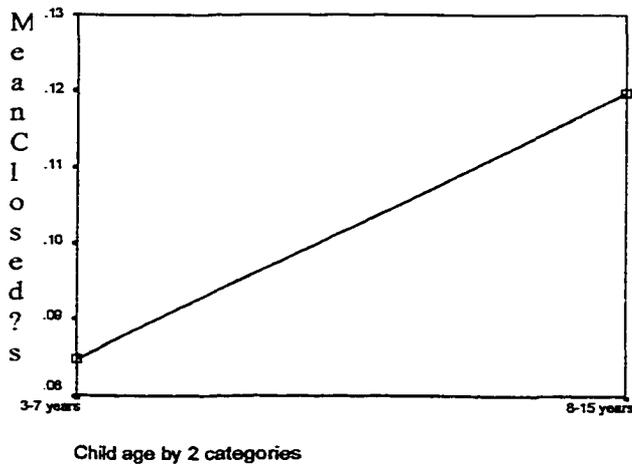


Accordingly, children were rated in the internalizing category more often (mean = .86, SD = .09) than in the externalizing category (mean = .20, SD = .15).

Similarly, VITAL interviewer ratings over the whole sample varied by question-type. As hypothesized, the most frequently asked question-types were specific (mean = .17, SD = .08) and closed (mean = .10, SD = .05), and least frequently were open (mean = .02, SD = .03) and leading (mean = .05, SD = .04). See **FIGURE 2**, Mean Interviewer VITAL Question-Type ratings:

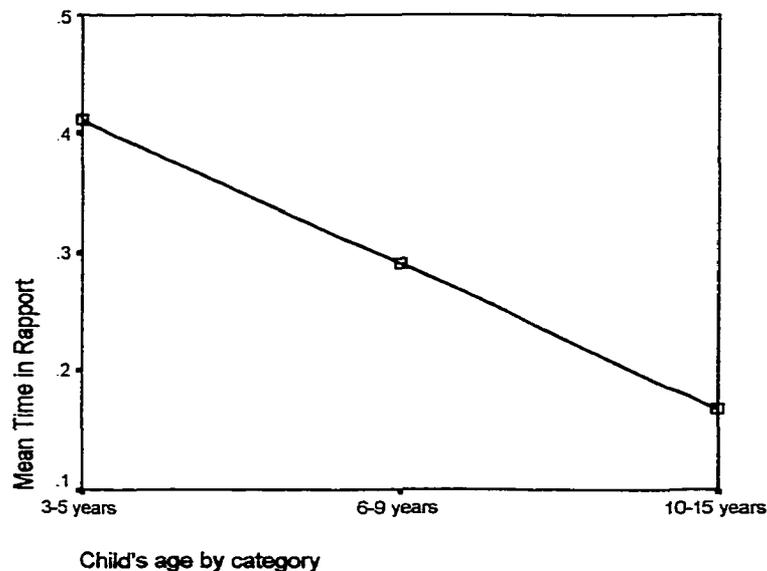


Next, it was hypothesized that ratings by interviewer question-type would vary by the age of the child. The hypotheses that interviewers would use fewer open-ended questions ($F = 1.179, p > .05 (p = .392)$), more specific questions ($F = 1.179, p > .05 (p = .287)$), and more closed questions ($F = 3.363, p > .05 (p = .077)$) with younger children than older children were not supported. In fact, there was a trend toward the use of fewer closed questions with younger children as compared to older children. See **FIGURE 3**, The effect of age on the proportion of closed question-types:



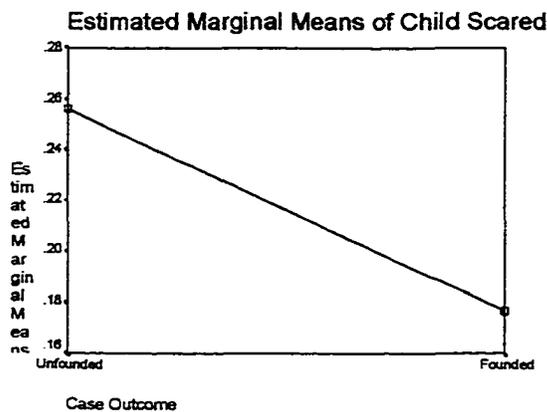
There was no statistically significant difference between the use of leading questions and the age of the child ($F = .591, p > .05 (p = .802)$). Examining the distribution of frequencies of this data, the highest proportions of time asking leading questions (10-16%) were spent with younger children than older children, as hypothesized.

Next, as hypothesized, interviewers spent significantly more time in rapport with younger children as compared to older children, especially pre-adolescent and adolescents ($F = 6.267, p < .05 (p = .006)$). Bonferroni post-hoc analyses revealed statistically significant differences between the scores of the youngest and oldest children (mean difference = -2.433, $p < .05 (p = .005)$). See **FIGURE 4**, Time in rapport by 3-age groups of children:



Regarding the relationship between child VITAL ratings and the outcome of the case, it was hypothesized that ratings should not differ when the child's case was considered founded as compared to when it was considered unfounded by the center,

although trends might be indicated. The main hypothesis of no differences between the founded and unfounded groups was supported. Mixed support was found for hypothesized trends. As hypothesized, very slight, non-statistically significant higher ratings were found in the founded group on anger ($F = .409$, $p > .05$ ($p = .530$), $\eta^2 = .019$) and depressed ($F = .591$, $p > .05$ ($p = .448$), $\eta^2 = .021$) than in the unfounded group. However, contrary to the trend hypothesis that higher scared ratings would be found in the founded group, there was a trend toward lower scared ratings in the founded group and higher ratings in the unfounded group ($F = 2.958$, $p > .05$ ($p = .097$), $\eta^2 = .102$). See **FIGURE 5**, Child scared ratings by case outcome:

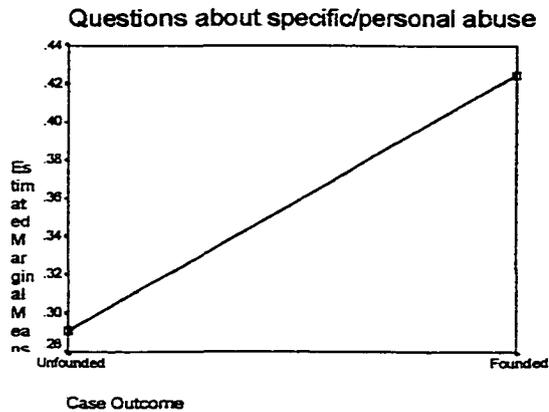


Also contrary to the trend hypothesis that higher ratings on positive emotions/behaviors would be found for children in the unfounded group, results indicate no difference between the groups, with non-significantly higher positive ratings in the founded group ($F = .348$, $p > .05$ ($p = .564$), $\eta^2 = .023$). Finally, contrary to the trend hypothesis that little difference would be found on the uncomfortable ratings for both groups, children were given non-significantly lower uncomfortable ratings when their

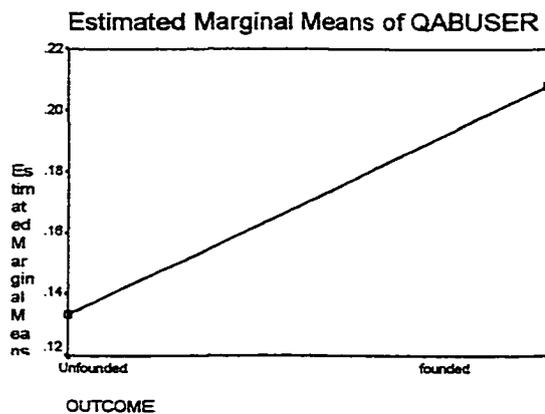
cases were deemed founded than when unfounded ($F = .890$, $p > .05$ ($p = .354$), $\eta^2 = .031$).

Similarly, regarding the relationship between interviewer VITAL ratings and the outcome of the case, it was hypothesized that time spent in rapport and general questions should not differ when the child's case was considered founded as compared to when it was considered unfounded by the center. No differences in rapport were found ($F = 1.582$, $p > .05$ ($p = .219$), $\eta^2 = .053$), although non-significantly less time was spent in rapport during founded cases as compared to unfounded cases. No differences were found in the proportion of time spent in asking general and education questions by case outcome ($F = 1.108$, $p > .05$ ($p = .301$), $\eta^2 = .038$), although in the non-significant direction of less time during founded cases as compared to unfounded cases.

In addition, it was hypothesized that more time would be spent asking personal abuse-related questions and alleged abuser questions in founded cases than in unfounded cases. As hypothesized, the amount of personal questions ($F = 12.430$, $p < .05$ ($p = .001$), $\eta^2 = .307$) was statistically significantly related to the case outcome, such that a higher proportion of personal abuse-related questions were asked in founded than in unfounded cases. See **FIGURE 6**, The proportion of personal interviewer questions by case outcome:



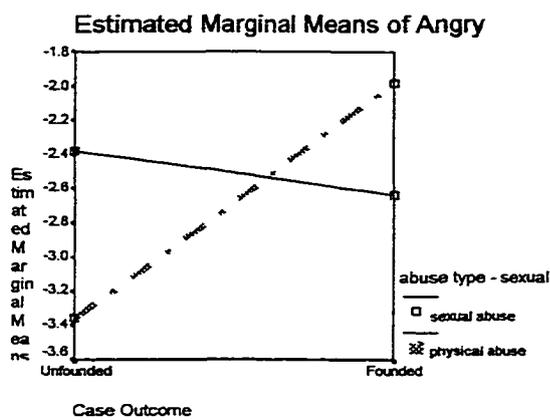
Similarly, as hypothesized, the amount of alleged abuser questions ($F = 4.405$, $p < .05$ ($p = .045$), $\eta^2 = .136$) was statistically significantly related to the case outcome, such that a higher proportion of alleged abuser questions were asked in founded than in unfounded cases. See **FIGURE 7**, The proportion of alleged abuser interviewer questions by case outcome:



Another hypotheses concerned the interaction between the case outcome and the type of abuse that the child allegedly suffered. One specific hypothesis was tested; that children from founded cases of physical abuse would be rated higher on anger than children from unfounded physical abuse cases, and children from both types of sexual

abuse cases. As hypothesized, there was a statistically significant interaction among anger, outcome and the type of abuse ($F = 4.279$, $p < .05$, $\eta^2 = .184$), such that children from founded cases of physical abuse were rated as more angry than children in unfounded cases of physical abuse, and than children in founded or unfounded cases of sexual abuse.

See **FIGURE 8**, Child angry ratings by type of alleged abuse and case outcome:

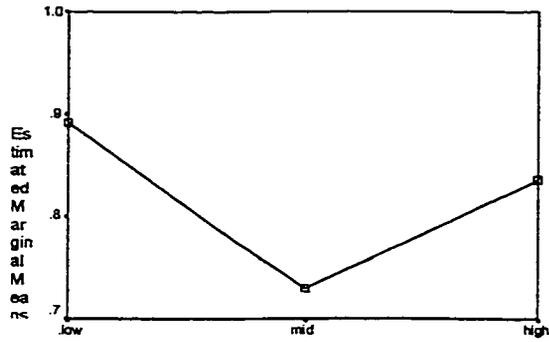


The last hypothesis regarding concurrent criterion validity concerns the interaction between case outcome and the child and interviewer VITAL ratings. Based on *Maryland v. Craig* rationale, it was hypothesized that children in the founded group would display higher amounts of emotional distress when asked personal questions about the alleged abuse, the alleged perpetrator, and court/criminal justice intervention, and lower amounts of emotional distress when asked about abuse in general, or about neutral/other issues. As previously discussed, the proportion of time spent in asking about criminal justice intervention was combined with the proportion of time asking about personal questions. Also because of low base rates, the proportion of time spent in asking about neutral/other questions was combined with the proportion of time asking about general questions. In

addition, the proportion of time for each category of question was subdivided into three groups for subsequent analyses. Low, medium, and high categories of time were selected based on approximate equal divisions (1/3) of the frequency of that question category. After these data transformations, three hypotheses were tested.

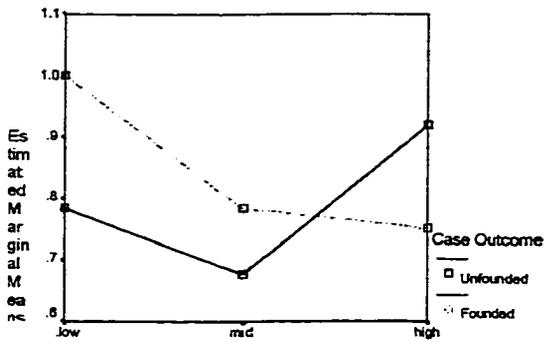
First, the hypothesis that higher proportions of emotional distress would be found when children from founded cases were asked personal questions about the abuse was partially supported. The results of a MANOVA indicated that both internalizing ($F = 4.426, p < .05 (p = .024), \eta^2 = .278$) and externalizing ($F = 3.609, p < .05 (p = .043), \eta^2 = .239$) were statistically significantly associated with the amount of personal abuse questions, and with the interaction between personal questions and the case outcome (*internalizing*, $F = 5.176, p < .05 (p = .014), \eta^2 = .310$; *externalizing* ($F = 4.823, p < .05 (p = .018), \eta^2 = .295$)). The relationship between case outcome and internalizing ($F = 1.066, p > .05 (p = .313), \eta^2 = .044$) or externalizing ($F = .650, p > .05 (p = .428), \eta^2 = .027$) was not statistically significant in the MANOVA. Examination of the marginal mean plots reveals that the hypothesis was more supported for child externalizing emotion/behaviors than for internalizing ones. See **Figure 9**, Child internalizing ratings by personal abuse questions; **Figure 10**, Child internalizing ratings by personal abuse questions and case outcome; **Figure 11**, Child externalizing ratings by personal abuse questions; and **Figure 12**, Child externalizing ratings by personal abuse questions and outcome:

Estimated Marginal Means of Internalizing



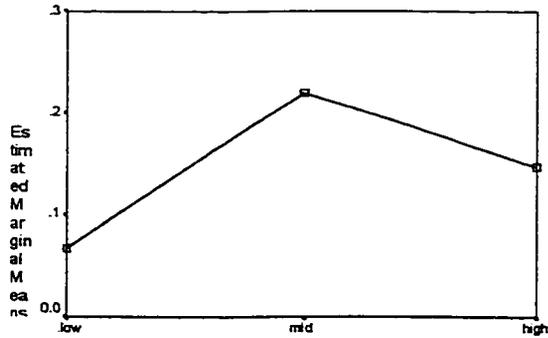
Specific questions by 3 amount categories

Estimated Marginal Means of Internalizing

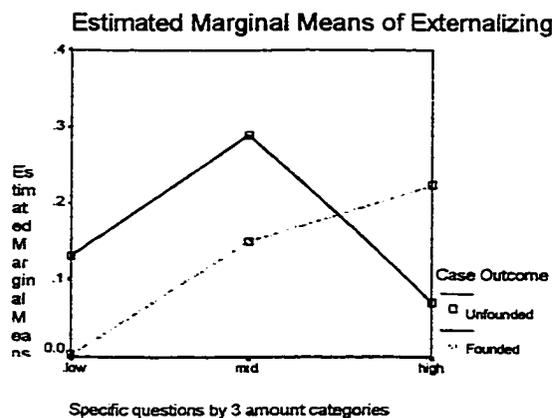


Specific questions by 3 amount categories

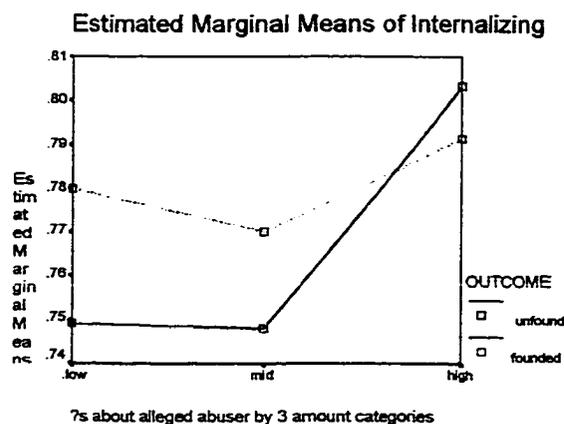
Estimated Marginal Means of Externalizing

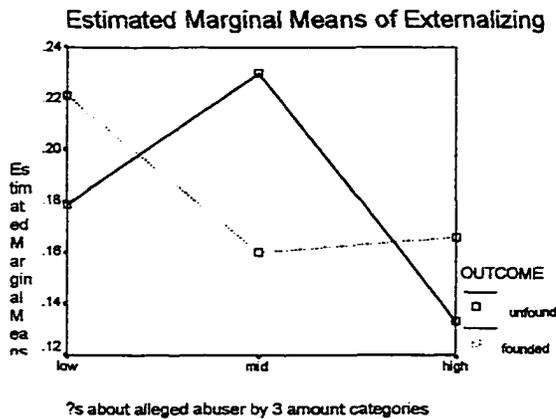
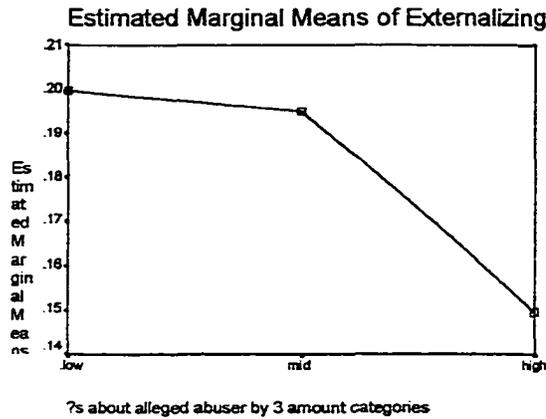


Specific questions by 3 amount categories



Second, the hypothesis that higher proportions of emotional distress would be found when children from founded cases were asked more questions about the alleged perpetrator was not supported. No statistically significant differences were found in the MANOVA analysis of the effect of alleged abuser questions and outcome on internalizing and externalizing ratings. See **Figure 13**, Child internalizing ratings by alleged abuser questions and case outcome; **Figure 14**, Child externalizing ratings by alleged abuser questions; and **Figure 15**, Child externalizing ratings by alleged abuser questions and outcome:

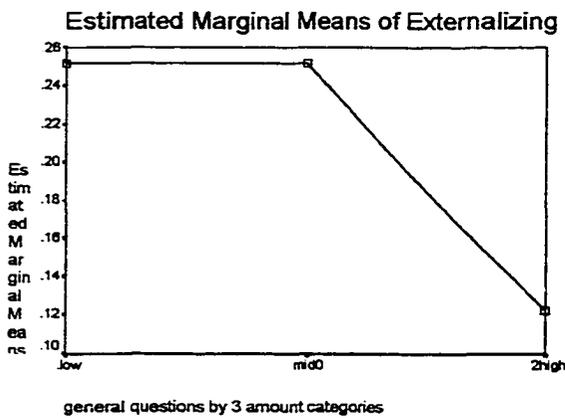
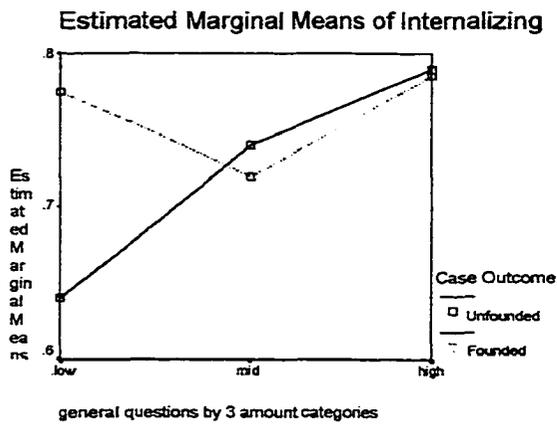
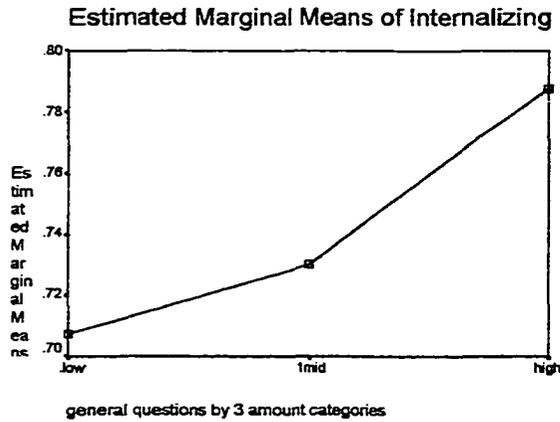


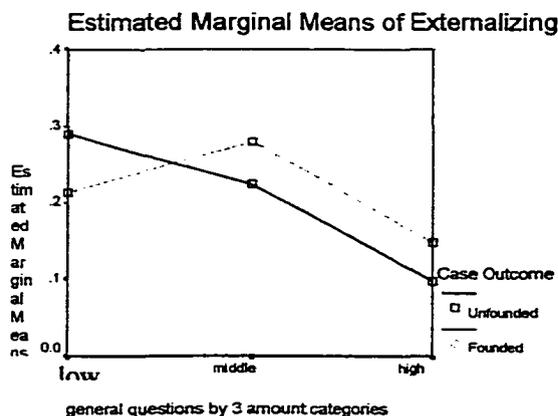


The final *Maryland v. Craig*-related hypothesis posited that lower proportions of emotional distress would be rated in children from founded cases who were asked general questions. To assess the effect of general questions and case outcome on internalizing and externalizing ratings, a MANOVA was conducted. No statistically significant differences were found, although there was a slight trend in the association between general questions and externalizing ($F = 2.584, p > .05$ ($p = .107$), $\eta^2 = .244$). See **Figure 16**, Child internalizing ratings by general questions; **Figure 17**, Child internalizing ratings by general

questions and case outcome; **Figure 18**, Child externalizing ratings by general questions;

and **Figure 19**, Child externalizing ratings by general questions and outcome:





For a summary of the results from this phase of the analyses, see TABLE 5.1, Concurrent Criterion Validity Results.

The Effect of Risk Factors on VITAL Ratings

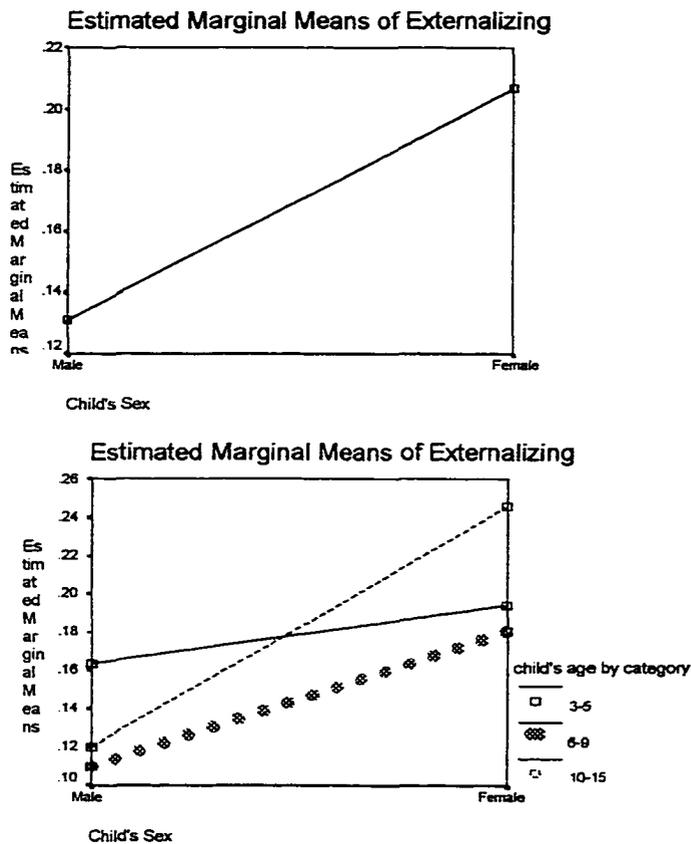
Mediating factors involving demographic, abuse-specific, and systemic variables were IVs next assessed in the present study.

The Effect of Demographic Risk Factors

The first phase of risk factor analyses involved assessing the effect of demographic factors (child's age and gender) on child VITAL ratings. As noted previously, ethnicity of the child could not be analyzed due to the majority of sample cases involving Caucasian children. In addition, the association between demographic risk factors and the determination of the case as founded or unfounded was explored.

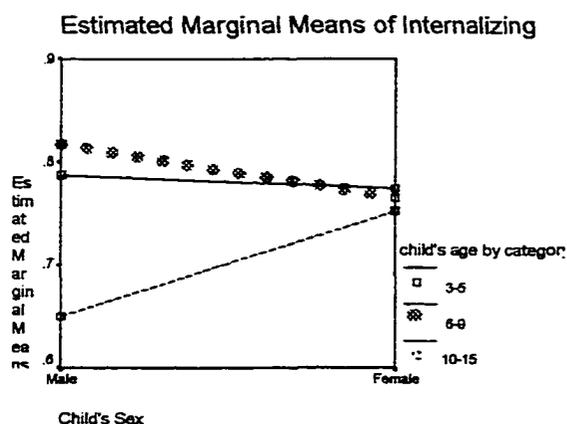
In general, as hypothesized the child's age and gender had minimal effect on the child's VITAL ratings. When comparing across allegedly abused children on externalizing by gender and age using a MANOVA, no differences were found on externalizing by *gender* ($F = 1.889$, $p > .05$ ($p = .183$), $\eta^2 = .076$), by *age* (using 3 categories) ($F =$

.465, $p > .05$ ($p = .634$), $\eta^2 = .039$), or by the interaction between *gender and age* ($F = .191$, $p > .05$ ($p = .828$), $\eta^2 = .016$). A slight trend was shown for the association between externalizing and gender in the opposite direction of the trend hypothesis (i.e., boys would be rated higher on externalizing than girls). See **FIGURE 20**, The proportion of externalizing by child gender, and **FIGURE 21**, The proportion of externalizing by child gender and age.

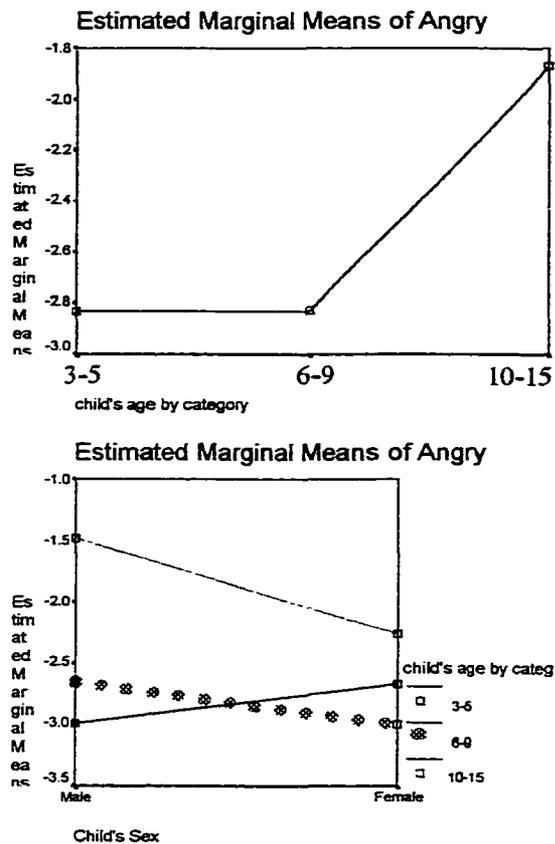


Similarly, no statistically significant differences or trends were found on internalizing ratings, as measured by the same MANOVA, by *gender* ($F = .148$, $p > .05$ ($p = .704$), $\eta^2 = .006$), by *age* (using 3 categories) ($F = .273$, $p > .05$ ($p = .763$), $\eta^2 =$

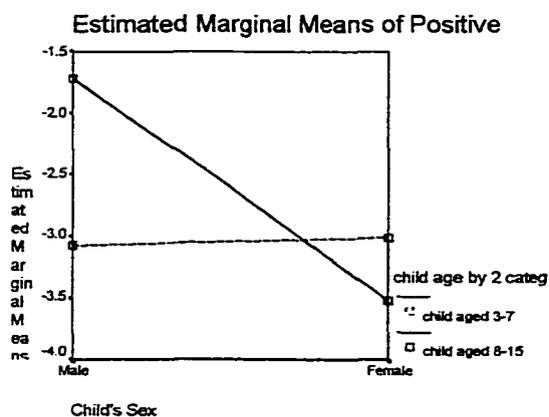
.023), or by the interaction between *gender and age* ($F = .476, p > .05$ ($p = .627$), $\eta^2 = .040$). See **FIGURE 22**, The proportion of internalizing by child gender and age.



Due to the trend in the association between externalizing and child gender, a univariate ANOVA was conducted to determine the association between the related 5-category child VITAL rating of anger, and the demographic variables. No statistically significant results were found for *gender* ($F = .346, p > .05$ ($p = .564$), $\eta^2 = .020$), or *gender by age* ($F = .406, p > .05$ ($p = .672$), $\eta^2 = .046$). However, a trend for the association between anger and age was indicated, such that older children were rated slightly higher on anger than the two groups of younger children (*age*, $F = 2.241, p > .05$ ($p = .137$), $\eta^2 = .049$). See **FIGURE 23**, The proportion of anger by child age, and **FIGURE 24**, The proportion of anger by child age and gender:



In addition, the positive child rating was tested for association with child gender and age using a univariate ANOVA. A trend was found for higher positive ratings being associated with the child's gender ($F = 2.228$, $p > .05$ ($p = .159$), $\eta^2 = .146$) and with the interaction between gender and age ($F = 2.547$, $p > .05$ ($p = .134$), $\eta^2 = .164$), but not with age alone ($F = .529$, $p > .05$ ($p = .480$), $\eta^2 = .039$). See **FIGURE 25**, The proportion of positive by child age and gender:



Finally, the gender of the child did not affect the determination of the case as founded or unfounded (Mann-Whitney $U = -.392$, $p = > .05$ ($p = .695$)). But, the different age groups of the children (3-5 years; 7-9 years; 10-15 years) were associated with the determination of the cases as founded or unfounded. Specifically, the 10-15 year old children were more likely to be from founded cases of abuse than the 3-5 year olds (Mann-Whitney $U = -2.197$, $p < .05$ ($p = .028$)), and the 7-9 year olds (Mann-Whitney $U = -2.152$, $p < .05$ ($p = .031$)). See **TABLE 3.3**, Case outcome by child age.

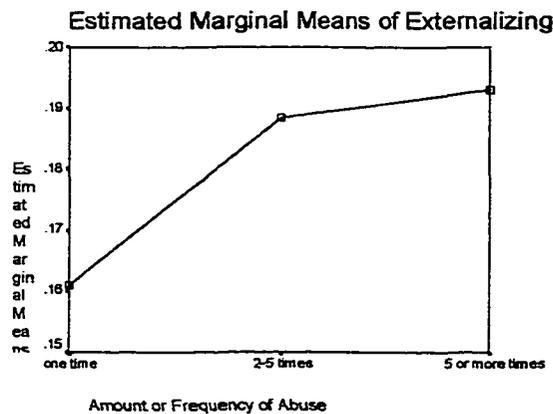
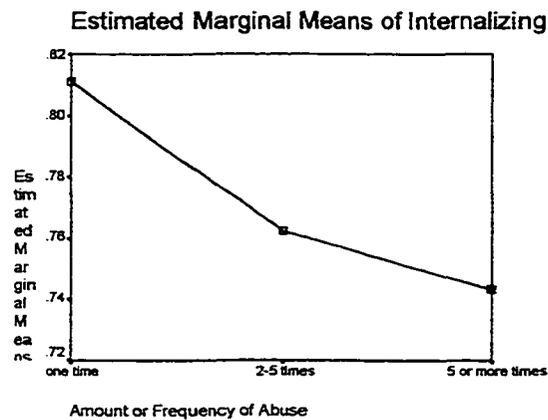
For a summary of the results from this phase of the analyses, see **TABLE 5.2**, Risk and protective factors results.

The Effect of Abuse-Specific Risk Factors

The second phase of risk factor analyses involved assessing whether five risk factors (i.e., greater frequency, longer duration of abuse, the use of threat or force, more invasive alleged sexual abuse, and closer relationship between the child and the alleged abuser) had the hypothesized minimal effect on child VITAL ratings. In addition, the

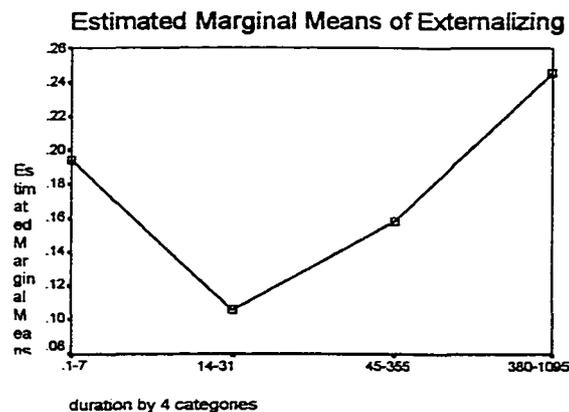
association between these risk factors and the determination of the case as founded or unfounded was explored.

First, as hypothesized, greater frequency of alleged abuse was minimally related to child VITAL ratings. No statistically significant difference in internalizing ratings ($F = .809, p > .05$ ($p = .456$), $\eta^2 = .059$) or externalizing ratings ($F = .189, p > .05$ ($p = .829$), $\eta^2 = .014$), was found across the three frequency of abuse groups (i.e., 1 time; 2-5 times; more than 5 times). See **FIGURE 26**, The proportion of internalize by frequency of alleged abuse, and **FIGURE 27**, The proportion of externalize by frequency of alleged abuse:



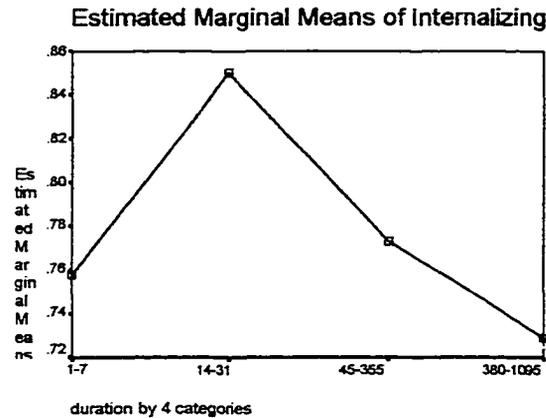
Moreover, the frequency or amount of abuse was not related to the case determination of founded or unfounded (Mann Whitney $U = -1.390$, $p > .05$ ($p = .164$)).

Second, there was mixed support for the hypothesis that a longer duration of abuse would be associated with higher VITAL ratings in the sample. Assessed using a MANOVA, externalizing ratings tended to differ by duration of abuse as measured in four categories (i.e., 1-7 days; 14-31 days; 45-355 days; 365-1095 days) ($F = 1.636$, $p > .05$ ($p = .207$), $\eta^2 = .164$). Examination of the estimated marginal means shows higher externalizing ratings in the over a year, and 0 to 1 week duration categories, as compared to the lowest ratings in the 1 week to 1 month period. However, a Bonferroni post hoc test revealed no statistically significant differences between each of the 4 categories, with the largest non-statistically significant difference between the 1 week to 1 month period and the over a year period (mean difference = $.1397$, $p > .05$ ($p = .276$)). See **FIGURE 28**, The proportion of externalize by duration of alleged abuse:

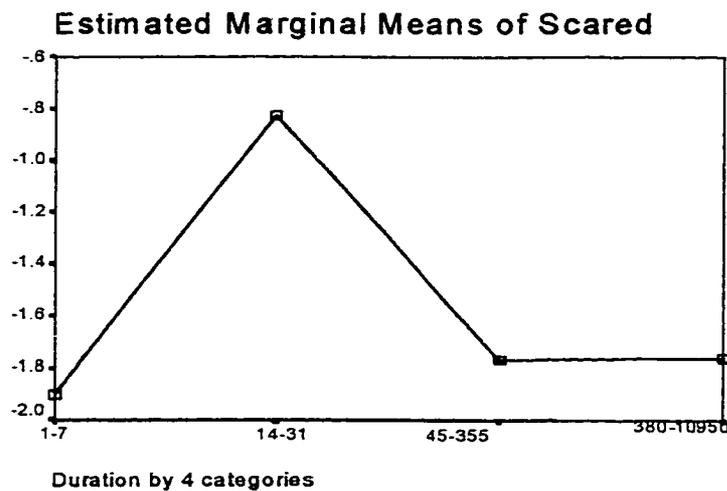
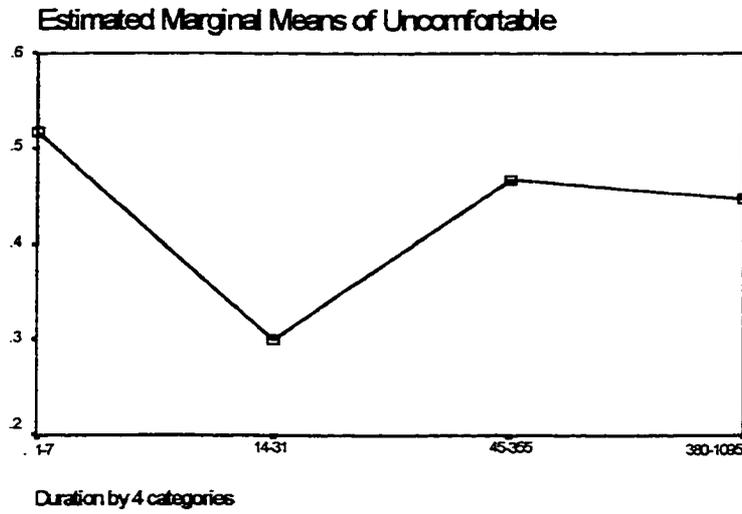


According to the same MANOVA, a slight trend toward higher internalizing ratings was found when the duration was one week to one month, as compared to lower interview

ratings in the other three time periods ($F = 1.087$, $p > .05$ ($p = .373$), $\eta^2 = .115$). Yet, a Bonferroni post hoc analysis revealed no statistically significant differences or trends between each of the 4 categories. See **FIGURE 29**, The proportion of internalize by duration of alleged abuse:



Because the results indicated a trend association between externalizing and internalizing ratings and duration of alleged abuse, the specific 5-category ratings were assessed. Both scared ($F = 2.341$, $p > .05$ ($p = .097$), $\eta^2 = .213$) and uncomfortable ($F = 2.111$, $p > .05$ ($p = .097$), $\eta^2 = .196$) showed a trend toward an association with duration. Specifically, higher scared ratings were observed in the one week to one month duration as compared to the other periods, and higher uncomfortable ratings were seen in the shortest and longer two duration periods as compared to lower ratings in the one week to one month time period. See **FIGURE 30**, The proportion of uncomfortable by duration of alleged abuse, and **FIGURE 31**, The proportion of scared by duration of alleged abuse:



Ratings did not differ on the remaining child VITAL categories between the four duration groups: angry ($F = .543$, $p > .05$ ($p = .658$), $\eta^2 = .079$); depressed ($F = .547$, $p > .05$ ($p = .655$), $\eta^2 = .059$); and positive ($F = .692$, $p > .05$ ($p = .573$), $\eta^2 = .138$). Finally, the duration of abuse was not related to the case determination of founded or unfounded (Mann Whitney $U = -1.557$, $p = > .05$ ($p = .120$)).

Third, as hypothesized, the use of threat or force was minimally related to child VITAL ratings. No significant differences were found on child ratings when the use of

threat or force was reported as compared to when it was not reported in the written files. Specifically, no statistically significant difference in internalizing ratings for use of threat or force was found in this sample ($F = .785$, $p > .05$ ($p = .383$), $\eta^2 = .028$). The MANOVA results for externalizing could not be used because the Levene test of equality of error variances indicated that error variances were not equal ($L = 4.662$, $p < .05$ ($p = .040$)). This statistical tool tests the null hypothesis that the error variance of the DV is equal across groups. Because test assumptions were violated, nonparametric analysis was employed to test the association between externalizing and threat or use of force. The Mann Whitney U test was not statistically significant, but indicated a trend association such that lower externalizing ratings were given to children whose cases included threat or use of force ($z = -1.580$, $p > .05$ ($p = .114$)). Also, the use of threat or force was associated with a founded or unfounded case determination (Mann Whitney U = -2.437, $p < .05$ ($p = .015$)).

Fourth, the degree of intrusiveness of the abuse was assessed in the subset of alleged sexual abuse cases. No statistically significant differences were found on internalizing ($F = .229$, $p > .05$ ($p = .960$), $\eta^2 = .096$) or externalizing ($F = .529$, $p > .05$ ($p = .777$), $\eta^2 = .196$) among ratings on the original 12 intrusiveness categories (see Appendix A).

To increase the cell count in the intrusiveness categories, a rational method was used to reduce them from 12 to 3 categories. The three new categories were fondling, oral contact, and penetration. This data transformation, however, did not result in statistically significant differences for internalizing ($F = .035$, $p > .05$ ($p = .966$), $\eta^2 = .004$) or

externalizing ($F = 1.107$, $p > .05$ ($p = .353$), $\eta^2 = .115$) and the three new categories. However, the degree of intrusiveness was associated with founded or unfounded sexual abuse cases (Mann Whitney $U = -2.778$, $p < .05$ ($p = .005$)).

Finally, as hypothesized, a closer relationship between the child and the alleged abuser was minimally related to child VITAL ratings. Because of low cell means in the original 7 categories (see Appendix A), the relationship categories were rationally reduced in two ways. First, no statistically significant differences were found among ratings on the internalizing ($F = .820$, $p > .05$ ($p = .495$), $\eta^2 = .090$) or externalizing ($F = .160$, $p > .05$ ($p = .922$), $\eta^2 = .019$) categories when compared across 4-categories of relationship types (acquaintance; in-home sibling or adult; out of home parent; in-home caretaker) using a MANOVA. Similarly, with another MANOVA, no statistically significant differences were found among ratings on internalizing ($F = .446$, $p > .05$ ($p = .510$), $\eta^2 = .016$) or externalizing ($F = .431$, $p > .05$ ($p = .517$), $\eta^2 = .016$) for cases involving alleged abusers who lived in the child's home as compared to those who lived out of the child's home. Lastly, relationship was not associated with a founded or unfounded case determination (Mann Whitney $U = -.111$, $p > .05$ ($p = .911$)).

For a summary of the results from this phase of the analyses, see TABLE 5.2, Risk and Protective Factors Results.

The Effect of Systemic Risk Factors

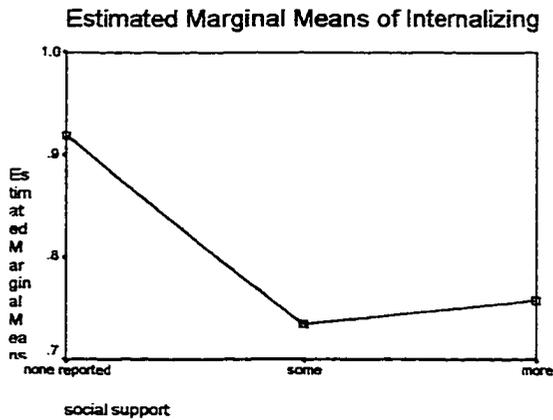
The third phase of the risk factor analyses concerned variables associated with systemic factors. Two hypotheses involved the child's family environment and social support, and two involved secondary traumatization risk factors.

First, it was hypothesized that children from unstable homes would be rated higher on the child VITAL. A MANOVA analysis indicated that children from unstable homes did not have higher VITAL ratings than did children from stable homes. Specifically, internalizing ($F = .012$, $p > .05$ ($p = .914$), $\eta^2 = .000$) and externalizing ($F = .121$, $p > .05$ ($p = .731$), $\eta^2 = .004$) ratings did not differ between the two groups on ratings of the stability of the child's family environment.

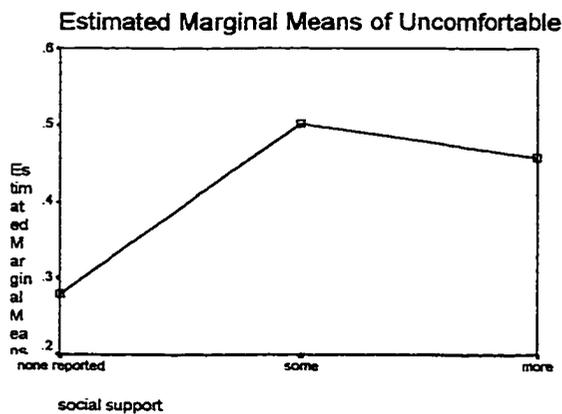
Family environment was also tested against child positive emotion/behaviors using an ANOVA. However, the results (which indicated only very minimal evidence for unstable homes being associated with less positive ratings) could not be used because the Levene's test of equality of error variance was statistically significant ($z = 5.332$, $p < .05$ ($p = .036$)). Because test assumptions were violated, a non-parametric analysis was employed to test the association between positive and family environment, but it was not statistically significant (Mann Whitney U, $z = -.557$, $p > .05$ ($p = .577$)). Moreover, family environment was not associated with a founded or unfounded case determination (Mann Whitney U, $z = -.392$, $p > .05$ ($p = .695$)).

Second, the hypothesis that children with none or limited sources of social support would have higher VITAL ratings than children with more social support was also partially supported. As hypothesized, children with no reported sources of social support had statistically significantly higher internalizing ratings ($F = 4.828$, $p < .05$ ($p = .016$), $\eta^2 = .271$) than children who had one or more support sources. Bonferroni post hoc analysis revealed statistically significant differences between the none and some categories (mean difference = .1860, $p < .05$ ($p = .016$)), and the none and more categories (mean difference

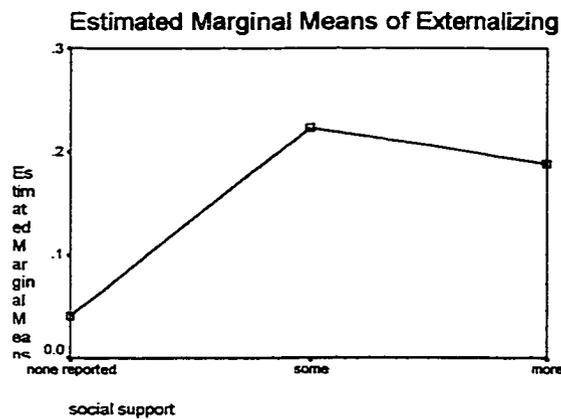
= .1629, $p < .05$ ($p = .033$) of social support. See **FIGURE 32**, The proportion of internalizing by social support:



However, children with no reported sources of social support showed a trend toward lower uncomfortable ratings ($F = 3.001$, $p < .05$ ($p = .066$), $\eta^2 = .182$), than children who had one or more support sources. Bonferroni post hoc analysis on uncomfortable also revealed a trend differences between the none and some categories (mean difference = .2218, $p > .05$ ($p = .068$)). **FIGURE 33**, The proportion of uncomfortable by social support



In addition, children with no sources of social support were rated as statistically significantly lower on externalizing than children one or more sources of social support ($F = 4.513$, $p < .05$ ($p = .021$), $\eta^2 = .258$). Bonferroni post hoc analysis showed statistically significant differences between the none and some categories (mean difference = .1829, $p < .05$ ($p = .018$)), and a trend differences between the none and more categories (mean difference = .1478, $p > .05$ ($p = .059$)). **FIGURE 34**, The proportion of externalizing by social support



Ratings did not differ between the three groups on the remaining child VITAL categories: angry ($F = .567$, $p > .05$ ($p = .576$), $\eta^2 = .054$); depressed ($F = 1.035$, $p > .05$ ($p = .369$), $\eta^2 = .071$), and scared ($F = .826$, $p > .05$ ($p = .449$), $\eta^2 = .058$). Moreover, social support was not associated with case outcome of founded or unfounded (Mann Whitney U, $z = -1.234$, $p > .05$ ($p = .217$)).

Regarding systemic secondary traumatization risk factors, the hypothesis that an increase in the number of investigative interviews or interviews conducted by trained professionals would be associated with higher child VITAL ratings was not supported.

When comparing one interview with more than one interview (range of 2-4 interviews) on 2 category child VITAL ratings with a MANOVA, there was no significant difference between a child being rated as internalizing ($F = 1.319, p > .05$ ($p = .261$), $\eta^2 = .047$) or externalizing ($F = 1.512, p > .05$ ($p = .229$), $\eta^2 = .053$). A very slight non-significant trend indicated that children who experienced more than one interview were rated higher on internalizing, and lower on externalizing. In addition, a univariate ANOVA indicated a trend toward fewer positive ratings ($F = 1.638, p > .05$ ($p = .220$), $\eta^2 = .098$). But, the number of interviews was not associated with a determination of founded or unfounded (Mann Whitney $U = -.888, p > .05$ ($p = .375$)).

The final systemic risk factor concerned the use of rapport during investigative interviews. The hypothesis that children would have lower VITAL ratings of uncomfortable when higher proportions of time were spent in rapport was not supported ($F = .271, p > .05$ ($p = .607$)). This finding was not affected by the age of the child. Specifically, when the effect of age was controlled (ANCOVA), there was no difference in the uncomfortable ratings (age $F = .852, p > .05$ ($p = .364$), η^2 ; rapport $F = .000, p > .05$ ($p = .999$)).

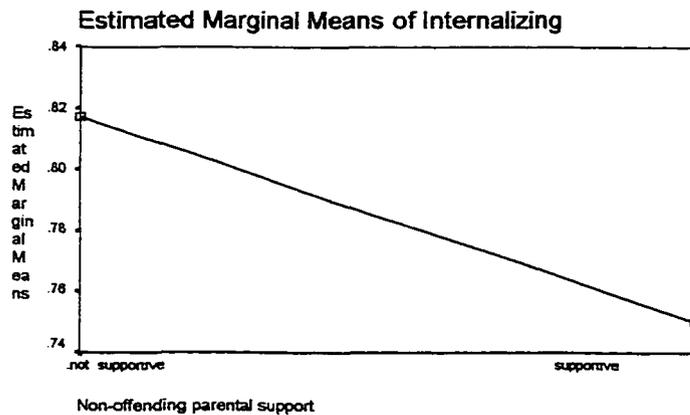
For a summary of the results from this phase of the analyses, see **TABLE 5.2, Risk and Protective Factors Results.**

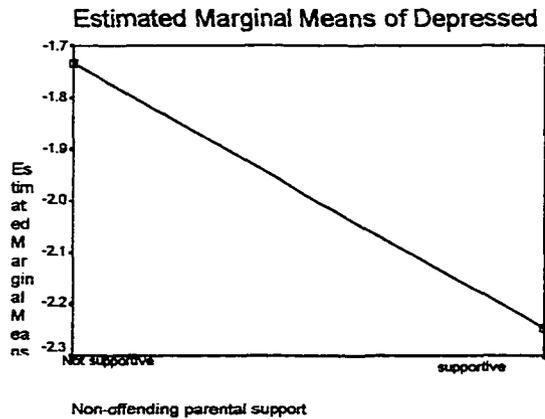
The Effect of Protective Factors on VITAL Ratings

The final phase of the analyses in the present study involved assessing the effects of hypothesized protective factors (i.e., parental support; social support) on child VITAL ratings. Specifically, social support and non-offending parental support were hypothesized

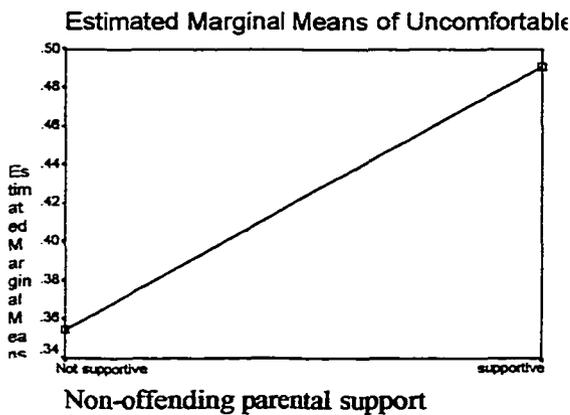
to be associated with lower child VITAL ratings of emotional distress in the entire sample, with small differences between the founded and unfounded abuse groups. Mixed support for the hypotheses was found.

First, some support was found for the hypothesis that non-offending parental support was related to lower proportions of emotional distress ratings in children. Specifically, there was a trend to support the finding that the presence of such support was associated with lower ratings of internalizing ($F = 2.061$, $p > .05$ ($p = .163$), $\eta^2 = .071$), and depression ($F = 2.684$, $p > .05$ ($p = .113$), $\eta^2 = .087$). See **FIGURE 35**, The proportion of internalizing by non-offending parental support, and **FIGURE 36**, The proportion of depressed by non-offending parental support:





However, parentally-supported children were found to have statistically significantly higher ratings of uncomfortable ($F = 4.814$, $p < .05$ ($p = .037$), $\eta^2 = .147$), and a trend toward higher angry ratings ($F = 1.534$, $p > .05$ ($p = .229$), $\eta^2 = .068$), than children apparently without such support. See **FIGURE 37**, The proportion of uncomfortable by non-offending parental support:



Ratings did not differ between the two groups on the remaining child VITAL categories: externalizing ($F = .413$, $p > .05$ ($p = .526$), $\eta^2 = .015$); scared ($F = .177$, $p > .05$ ($p = .677$), $\eta^2 = .006$); and positive ($F = .539$, $p > .05$ ($p = .474$), $\eta^2 = .035$). Moreover,

parental support was not associated with founded or unfounded cases (Mann Whitney U, $z = -1.175$, $p > .05$ ($p = .240$)).

Finally, the hypothesis that children with one or more sources of social support would have lower VITAL ratings than children with no reported social support was tested. As reported in the risk factor section, children with one or more support sources had statistically significant lower internalizing ratings than children who had no reported sources of social support. Yet, contrary to expectations, children with one or more sources of social support were rated as statistically significant higher on externalizing (and a trend toward being more uncomfortable) than children with no sources of social support. Moreover, social support was not associated with case outcome of founded or unfounded (Mann Whitney U, $z = -1.234$, $p > .05$ ($p = .217$)).

For a summary of the results from this phase of the analyses, see **TABLE 5.2, Risk and Protective Factors Results.**

DISCUSSION

The results of the present study indicate that the VITAL meets appropriate standards of reliability and concurrent criterion validity. Results also reveal the ability of the VITAL to assess the effect of risk and protective factors on the emotional distress ratings of allegedly abused children during investigative interviews.

Interrater Reliability

The chance-corrected measures of interrater agreement revealed moderate to excellent interrater reliability for the overall VITAL, and fair to substantial/excellent interrater reliability for the individual categories. As expected, interrater reliability was negatively affected by the complexity of the coding system (especially for the child VITAL) and by the frequency of the target behaviors (Dorsey, Nelson, & Hayes, 1986). To address these common problems associated with behavioral analyses, the present study analyzed VITAL categories that were modified in accordance with the interrater reliability data, and with rational methods informed by research and other measures of child emotional distress. Future research on the VITAL should explore the use of the modified categories with larger samples of videotaped child-interviewer dyads.

Concurrent Criterion Validity of Child and Interviewer VITAL Ratings

The results of the present study reveal support for the concurrent criterion validity of the VITAL. In other words, results from the present study indicate that the VITAL can be used to measure children's existing status, and the impact of interviewer questions and statements on children's emotions and behaviors. Given this basis, future research can now be pursued to follow open cases of alleged child abuse in order to test the VITAL's

predictive criterion validity (to predict future outcomes) and construct validity (through correlation with other tests and convergent and discriminant validation) (Anastasi, 1988). The future outcome of primary interest for this measure is the child's discussions of alleged abuse in interviews with attorneys or during in-court testimony. Once the process of validation is completed, the VITAL should be generalizable and ecologically valid for use in courtroom accommodation decisions.

Specifically, results of VITAL ratings over the entire sample of allegedly abused children were in accordance with previous research regarding the emotions/behaviors of children and interviewers during child abuse investigative interviews. As hypothesized, the interviews were generally stressful for the children, as reflected by relatively high average proportions of time during which children were rated as uncomfortable (45%), and by low proportions of positive emotions/behaviors (5%) as measured by the child VITAL. In addition, children were observed expressing a variety of more severe negative emotions/behaviors during the investigative interviews at fairly high rates: scared (25%), depressed (16%), and angry (9%).

When comparing children from founded and unfounded cases, however, no significant differences were found on these ratings, although some directions were slightly indicated (i.e., higher anger and depressed ratings but lower scared and uncomfortable ratings in founded cases). One explanation for the lack of statistically significant differences involves the limitations of the sample (i.e., small sample size; cases labeled unfounded may contain "real" abuse and witnessing abuse experiences). Another reason may be due to the interaction with the type of abuse. In the one hypothesis that was tested

using this rationale, statistically significant results were found; namely, children from founded physical abuse cases were rated higher on anger than children from unfounded physical abuse cases and children from both outcomes of sexual abuse cases. This finding lends support to the work of Shields and Cicchetti (1998) who posited that physically maltreated children may be at higher risk for reactive aggression than other abused children and non-maltreated children.

Also consistent with measurements in other samples of investigative interviews, overall results from the interview VITAL indicated that the most frequently asked question-types were specific (17%) and closed (10%) and the least frequent were open (2%) and leading (5%). For example, the proportion of leading questions in the present study was similar to that made by prosecutors in recent, actual criminal child sexual abuse cases (approximately 5%), but lower than asked by the defense attorneys (approximately 30%) (Zajac, Gross, & Hayne, 2000). A low frequency of leading questions may be viewed as a positive result because leading questions can have a detrimental effect on the accuracy of children's reports of abuse (Ceci & Bruck, 1993). However, a low proportion of open questions can be seen as problematic because open questions are the best way to "create an opportunity for a child to provide the most independent and complete description of an event as possible in their own words" (Saywitz & Camparo, 1998, p. 831).

Some results from the interviewer VITAL indicated a departure from previously-reported interviewing practices. It had been hypothesized that fewer open-ended, and more specific and closed questions would be used with younger children. However, results

indicated no differences between these question-types and child age. In fact, there was a trend toward interviewers asking *fewer* closed questions with younger children. Because closed questions (e.g., yes/no; forced choice) have been shown to be the most problematic for young children (Peterson & Biggs, 1997), a trend toward the use of fewer closed questions with this age group is developmentally appropriate and may result in more accurate interviews (Jung & Bavelas, 2000). Thus, such results should not lessen the criterion validity of the VITAL because the findings indicate that interviewers may have been conducting developmentally appropriate interviews, consistent with national guidelines. Previous survey research has shown that many interviewers do not follow such guidelines (Oberlander, 1995). Thus, it is probable that research data from older interviews will reflect a different level of practice, and produce some different results from the present study.

Further support for the developmental appropriateness of the interviews in the present study was shown by overall results on rapport. Interviewers spent higher proportions of time in rapport-related questions/statements with younger children as compared to older children, especially adolescents. This result may reflect appropriate investigative interviewing practices because more time is needed to assess the child's developmental level and ability to communicate with the investigative interviewer (Saywitz & Camparo, 1998). In addition, it appears that the interviewers in the present study appreciated the importance of the rapport phase because results indicated that an average of 30% of the interview was spent asking rapport-related questions, and no statistically significant differences in report were found based on the outcome of the case. The

development of rapport has been associated with an increase in children's ability to cognitively process information at an high level, which can result in better memory retrieval and statement accuracy (Ceci & Bruck, 1993; Saywitz & Camparo, 1998).

In addition, as hypothesized, interviewers spent statistically significantly higher proportions of time asking about personal abuse-related issues and about the alleged abuser in founded cases as compared to unfounded cases. An explanation for this finding might be that interviewers spent more time talking about abuse-related issues with children who were responsive and/or whose explanations involved more detail, and less time with children who were not responsive and/or less detailed. Based on discussions with the Center's investigative interviewers and their supervisors, the latter type of case would probably result in an unfounded determination, unless other corroborating information were obtained (e.g., confession; independent medical evidence). Notwithstanding, avoiding repetitive (and presumably longer) questioning with non-responsive children is a sound practice, because such techniques are considered suggestive and may lead to less accurate, and even false or exaggerated reports of abuse (Bruck, Ceci, & Melnyk, 1997).

It is important to note, however, that conclusions cannot be drawn about the quality of investigative interviews in the present study. The analyses in the present study of the VITAL simply measure the proportion of time spent during these interviews in asking different categories of questions/statements - the content of the question is not assessed. In addition, the impact of different interviewing techniques on the child is not fully addressed in the present study. Sequential analyses that would connect interviewer question/statements with the child's emotional response over time reflect a more

appropriate statistical technique that could be employed to address this issue. In addition, other studies that specifically assess adherence to national investigative interview guidelines could address this issue. Future studies with the VITAL should explore these other related areas of research, and are planned by the author.

The final aspect of concurrent criterion validity concerns the ability of the VITAL to address *Maryland v. Craig* related-issues. Based on the Supreme Court's decision and the empirical research supporting its rationale, the present study hypothesized that children in the founded group would display higher amounts of emotional distress when asked personal questions about the alleged abuse, the alleged abuser, and court/criminal justice intervention, and lower amounts when asked about abuse in general and neutral/other issues. As previously mentioned, the low base rates of ratings in the court/criminal justice intervention categories precluded separate analysis of this issue. The ratings from those categories were combined with ratings in the personal abuse-related questions. This data transformation was selected so that comparisons could be made between that category and the alleged abuser questions category, in rough accordance with one of the *Maryland v. Craig* tests – that the child would suffer emotional distress due to presence of the alleged abuser rather from the courtroom generally.

The results reveal mixed support for the hypotheses. First, the proportion of personal abuse-related questions was statistically significantly associated with both internalizing and externalizing ratings in the children overall, and in founded versus unfounded cases. The hypothesis was somewhat supported with regard to externalizing emotion/behaviors, such that higher proportions of personal questions resulted in higher

externalizing emotion/behaviors for the allegedly abused children as a group, and for children from founded cases as a group. Children from unfounded cases, however, were rated similarly on externalizing at the low and high ends of proportions of personal abuse questioning. Regarding internalizing ratings, however children from the founded group were rated lower on internalizing when asked more personal abuse-related questions, and children from the unfounded group were rated as the most internalizing when asked the most personal questions.

Second, investigative interviewers' questions/statements about the alleged abuser did not have a statistically significant effect on the children's externalizing or internalizing ratings. In addition, no trends were indicated. For purposes of discussion, examination of the marginal means plots revealed the highest ratings of internalizing for the founded group when the most questions about the alleged abuser were asked (.79) as compared to when lesser amounts of such questions were asked (.78 and .77). A similar pattern of internalizing ratings was found for the unfounded group, but with slightly larger differences among ratings (.75 - .80). Conversely, lowest rated of externalizing were found when the most questions about the alleged abuser were asked of children from founded and unfounded groups.

Finally, the proportion of general and educational questions/statements did not have a statistically significant difference in internalizing or externalizing ratings, although a trend in the hypothesized direction was found for externalizing. That is, when the most general questions were asked of children from both founded and unfounded cases, the children were rated lowest on externalizing. Examination of the marginal mean plots

revealed the opposite findings for internalizing ratings, such that the highest proportions of general questions resulted in the highest internalizing ratings for children from founded and unfounded cases.

There are several reasons that might explain the lack of statistically significant results and somewhat conflicting findings regarding these hypotheses. First, it is likely that the way that the alleged abuser categories were coded might not have been refined enough to capture variation in child affect. However, a more conservative approach to rating references to the alleged abuser would lower the base rates of this response even further, and as a result, attempts to assess its effects on larger samples of children would be even more difficult.

Second, the analyses in the present study did not record the content of interviewer questions. It is possible that certain types of questions might evoke strong emotional/behavioral reactions in different children. Other studies that employ sequential analyses of each question and answer pair would be able to better account for the differences due to question content.

Third, the category transformations might have blurred distinctions between the groups in three ways. For one, the legal setting questions might have produced effects in a different direction than the personal questions. Second, the personal question category included the low base rate, personal-with-alleged-abuser subcategory, which probably reduced the differences between the larger category groups. Third, transforming some of the categories so that they were normally distributed removed some of the variation in the sample, and possibly these findings. While it was recognized that the general linear model

is fairly robust to skewed data (Iverson & Norpoth, 1987), the present study transformed the data because it sought to evaluate the new measure against more conservative analyses (i.e., those in which it would be harder to obtain statistically significant findings because outliers would not be able to influence findings in their direction).

Finally, while the analyses were able to compare children from founded and unfounded cases, they were not able to compare children within those groups due to small cell sizes. A future, larger study of the VITAL should perform such analyses because courtroom accommodations that implicate the defendant's right to confrontation are not available to assist all children from founded cases. Instead, the goal of assessment in these cases is to identify a *subset* of children within the founded group who are most in need of accommodation. Without such analyses, the VITAL could still function as a screening tool but would have less specific utility in these cases. In either case, once fully developed, the VITAL must be used in conjunction with other measures when assessing an individual allegedly abused child for courtroom accommodation purposes or for other, more general, clinical reasons.

The Effect of Risk and Protective Factors on VITAL Ratings

The second main objective of the present study was to assess the effect of mediating and moderating factors, including those related to secondary traumatization, on the emotions and behaviors of a sample of allegedly sexually and physically abused children during investigative interviews. In general, the risk factors did not have statistically significant effects on the emotional/behavioral states of allegedly abused children. These findings were consistent with the hypotheses of no difference between the

groups, which was based on the conflicting evidence for the effect of risk factors in the literature. Yet, support for the hypothesized protective factors was demonstrated.

First, no difference in the demographic factors of age and gender were found. Some trends were indicated, and not all in the trend-hypothesized direction. In general, there was a trend for girls to be rated higher on externalizing than boys. This result was different from the trend-hypothesis, but may be due to sampling differences (i.e., 21 girls and 9 boys). It may also be due to differences in the ways girls are raised and socialized and/or how girls are viewed by society. Since the past decade, girls can not only play sports with boys, but they may also join gangs with boys, or form their own. Now, girls who commit crimes are being recognized and adjudicated by a “get-tough” criminal justice system (Grisso, 1996; Hsia & Beyer, 2000), and statistics reflect higher numbers of juvenile delinquents/offenders who are female (Coolbaugh & Hansel, 2000).

Next, consistent with the trend-hypothesis, older children and adolescents (10 to 15 years of age) were rated more on angry than younger children (3 to 5 and 6 to 9 years of age). In addition, the oldest children were more likely to be from founded cases than were younger children (3 to 5 and 6 to 9 years of age). This result is consistent with studies that have found older children more credible (e.g., Wood et al., 1996) than younger children because it is believed that older children can communicate more effectively and give more complete statements than can younger children (Faller, 1988). At the same time, however, studies have found younger children to be more credible in sexual abuse cases because it is believed that younger children have less sexual knowledge than older children, and thus would be less able or likely to fabricate such knowledge

(Kovera & Borgida, 1996; Pence & Wilson, 1994). Regarding gender, no differences were found on case outcome by child gender. This result is unlike some studies that have found girls to be more credible than boys because boys gave less accurate accounts of abuse details and showed less emotion during disclosure (Faller, 1988; Wood et al., 1996).

Second, regarding the abuse-specific risk factors, results from the present study found that four of the five factors (i.e., higher frequency of alleged abuse, threat or use of force, closer relationship between the alleged abuser and the child, and higher degree of intrusiveness of sexual abuse) were not related to higher ratings of internalizing or externalizing in the sample of allegedly abused children. There was mixed support for the trend-hypothesis that a longer duration of abuse would be associated with higher VITAL ratings of emotional distress in the children.

Specifically, there was a trend finding that children whose abuse allegedly lasted the longest (1–3 years) were rated highest on externalizing. But, the time period of a week to a month of alleged abuse duration appeared to result in a trend toward the highest internalizing, scared and uncomfortable ratings, as compared to the other shorter and longer time periods. In addition, the direction of these results is consistent with the examination of the marginal means in the frequency data - higher externalizing and lower internalizing ratings for more frequent and longer duration of times. The connection between duration and frequency data has been shown in a review of previous studies (Williams, 1993). But, as previously discussed, the short- and long-term effects of abuse have not been consistently associated with duration and frequency. Some studies have shown that longer duration led to more severe behavioral and emotional problems (e.g.,

Urquiza, 1988), but other studies have shown no relationship or an effect in the opposite direction, as in the present study (e.g., Chandler, 1982; Courtois, 1979; Williams, 1993). More research is needed on this factor to resolve continuing differences, especially research that involves assessing the cumulative effect of a number of risk factors. For example, a review of findings on abuse-specific factors revealed that sexual abuse of greater frequency, greater force, and more physically harmful or threatening by multiple perpetrators results in more severe psychological problems (Barker-Collo, Melnyk, & McDonald-Miszczak, 2000).

Determinations of a case as founded or unfounded was associated with the threat or use of force, and the degree of intrusiveness of sexual abuse allegations, but not with frequency, duration, or the child's relationship with the alleged abuser. A possible reason for these findings is that investigators might more rigorously pursue cases with more serious or aggravating charges than those viewed as less serious or harmful. For example, in the Center's state, molestation of a child is a class 2 felony, while sexual abuse of a child involving only touching of the breast is a lower class 3 felony. Also, longer sentences may be imposed when use or threat of a weapon is involved.

The last category of risk factors tested were family dysfunction, lack of social support, increased number of interviews and lack of interviewer rapport. Determinations of a case as founded or unfounded was not associated with any of these systemic risk factors. Contrary to hypotheses, no statistically significant differences were found for the effect of family dysfunction, number of professional interviews, or time in rapport. However, there was a trend that more interviews were related to higher internalizing

ratings, lower positive ratings, and lower externalizing ratings. These results are consistent with prior research already discussed that has shown repeated investigative interviews to have negative, harmful effects on children. Instead of expressing anger toward the interviewers, children may tend to blame themselves and think that something is wrong with them or their account when interviewed repeatedly. Children who have suffered abuse may also be traumatized by having to re-tell and re-live their abusive experiences.

The lack of statistically significant findings for family dysfunction can be partially explained by the present study's method. The use of closed cases protected the confidentiality of the information, but did not allow for information to be gathered from the children and their families. Moreover, the detectives' written files were not the ideal source for information regarding the nature of the child's family environment. The corresponding child protective services files would be a better source of such information, but permission to access these files was not granted to the author.

Also, the results regarding rapport may be due to the fact that in general, interviewers spent a substantial amount of the total interview time in rapport. It may also be because the amount of time spent in rapport is not a good indication of the quality of the rapport. Other studies should examine not only the content of the rapport questions, but also the process of establishing rapport for a measure of its quality and resulting effect on the child interviewees.

Finally, social support was considered as both a risk and protective factor in the present study. Consistent with both sets of hypotheses, children with no reported sources of social support (i.e., zero to one supportive person) had statistically significantly higher

internalizing ratings than children who had one or more support sources. Put another way, children with one or more sources of social support were lower on internalizing. Similarly, the results from children whose non-offending parent(s) was supportive showed a trend toward lower internalizing and depression ratings than did children whose parent was not supportive. Results also showed that children with one or more sources of social support were rated as more externalizing and uncomfortable than were children without social support. Similarly, children with a supportive non-offending parent were statistically significantly higher on uncomfortable, and a trend toward being rated as higher on anger.

As hypothesized, these findings are consistent with the literature on protective factors. Although the present study could not assess the children's internal coping mechanisms and attributions (i.e., no researcher-conducted interviews with the children were permitted), it attempted to behaviorally measure emotions/affect that have been empirically associated with internalizing and externalizing. Thus, a possible explanation for the results of social and parental support may be that supported children may feel more comfortable at speaking out against or being angry at the non-supportive alleged perpetrator and/or at the interviewer who asks about the alleged offender and abusive event(s). For example, one child who went to the home of a church official after being allegedly physically abused by his father expressed high amounts of anger during the interview. Without such support, children may feel isolated, turn inward, and see themselves as at fault; thus expressing more internalizing affect. These children may also be depressed and anxious when thinking about the consequences of their disclosure to investigators. For example, another child who lived most of her life in foster care because

her parents abandoned or neglected her, expressed high amounts of depression. Because the case against her father who allegedly sexually abused her was founded, the child would again be placed in temporary care because she did not have any other family to assist her.

In sum, a variety of risk and protective factors were assessed using the VITAL on the present study's sample of actual child abuse investigative interviews. The results obtained using this new assessment instrument were consistent with the research on investigative interviewing and on risk and protective factors, thus lending support to the concurrent criterion validity of the VITAL. Although the interviews were generally an uncomfortable experience for the children in the present study, the presence of one or more sources of social support was found to result in lower internalizing ratings (i.e., sadness, anxiety, embarrassment, withdrawal). Indeed, the hypothesized protective effect of social support was the most robust finding in the second phase of the present study.

Thus, future research, including research on the VITAL, should be conducted to continue improving the assessment and measurement of the effects of social support as well as other resiliency, coping and risk factors in abused children and adolescents. Behavioral instruments like the VITAL are virtually absent and much needed in the field of psycho-legal research (McAuliff & Kovera, 2000), and in the world of actual practice with allegedly abused children and their families. It is hoped that such research will be used to improve assessment and treatment programs as well as inform public policies and practices that can help these young members of our society when they become involved in the criminal justice system – and avoid causing them harm.

APPENDIX A
VITAL - Record Review Form (Short Version)

Project Code Number: M _____

Initials of Data Coding Person & Date: _____

Initials of Data Entry Person & Date: _____

Data Entry Code	Data Entry Heading	Code	Definitions of Code		
cage	Child's age at interview				
cgen	Child Gender		1=Male 2=Female		
ceth	Child Ethnicity		1=European American; 2=African Am. 3=Native Am.; 4=Hispanic; 5=Asian 10=Other; 11=Unknown		
czip	Child Zip Code				
igen	Interviewer Gender		1=Male 2=Female		
ieth	Interviewer Ethnicity				
cons	Consent Obtained		1=No; 2=Yes		
assent	Child's Assent Obtained		1=No; 2=Yes		
intnum	Number of interviews				
Famdys	Family Dysfunction		1=No 2=Yes		
Famdtyp	Family Dysfunction Type		1=divorce 2=Parental Criminal Behavior; 3=Parental Psych Dx; 4=Parental Abuse; 5= Neglect; 10=Other; 11=Unknown		
Sup	Support Person Present		1=No; 2=Yes		
Socsup	Social support		1=No; 2=1 person; 3=>1 person		
Vqual	video quality problems		1=No; 2=Sound; 3=Picture; 10=Other		
Pnum	# of Allegd Perps		11=Unknown		
Valid	Founded/corroboratd?		1=No; 2=Yes (2.1=perp. Confession; 2.2=medical ev.; 2.3=other corrob)		
Type: 1=sexually suggestive talk, unwanted hugs&kissing; 2=vicars exposure to sexuality; 3=visual exposing of genitals, voyeurism or viewing porn; 4=Fondle over ch's clothes; 5=Fondle under clothes; 6=Simltd intercourse over clothes 7=Child masturbating abuser or simulated intercourse under cloths 8= Oral-abuser to ch; 9=Oral-ch to abuser; 10=Digital or object pen; 11=Vaginal or Anal intercourse; 12=paraphilic sex/explt 13=Ritual or satanic	Alleged Perpetrator Type; Relationship; Age; Gender; Ethnicity Relationship: 1=Stranger; 2=neighbr; acquaint, babysitter; 3=Involved adult living outside home-teacher, minister, daycare, relative (e.g., gparent); 4=in-home sibling or step-sib; 5=In-home adult, except parent/step-p; 6=Out of home parent; 7=In-home prim. caretaker (parent, guardian, step-p; 10=Other; 11=Unkwn T R A G E	<i>Number of Occurrences of Each Type</i>	Coercion Used & Type: 1=Verbal/Threat(1.1=secrecy; 1.2=phys harm; 1.3=other; 1.4=unkn); bribery, trick, authority figure 2=Weapon (2.1=gun; 2.2=knife; 2.3=other; 2.4=unkn); 3=Physical 10=Other; 11=Unkn; 12=N/A	Time of Most Recent Occurrence: 1=in past week; 2=in past mo.; 3=1-6mos 4=6-12mo. 5=12-24 mo 6=>than 24 mos. 11=Unk Duration of Most Recent Occurrence: [Duration= Age at most recent occurrence minus Age at onset]	
ptl		ptloc	ptlco	liti	ld

APPENDIX B

VITAL CHILD CODING DEFINITIONS AND PRECEDENCE RULES (10/28/98)**Precedence (Decision) Rules:**

1. All codes are mutually exclusive and exhaustive – code only one category per 10-second interval.
2. When 2 or more codes are witnessed during a 10-second interval, *code the one that is seen the most...*
3. *Negative Codes take precedence over positive codes;*
4. *Tense/Anxious/Fear takes Precedence over all other negative codes;*
5. *The lower intensity code (1) is chosen over a higher intensity (2) if you are unsure within a code;*
6. *Neutral* is coded when you are not sure or have a question about a code.

Definitions:

1. UNOBSERVABLE is coded when the child's emotions and behavior cannot be rated by the observer due to lack of visibility of the child's face, body posture, and/or audio of the child's verbal content.
 2. NEUTRAL: FACS = 0
No other code applies – no obvious emotional reaction by facial expression or body language; information exchange; matter of fact discussion of abuse allegations; voice tone is flat, even, and evenly paced; "I don't know what it is" code.
 3. ANGER (1): FACS = 4 + 5; 4 + 23; 4 + 5 + 7 + 10 + 23/25; 7 (close scrutiny) (the eyebrows are lowered and drawn together, the eyelids are tensed, and the eye appears to stare in a hard fashion; the lips are either tightly pressed or parted in a square shape); a strong and in context 24 (lip press).
Words may be biting, abrupt, often with a key word or syllable highly stressed; Voice tone may be raised or lowered - controlling type anger is shown by a lowering in the volume and uniformity of syllables with a definite edge to the voice; clenched jaw.
- BELLIGERENT/AGGRESSIVE (2): FACS = 2; sticks out tongue; yelling; 4
Child's anger is directed outward – to the interviewer; attempt to challenge the interviewer by use of force or to get a rise out of him/her; physical aggression and contact, such as hitting, throwing, biting, punching, kicking, and other destruction of the environment.

4. **DISGUST** (1) FACS = 9 (nose wrinkle) + 10 (upper lip raised) (may be unilateral); 9 + 17 (nose wrinkled and chin raised).
An involuntary reaction related to rejection of a perceived noxious stimuli, situation, or person's behavior; the child sounds sickened, repulsed; may say "gross," "eew," "yuk."

CONTEMPT/SCORN (2): FACS = eye roll; 14 (dimpler, may be unilateral); 10 + 14.
Intended to communicate a clear lack of respect or hatred; an icy, distant stance that may
Include derision, mockery, put downs, sarcasm, name calling, and hostile humor.

5. **SAD** (1): FACS = 1 (inner brow raised); 1 + 4 + 15 (lip corners down) + 17 (chin raised); generally characterized by a low volume of voice and slowness of speech; voice tone may quiver or be abnormally high in pitch; the child may sigh (not a huff); look for pauses in places where you wouldn't expect them or just long pauses as if it takes effort; resigned/passive; head and eyes may be cast down; child may express grief or remorse; "I'm sad/depressed"

PAIN/HURT (2): FACS = 1 + 4; cry = 6 + 15 + 25 (mouth open); 4 + 10 + 20 + 25 (grimace); affect more pronounced in that it is more plaintive, desperate; crying. Also includes expressions of physical pain, partial or guarded movement, and sounds such as moans, grunts, and groans.

6. **TENSE (FIDGET)** (1) a low level expression of tension as shown by fidgeting (moving/rocking parts of the body, nail biting, hair twisting, thumb sucking, picking at carpet/toys/clothes/sofa), shifting positions, and speech disturbances (uhs or ahs, stuttering, several incomplete or unfinished thoughts in a speaking turn).

ANXIOUS (2) a higher level expression of anxiety shown through elevated amounts of the tense code, and including additional behaviors such as pacing, nervous or inappropriate laughter, fixed or wary gaze; asks to go to the bathroom; tells you s/he is nervous or worried.

FEAR (3) FACS = 1 + 2 + 5 + 20, and 25/26/27.

The highest level expression by the child that s/he feels unsafe, scared, fearful; wets pants; child may go into a fetal position and start rocking back and forth; lip may quiver or be "swallowed up" by the child.

7. **EMBARRASSED** (1) expression of concern with negative thoughts and reactions of others ("what will they say?"); child may look away or down & avoid eye contact; blushes

SHAME (2) clear expression of self-blame or self-deprecation or putting self down (“it’s all my fault”); child may look away or down & avoid eye contact; blushes

8. WHINE (1): non-defensive complaint or reflection of dissatisfaction and fatigue “I’m tired”; characterized by a thin edge to the voice, irritating nasal quality, high-pitched, and a sing-song voice quality, often with one syllable stressed at the end of the sentence.

FRUSTRATED/CRANKY (2) intensified expression such that the child may be expressing frustration, defensiveness “nu-uh, but I didn’t do that”; or having an all out temper tantrum.

9. DISSOCIATION WITHDRAWL (1) the child is withdrawing from the interaction but not in a way that appears active or intentional. Appears as a daze or trance-like state; appears “spaced out.” May also be shown by a rapid age regression such that the child may refer to self in 3rd person, talk to an imaginary companion, talk to self, or use a different voice.

STONEWALL WITHDRAWL (2) the child is withdrawing from interaction in an active way, as if to say, “I’m not going to listen to you!” Turn head; clench jaw; look disinterested; engage in away behavior (pick, fidget, perhaps in an exaggerated manner); offers no connectors (e.g., nod, uh huh); flat facial expression such that face appears flat and frozen; in extreme cases they may have a clenched jaw and more pronounced neck muscles.

10. INTEREST (1) Positive energy expressed in relation to what the other person has said; communicating an active interest in the other person or activity; curiosity

AFFECTION (2) Direct expression of affection and supportiveness; voice may be slow with a drop in amplitude or pitch and relaxation in voice; expression of understanding.

11. HAPPY (1): FACS = 12 (smile) + 6 (making it a genuine smile) + sometimes 7. Look for twinkle in eye. Giggle/laugh

EXCITED (2) FACS = 1 +2; Exclamation Point! Burst of positive energy, often with rapid changes in physical features like volume, pitch, stress, and rate of speech; includes expressions of surprise (“oh my God!” “You’re kidding!”) seen as FACS 1 + 2 + 5 + 26

APPENDIX C

VITAL INTERVIEWER CODING DEFINITIONS AND PRECEDENCE RULES
(11/3/98)

Precedence Rules:

1. When a partial and a full question/statement are made during a 10 second interval, *code the full one.*
2. When 2 or more questions/statements are made during a 10 second interval, *code the latter one.*

Definitions:

1. NONE is coded when the interviewer is not speaking during the interval.
2. OTHER is coded as a default when no abuse-related material or other codes (e.g., rapport) are being discussed.
3. CONNECTOR is coded when the interviewer makes a non-verbal (e.g., mm-hmm, ah, oh) or verbal (e.g., I see) communicator that serves to keep the child talking. No indication made of the most desirable response.
4. RAPPORT is coded when the interviewer asks or talks to the child about neutral topics in an attempt to put the child at ease as well as to assess the child's developmental level. This may include questions about school, pets, activities, knowledge of numbers and alphabet, as well as identification of body parts, assessments of the difference between truth and lies, etc. It also includes introduction to the interviewer and the interview process.
Although possibly a questionable interviewing practice, use this code when the interviewer gives the child praise or a verbal affirmation ("good," "you're smart").
5. ABUSE GENERAL is coded when the interviewer is talking about or asking the child about child abuse in any form (e.g., physical, emotional) to any real or hypothetical person other than the child. This includes any (non-child-specific) question about "good/bad touch," and questions about or references made to "uncomfortable," "bad," "hurtful," or such other negative description of physical touching. Also questions/statements about secrets and coercion.
6. ABUSE PERSONAL is coded when the interviewer is talking about or asking the child about his/her own alleged abuse experience(s). May include general ?s as well as specific incidents (e.g. specific day, time, place)
7. LEGAL SYSTEM is coded when the interviewer asks or talks to the child about any aspect of the legal system, including past, present, or future contact or feelings about

or experiences with the law, legal rules, (but not telling the difference between the truth and a lie), attorneys, judges, courtrooms, testimony, juries, probation

*****A for alleged abuser is coded in ANY category when the interviewer refers to the alleged perpetrating abuser (NOT the nonoffending caregiver – unless s/he was not protective) in ANY manner.** This would include questions/statements about the name/identity of the alleged abuser, what s/he looks like, what s/he said, his/her body parts, what the child feels or thinks about him/her, etc.

There is no generally accepted way to categorize questions. So, our research is helping to fill this gap in the literature. So, these definitions are subject to revision and review. Thanks for your help! See Poole & Lamb (1998), Memorandum of Good Practice (see Westcott, Davies, & Horan, 1998), and Davies & Seymour (in press).

Question Types:

1. OPEN-ENDED:

In general, open-ended questions are intended to produce a multiple word response. While technically they might be answered with a yes or no, they could also be answered with a narrative response. In addition, open-ended questions can be truly open free-recall questions, or can be more focused when topics have been brought up by the child. There are theoretically no limits on the range of answers and no indication of the most desirable response.

Some examples:

Tell me about that. Can you tell me about that? I wonder if you could tell me about that.
Can you tell me more about the man in the playground? – only if child mentioned man in playground
And then what happened? Is there anything (more/else) that you would like to tell me?

2. SPECIFIC or FOCUSED yet Non-Leading:

Specific questions ask about a particular detail or concept that has already been brought up. They are used to guide the child to provide relevant evidence, but must do so non-directively. They may be answered in a single word & are most often be seen as who, what, when, where, how & why questions.

Some examples:

What color was her hair? What was his name? What time is your recess? Who else was there?
Where in the playground did you see the man? Why was he standing there? How did he get there?
How did that make you feel?
Do you remember what you were doing when he came over to you?

3. CLOSED:

Closed questions are a subset of specific questions that have a limited number of response alternatives. They include multiple choice, forced choice, and yes-no questions that require a yes or no answer. There is no indication of the most desirable response.

Some examples:

Was he inside or outside your house? Was he in the living room, or bathroom, or...?

Was the man in the playground's jacket red, blue, or another color? Or can't you remember?

Did she offer you some candy? Did he do that? Was his name John? Was your mom there?

4. LEADING

Leading questions are questions in which the question implies the answer or the most desirable response. They may be of the following types:

Open loaded: No limits on the range of answers, but most desirable response indicated – “Why didn't you say to him ‘stop that’?”

Specific or closed loaded: Limited range of responses, but one response is indicated as more desirable than another. New information is introduced into the question. Did it happen in the bedroom or someplace else? May include yes/no questions. “He would come over, right?” “And then he hit you, didn't he?”

5. RECAP - The RECAP category includes both

REPEATED STATEMENTS: Repetition of information just given by the child, using the child's words – Child says, “He held my hand;” interviewer says, “He held your hand.”

And **RECAPPING STATEMENTS**: Repetition of information already given by the child earlier in the series of questions with the same interviewer. Summarizes or recaps what the child has said – “And you said that your father was over at the school?”

TABLE 1.1, Proportion of time in original child VITAL categories

Rater1	ang	bellig	disgust	contp	sad	pain	tense	anx	fear
M1	0.1	0	0	0	0	0	0.5	0	0
M2	0	0	0	0	0	0	0.3	0	0
M3	0	0	0	0	0	0	0.6	0	0
M4	0	0	0	0	0.1	0	0.6	0	0
M5	0	0	0	0.1	0.2	0.1	0.2	0	0
M6	0	0	0	0	0	0	0.2	0	0
M7	0.1	0	0	0.1	0	0	0.4	0	0
M9	0	0	0	0.1	0	0	0.4	0	0
M12	0	0	0	0	0.1	0	0.5	0	0
M13	0	0	0	0	0.1	0	0	1	0
M14	0	0	0	0	0	0	0.1	1	0.1
M19	0	0	0.1	0.3	0.1	0	0.3	0	0
M21	0	0	0	0	0.3	0.1	0.3	0	0
M23	0	0	0	0	0.1	0.2	0.4	0	0
M24	0	0	0	0.1	0	0	0.5	0	0
M26	0	0	0	0.1	0	0	0.4	0	0
M28	0	0	0	0	0	0	0.3	0	0
M29	0	0	0	0	0.2	0	0.3	0	0
M34	0	0	0	0	0.1	0	0.3	0	0
M35	0	0	0	0	0	0	0.3	0	0
M36	0	0	0	0	0.1	0	0.2	0	0
M37	0	0	0	0.1	0.1	0	0.5	0	0
M38	0	0	0	0.2	0	0	0.4	0	0
M39	0	0	0	0	0.1	0	0.2	0	0
M41	0	0	0	0.4	0	0	0	0	0
M42	0	0	0	0	0	0	0.4	0	0.1
M43	0	0	0	0	0	0	0.3	0	0.1
M44	0	0	0	0.1	0	0	0.2	1	0
M45	0	0	0	0	0	0	0.3	0	0
M49	0	0	0	0.1	0	0	0.7	0	0
M50	0	0	0	0	0	0	0.7	0	0
Mean of F	0	0	0	0.1	0.1	0	0.3	0	0

TABLE 1.1, Continued

F	emb	sham	whine	crnk	w/dr	int	afect	happ	excited
m1	0	0	0	0	0	0	0	0	0
m2	0	0	0.1	0.2	0	0	0	0	0
m3	0	0	0	0	0.1	0	0	0	0
m4	0.1	0	0.1	0	0	0	0	0	0
m5	0	0	0	0.1	0	0	0	0	0
m6	0	0	0.2	0.2	0.1	0	0	0	0
m7	0.2	0	0	0	0	0	0	0	0
m9	0.1	0	0.1	0	0.1	0	0	0	0
M12	0	0	0	0	0	0	0	0	0
M13	0	0	0	0	0.1	0	0	0	0
M14	0	0	0	0	0	0	0	0	0
M19	0	0	0	0	0	0	0	0	0
M21	0	0	0.1	0.1	0	0	0	0	0
M23	0.1	0	0	0	0.1	0	0	0	0
M24	0	0	0.1	0.1	0	0	0	0	0
M26	0	0	0	0	0.1	0	0	0	0
M28	0	0	0.2	0	0.4	0	0	0	0
M29	0	0	0	0	0.1	0	0	0	0
M34	0	0	0	0	0	0	0	0	0
M35	0	0	0.1	0	0.1	0	0	0	0
M36	0	0	0.2	0.1	0.1	0	0	0	0
M37	0	0	0	0	0	0	0.1	0	0
M38	0	0	0	0	0	0	0	0	0
M39	0	0	0.2	0	0.1	0	0	0	0
M41	0	0	0	0.2	0	0	0	0	0
M42	0	0.1	0	0.1	0	0	0	0	0
M43	0	0	0.1	0	0.1	0	0	0	0
M44	0	0	0.1	0.1	0	0	0	0	0
M45	0.1	0	0.3	0	0	0	0	0	0
M49	0	0	0	0	0	0	0	0	0
M50	0	0	0	0	0	0	0	0	0
Mean	0	0	0.1	0	0	0	0	0	0

TABLE 1.2, Proportion of time in modified child VITAL categories

	3	4	5	6	7	8	9	10	11
	ang	bellig	disgu	contp	sad	pain	tens	anx	fear
			st				e		
M1	0.08		0.03		0.05		0.68		
M2f	0		0.05		0.05		0.57		
M3	0		0		0		0.82		
M4	0		0.02		0.05		0.64		
M5	0.02		0.06		0.28		0.52		
M6	0		0.04		0.03		0.44		
M7	0.07		0.09		0.02		0.55		
M9	0		0.08		0		0.52		
M12	0.02		0.01		0.1		0.75		
M13	0.02		0.04		0.09		0.61		
M14	0		0		0		0.96		
M19	0.02		0.37		0.07		0.45		
M21	0.01		0.04		0.34		0.36		
M23	0		0		0.37		0.43		
M24	0		0.06		0.05		0.7		
M26	0.02		0.14		0.04		0.67		
M28	0		0		0		0.28		
M29	0		0		0.24		0.58		
M34	0.04		0.04		0.13		0.51		
M35	0		0.05		0.01		0.69		
M36	0.01		0.03		0.11		0.42		
M37	0.01		0.1		0.07		0.57		
M38	0.06		0.23		0.03		0.49		
M39	0.03		0.02		0.11		0.52		
M41	0.06		0.41		0.01		0.28		
M42	0		0.01		0.04		0.71		
M43	0		0		0		0.72		
M44	0.04		0.06		0		0.73		
M45	0		0		0.01		0.63		
M49	0.03		0.06		0.06		0.81		
M50	0		0		0.02		0.92		
Mean	0.02		0.07		0.08		0.6		

TABLE 1.2, Continued

	12	13	14	15	16	17	18	19	20
	emb	sham	whine	crnk	w/dr	int	afect	happ	excited
M1	0.02		0		0	0.12		0.02	
M2f	0.01		0.3		0.01	0		0	
M3	0		0		0.05	0.13		0	
M4	0.13		0.13		0	0.03		0	
M5	0.04		0.08		0.01	0		0	
M6	0.07		0.33		0.08	0.01		0	
M7	0.17		0		0	0.1		0	
M9	0.09		0.18		0.12	0		0	
M12	0.02		0.06		0	0.02		0.01	
M13	0		0.02		0.11	0.05		0.07	
M14	0		0		0.04	0		0	
M19	0.07		0.02		0	0		0	
M21	0.01		0.23		0.01	0		0	
M23	0.13		0		0.06	0		0	
M24	0.03		0.13		0.02	0.01		0	
M26	0.02		0.04		0.06	0.01		0	
M28	0		0.21		0.36	0.15		0	
M29	0		0.05		0.08	0.05		0	
M34	0.01		0.02		0	0.19		0.06	
M35	0.04		0.06		0.11	0.02		0.02	
M36	0		0.32		0.11	0		0	
M37	0.01		0.01		0	0.19		0.03	
M38	0.02		0.04		0.02	0.09		0.01	
M39	0		0.23		0.09	0.01		0	
M41	0.02		0.23		0	0		0	
M42	0.05		0.14		0	0		0.04	
M43	0.02		0.07		0.13	0.06		0.01	
M44	0		0.15		0.02	0		0	
M45	0.05		0.3		0	0		0	
M49	0		0.02		0	0		0	
M50	0.03		0.02		0	0		0	
Mean	0.03		0.11		0.05	0.04		0.01	

TABLE 1.3, Proportion of time in 2-category child VITAL

	Externalizing	Internalizing
M1	0.11	0.74
M2	0.35	0.64
M3	0	0.87
M4	0.15	0.82
M5	0.15	0.85
M6	0.36	0.63
M7	0.16	0.74
M9	0.27	0.73
M12	0.09	0.87
M13	0.07	0.81
M14	0	1
M19	0.41	0.6
M21	0.28	0.72
M23	0	0.99
M24	0.18	0.81
M26	0.2	0.79
M28	0.21	0.64
M29	0.05	0.91
M34	0.1	0.65
M35	0.11	0.85
M36	0.36	0.64
M37	0.13	0.65
M38	0.33	0.57
M39	0.28	0.71
M41	0.7	0.3
M42	0.15	0.8
M43	0.07	0.87
M44	0.25	0.75
M45	0.3	0.7
M49	0.11	0.87
M50	0.02	0.98
Mean	0.19	0.76

TABLE 1.4, Proportion of time in 5-category child VITAL

	<i>ang/agg+dis/co mpt</i>	<i>tense+whin e/cranky</i>	<i>anx + fear</i>	<i>sad/pain+emb/shame</i>	<i>int/aff+hap/exc it</i>
M1	0.11	0.50	0.18	0.06	0.14
M2	0.05	0.64	0.24	0.07	0.00
M3	0.00	0.64	0.18	0.05	0.13
M4	0.02	0.69	0.07	0.18	0.03
M5	0.08	0.23	0.37	0.33	0.00
M6	0.04	0.53	0.23	0.19	0.01
M7	0.16	0.44	0.11	0.19	0.10
M9	0.08	0.58	0.12	0.21	0.00
M12	0.03	0.55	0.25	0.13	0.03
M13	0.05	0.05	0.58	0.19	0.12
M14	0.00	0.13	0.83	0.04	0.00
M19	0.39	0.29	0.18	0.14	0.00
M21	0.05	0.56	0.03	0.36	0.00
M23	0.00	0.39	0.04	0.57	0.00
M24	0.06	0.61	0.22	0.10	0.01
M26	0.16	0.42	0.29	0.12	0.01
M28	0.00	0.46	0.03	0.36	0.15
M29	0.00	0.35	0.28	0.32	0.05
M34	0.07	0.27	0.26	0.14	0.25
M35	0.05	0.34	0.40	0.16	0.04
M36	0.04	0.56	0.18	0.22	0.00
M37	0.11	0.49	0.09	0.08	0.23
M38	0.29	0.46	0.08	0.08	0.10
M39	0.05	0.43	0.31	0.20	0.01
M41	0.47	0.26	0.24	0.02	0.00
M42	0.01	0.56	0.29	0.09	0.04
M43	0.00	0.34	0.45	0.15	0.07
M44	0.10	0.35	0.52	0.02	0.00
M45	0.00	0.63	0.30	0.07	0.00
M49	0.10	0.68	0.14	0.06	0.00
M50	0.00	0.71	0.24	0.05	0.00
Mean	0.08	0.46	0.25	0.16	0.05

TABLE 2.1, Proportion of observed time in original interviewer VITAL categories

	Other		Connector		Rapport	
	3	4	5	6	7	8
M1	0.06	0.04	0.21	0.01	0.22	0
M2	0	0	0.23	0	0.43	0.01
M4	0.01	0	0.23	0	0.34	0
M5	0.03	0	0.51	0	0.15	0
M6	0	0	0.36	0	0.27	0
M7	0.01	0	0.19	0	0.4	0
M9	0.01	0	0.25	0	0.18	0
M12	0.04	0	0.18	0	0.22	0.03
M13	0.04	0	0.09	0	0.16	0
M14	0	0	0.09	0	0.39	0
M19	0.11	0.02	0.18	0	0.04	0
M21	0.08	0	0.24	0	0.09	0
M23	0.06	0.01	0.22	0	0.42	0
M24	0.02	0	0.24	0	0.28	0
M26	0.03	0	0.45	0	0.19	0
M28	0	0	0.18	0	0.73	0
M29	0.02	0	0.17	0	0.34	0
M34	0.04	0	0.25	0	0.38	0
M35	0.04	0	0.17	0	0.42	0
M36	0.02	0	0.24	0	0.41	0
M37	0	0	0.28	0	0.37	0
M38	0.02	0	0.26	0	0.26	0
M39	0.1	0	0.18	0	0.39	0
M41	0.05	0	0.21	0	0.25	0
M42	0.28	0.27	0	0	0	0
M43	0	0	0.1	0	0.66	0
M44	0.04	0	0.11	0	0.44	0
M45	0	0	0.21	0	0.31	0
M49	0.06	0	0.42	0	0.08	0
M50	0	0	0.34	0	0.15	0.02
Mean	0.04	0.01	0.23	0	0.3	0

TABLE 2.1, Continued

General Abuse Questions

Open		Specific		Closed		Leading		Recap	
9	10	11	12	13	14	15	16	17	18
0	0	0	0	0.01	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0.02	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0.01	0	0.01	0	0	0	0	0
0	0	0.01	0	0.03	0	0.01	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0.11	0.1	0.04	0	0.01	0	0.04	0
0.07	0	0.11	0	0	0	0.02	0	0.02	0
0	0	0.04	0.1	0	0	0	0	0.04	0
0	0	0.03	0	0	0	0.01	0	0.01	0
0	0	0	0	0	0	0	0	0	0
0	0	0.02	0	0.02	0	0	0	0.01	0
0	0	0.01	0	0.01	0	0	0	0	0
0	0	0.03	0	0	0	0	0	0.02	0
0	0	0	0	0.05	0	0	0	0	0
0	0	0.01	0	0	0	0	0	0	0
0	0	0.04	0	0.06	0	0.04	0	0.04	0
0	0	0.06	0	0.06	0	0.03	0	0	0
0	0	0	0	0.01	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0.01	0	0.03	0	0.01	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0.01	0	0.18	0.1
0	0	0	0	0.01	0	0	0	0	0
0	0	0.04	0	0.02	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0.02	0	0.01	0	0	0	0.01	0

TABLE 2.1, Continued

Specific Abuse Questns											
Open		Specific		Closed		Leading		Recap		Legal System	
19	20	21	22	23	24	25	26	27	28	29	30
0.03	0	0.04	0.1	0	0.01	0.03	0.01	0.03	0.01	0	0.01
0.01	0	0.1	0.05	0	0.03	0.02	0.01	0.05	0.03	0.01	
0	0	0.04	0.11	0	0.06	0.01	0.02	0.02	0.09	0	
0.01	0	0.06	0.07	0	0.08	0	0.02	0.03	0.06	0	
0	0	0.02	0.02	0.1	0.01	0.12	0.03	0.01	0	0	
0.01	0	0.04	0.11	0	0.06	0	0.04	0.02	0.02	0	
0.04	0	0.04	0.08	0.1	0.08	0.05	0.03	0.06	0.06	0	
0.01	0	0.07	0.03	0	0.07	0.01	0.01	0.02	0.02	0	
0.04	0	0.11	0.25	0	0.02	0	0	0.02	0.04	0	
0	0	0.17	0.09	0	0	0	0.04	0	0	0	
0.01	0	0.1	0.1	0.1	0.11	0.01	0.08	0.05	0.05	0	
0.02	0	0.12	0.05	0.1	0.08	0.04	0.03	0.07	0.03	0.01	0.01
0.01	0	0.05	0.02	0.1	0.01	0	0	0.02	0	0	
0.01	0	0.1	0.08	0.1	0.02	0.02	0.01	0.07	0.06	0	
0.06	0	0	0.1	0	0	0	0	0.04	0.02	0.01	
0	0	0	0	0.1	0	0	0	0	0	0	
0	0	0.16	0.08	0	0.07	0.03	0	0.04	0.03	0.01	
0.02	0	0.05	0.04	0	0.01	0.02	0	0.01	0.01	0	
0	0	0.08	0.01	0.1	0.01	0.01	0.01	0.01	0.03	0	
0.02	0	0.06	0.01	0.1	0.06	0.03	0	0.04	0.03	0	
0.01	0	0.05	0.1	0	0.02	0.01	0	0.07	0.05	0	
0	0	0.07	0.09	0.1	0.01	0.03	0.03	0.07	0	0.01	
0	0	0.13	0.05	0.1	0.02	0.04	0.02	0.01	0.01	0	
0	0	0.2	0.06	0.1	0.03	0.05	0.01	0.02	0.03	0	0.01
0.1	0	0.05	0.01	0	0	0					
0.03	0	0.05	0	0	0	0.12	0	0.02	0	0	
0	0	0.09	0.07	0	0.04	0.07	0.04	0.02	0	0	
0	0	0.11	0.04	0.1	0.03	0.01	0	0.13	0.01	0	
0.02	0	0.28	0	0	0	0.02	0.02	0	0.04	0	
0	0	0.14	0.1	0.1	0.05	0	0	0.03	0.03	0	
0.01	0	0.09	0.06	0.1	0.03	0.02	0.02	0.03	0.03	0	0.01

TABLE 2.2, Proportion of observed time in 4-category interviewer VITAL

	Connector	Rapport	General ?s	Personal ?s
M1	0.21	0.22	0.13	0.36
M2	0.23	0.44	0	0.33
M4	0.23	0.34	0.03	0.4
M5	0.51	0.15	0.03	0.35
M6	0.36	0.27	0.04	0.33
M7	0.19	0.4	0.08	0.33
M9	0.25	0.18	0.01	0.55
M12	0.18	0.25	0.33	0.24
M13	0.09	0.16	0.27	0.47
M14	0.09	0.39	0.17	0.35
M19	0.18	0.04	0.21	0.56
M21	0.24	0.09	0.08	0.59
M23	0.22	0.43	0.12	0.23
M24	0.24	0.28	0.05	0.43
M26	0.45	0.19	0.09	0.27
M28	0.18	0.73	0.05	0.05
M29	0.17	0.34	0.03	0.46
M34	0.25	0.38	0.2	0.18
M35	0.17	0.42	0.19	0.22
M36	0.24	0.41	0.03	0.33
M37	0.28	0.37	0	0.35
M38	0.26	0.26	0.07	0.4
M39	0.18	0.39	0.1	0.33
M41	0.21	0.25	0.05	0.48
M42	0	0	0.78	0.22
M43	0.1	0.66	0.01	0.23
M44	0.11	0.44	0.11	0.33
M45	0.21	0.31	0.01	0.47
M49	0.42	0.08	0.06	0.42
M50	0.34	0.17	0	0.49
Mean	0.23	0.3	0.11	0.36
SD	0.11	0.16	0.15	0.12

TABLE 2.3, Proportion of observed time in interviewer question-type categories

	Open	Specific	Closed	Leading	Recap
M1	0.06	0.15	0.07	0.04	0.04
M2	0.01	0.15	0.07	0.03	0.08
M4	0.01	0.15	0.12	0.03	0.12
M5	0.01	0.12	0.11	0.02	0.08
M6	0.01	0.05	0.12	0.16	0.02
M7	0.01	0.18	0.1	0.06	0.04
M9	0.04	0.13	0.18	0.07	0.13
M12	0.01	0.27	0.14	0.03	0.1
M13	0.11	0.49	0.02	0.02	0.07
M14	0	0.39	0.04	0.04	0.04
M19	0.01	0.25	0.17	0.1	0.11
M21	0.02	0.17	0.22	0.08	0.09
M23	0.01	0.09	0.14	0	0.04
M24	0.01	0.19	0.09	0.03	0.13
M26	0.09	0.14	0	0	0.09
M28	0	0	0.1	0	0
M29	0	0.25	0.12	0.03	0.06
M34	0.02	0.12	0.08	0.06	0.06
M35	0	0.16	0.14	0.05	0.04
M36	0.02	0.07	0.14	0.03	0.07
M37	0.02	0.15	0.05	0.01	0.12
M38	0.01	0.2	0.09	0.06	0.07
M39	0	0.18	0.08	0.06	0.01
M41	0	0.27	0.1	0.05	0.05
M42	0.13	0.06	0.03	0.01	0.23
M43	0.03	0.05	0.03	0.12	0.02
M44	0	0.2	0.07	0.11	0.02
M45	0	0.15	0.17	0.03	0.14
M49	0.02	0.28	0.04	0.04	0.04
M50	0	0.24	0.17	0	0.07
Mean	0.02	0.18	0.1	0.05	0.07
SD	0.03	0.1	0.05	0.04	0.05

TABLE 2.4, Proportion of observed time concerning the alleged abuser

M1	0.25
M2	0.13
M4	0.29
M5	0.22
M6	0.07
M7	0.25
M9	0.26
M12	0.27
M13	0.33
M14	0.22
M19	0.41
M21	0.20
M23	0.05
M24	0.18
M26	0.15
M28	0.00
M29	0.18
M34	0.06
M35	0.06
M36	0.10
M37	0.19
M38	0.14
M39	0.09
M41	0.14
M42	0.35
M43	0.00
M44	0.16
M45	0.10
M49	0.06
M50	0.21
Mean	0.17

TABLE 3.1, Child gender by alleged abuse type

	Child's Sex			
	Male		Female	
	abuse type - sexual or physical		abuse type - sexual or physical	
	Count	%	Count	%
sexual abuse	4	44.4%	16	76.2%
physical abuse	5	55.6%	5	23.8%

TABLE 3.2, Child age groups and gender by case outcome and alleged abuse type

				Case Outcome				Total
				Unfounded		Founded		
				abuse type - sexual or physical		abuse type - sexual or physical		
				sexual abuse	physical abuse	sexual abuse	physical abuse	
Child's Sex	Male	child's age by category 1	1	1		1	3	
		2	1	1	1	1	4	
		3			1	1	2	
	Female	child's age by category 1	1	3	1	1	2	7
		2	2	4	2	3		9
		3	1	1	4		5	
Total			10	5	10	5	30	

TABLE 3.3, Case outcome by child age groups

		Case Outcome		Total
		Unfounded	Founded	
child's age by category	1.00	6	4	10
	2.00	8	5	13
	3.00	1	6	7
Total		15	15	30

TABLE 4.1, Global interrater reliability

Overall Chance-Corrected Agreement (Cohen's kappa) between observers/raters

	Child VITAL				Interviewer VITAL		
	rater1x2	rater1x3	rater2x3		rater1x2	rater1x3	rater2x3
M2	0.613	0.611	0.633		0.781	0.909	0.802
M3	0.715	0.75	0.677		0.848	0.94	0.789
M4	0.518	0.68	0.505		0.832	0.924	0.803
M5	0.328	0.461	0.328		0.824	0.831	0.809
M6	0.504	0.577	0.493		0.714	0.787	0.732
M7	0.448	0.569	0.571		0.807	0.903	0.737
M9	0.43	0.554	0.505		0.744	0.826	0.782
M12	0.528	0.604	0.54		0.678	0.797	0.703
M13	0.44	0.565	0.407		0.823	0.879	0.803
M14	0.381	0.25	0.392		0.742	0.948	0.691
M19	0.329	0.513	0.506		0.723	0.816	0.782
M21	0.537	0.585	0.572		0.775	0.812	0.812
M23	0.582	0.646	0.591		0.878	0.886	0.901
M24	0.555	0.621	0.518		0.777	0.873	0.804
M26	0.58	0.643	0.537		0.873	0.905	0.873
M28	0.431	0.498	0.542		0.878	1	0.878
M29	0.28	0.298	0.242		0.846	0.89	0.825
M34	0.498	0.667	0.498		0.829	0.9	0.871
M35	0.23	0.503	0.41		0.823	0.941	0.835
M36	0.547	0.572	0.496		0.806	0.922	0.78
M37	0.324	0.39	0.218		0.713	0.856	0.742
M38	0.742	0.77	0.797		0.826	0.851	0.852
M39	0.568	0.66	0.475		0.81	0.831	0.865
M41	0.397	0.625	0.401		0.852	0.914	0.852
M42	0.849	0.879	0.969		1	0.972	0.972
M43	0.704	0.644	0.621		0.722	0.812	0.844
M44	0.611	0.634	0.634		0.868	0.956	0.896
M45	0.497	0.472	0.401		0.764	0.899	0.746
M49	0.82	0.83	0.866		0.9	0.899	0.877
M50	0.189	0.556	0.382		0.779	0.739	0.752
Mean	0.5058	0.588	0.5242		0.80783	0.8806	0.814
Overall Mean		0.539				0.834	

TABLE 4.2, Specific interrater reliability for child VITAL categories

(null sets (0) removed from analysis)

(where proportion of time spent in emotion category was at least 1%)

	2	3	5	6	7	8	9	10	11	12	14	15	16	17
	neutr al	ang er	disg ust	cont emp t	sadn ess	pain /hur t	Tens e	anxie ty	fear	em barr ass	whin e	cran ky	with draw	inter est
M2 1x2				0.54	0.39	0.7	0.918	0.48	0	0.7	0.48	0.76	0.66	
M2 1x3				1.2	0.65	0	0.706	0.4		2	0.59	0.78	1	
M2 2x3				0.7	1	2	0.643	0.51	0.8	0.8	0.37	0.86	1	
M3 1x2	0						0.661	0.56					1	0.73
M3 1x3	1						0.722	0.73					1	0.73
M3 2x3	0						0.661	0.63					1	0.79
M4 1x2			0		0.32		0.611	0.49		0.1	0.52		0	0.31
M4 1x3			0.5		0.74		0.683	0.68		0.9	0.56		0	0.49
M4 2x3			0		0.49		0.592	0.63		0.5	0.38		-0	0.49
M5 1x2		0.3	-0	0.37	0.24	0.6	0.453	0.37	0.8	0.1	-0	0.08	-0	
M5 1x3		0.8	0	0.57	0.45	0.6	0.545	0.46	0	-0	-0	0.1	0	
M5 2x3		0.4	0	0.46	0.29	0.8	0.444	0.19	0	-0	-0	0.23	0	
M6 1x2			0.8	0.43	0.6		0.435	0.48		0.4	0.46	0.66	0.15	0.33
M6 1x3			0.7	0.24	0.83		0.538	0.61		0.2	0.45	0.61	0.58	0.44
M6 2x3			0.8	0.57	0.6		0.401	0.57		0.4	0.26	0.6	0.11	0.35
M7 1x2	1	0.7	0	0.22	-0		0.413	0.73		0.1				0.73
m7 1x3	1	0.6	0	0.22	0.33		0.56	0.75		0.4				0.81
M7 2x3	1	0.7	0	0.49	0.24		0.518	0.6		0.4				0.93
M9 1x2			1	0.39			0.4	0.33		0.5	0.39	0.01	0.72	
M9 1x3			1	0.64			0.524	0.56		0.5	0.53	0.29	0.65	
M9 2x3			1	0.43			0.516	0.3		0.6	0.47	0.27	0.73	
M12 1x2	0.48	0.9		1	0.54		0.47	0.51	0.66	0.4	0.79	0.89		-0
M12 1x3	0.43	1		1	0.66		0.577	0.63	1	0.2	0.57	0.72		0.56
M12 2x3	0.29	0.9		1	0.59		0.471	0.52	0.66	0.6	0.68	0.83		0.5
M13 1x2				0	0.88		0.126	0.59			-0	0	0.64	0.66
M13 1x3				-0	0.73		0.65	0.71			-0	0	0.78	0.3
M13 2x3				0	0.85		0.151	0.46			-0	1	0.54	0.38
M14 1x2							0.619	0.26	0			0	1	
M14 1x3							0.273	0.1	0.65			0	0.65	
M14 2x3							0.273	0.19	0			0.65	0.65	
M19 1x2		0.7	0.3	0.47	0.41		0.353	0.33		0.3	0	0.32		
m19 1x3		0.7	0.6	0.62	0.52		0.457	0.45		0.6	0	0.25		
M19 2x3		1	0.5	0.47	0.54		0.526	0.41		0.6	0.66	0.73		
M21 1x2	-0	0.2		0.47	0.6	0.8	0.559	0.32	-0	0	0.31	0.52	0.33	
m21 1x3	0.5	0.3		0.52	0.62	0.8	0.616	0.47	0	-0	0.38	0.63	0.19	
m21 2x3	-0	-0		0.53	0.66	0.9	0.579	0.43	0	0	0.34	0.47	0.25	

TABLE 4.2, Continued

	2	3	5	6	7	8	9	10	11	12	14	15	16	17
	neutr al	ang er	disg ust	cont emp t	sadn ess	pain /hur t	tense	anxie ty	fear	em barr ass	whin e	cran ky	with draw	inter est
M23 1x2					0.45	0.9	0.521	0.26		0.3	1		0.44	
m23 1x3					0.59	0.9	0.601	0.58		0.5	0.5		0.23	
m23 2x3					0.34	0.9	0.55	0.24		0.5	0.5		0.49	
M24 1x2			-0	0.45	0.46		0.561	0.47		0.1	0.5	0.39	0.25	-0
m24 1x3			-0	0.67	0.58		0.637	0.53		0.1	0.42	0.5	0.49	0
m24 2x3			-0	0.43	0.09		0.558	0.36		0.1	0.47	0.41	0.33	0
M26 1x2	0			0.59	0.59		0.464	0.38		0.3	0.32	0	0.94	-0
m26 1x3	-0.01			0.58	0.2		0.676	0.57		0.3	0.29	0.49	0.79	-0
m26 2x3				0.69	0.27		0.436	0.28		0.6	0.24	0.43	0.74	0
M28 1x2							0.294	0.66			0.36		0.73	-0.1
m28 1x3							0.333	1			0.63		0.65	0
m28 2x3							0.645	0.66			0.15		0.71	0
M29 1x2	0			0	0.4	0	0.143	0.16		0	0.32	0.66	0.52	0.37
m29 1x3	0			0	0.34	0	0.227	0.29		0	0.27	0	0.46	-0
m29 2x3	0.39			-0	0.17	1	0.175	0.27		-0	-0	0	0.6	-0
M34 1x2	0			0.55	0.55		0.547	0.61		0.7	0.66			0.47
m34 1x3	1			0.55	0.77		0.647	0.57		0.5	0.66			0.86
m34 2x3	0			0.74	0.63		0.511	0.61		0.4	0.39			0.36
M35 1x2				0.23	-0		0.222	0.2	0	-0	0.13	0	0.41	0
m35 1x3				0	-0		0.508	0.42	0.32	0.5	0.65	0	0.8	0.66
m35 2x3				0	-0		0.606	0.48	0	0.3	0.27	0.42	0.48	0
M36 1x2				0.66	0.76	0.7	0.604	0.36			0.49	0.37	0.55	
m36 1x3				0.28	0.38	0.6	0.681	0.5			0.65	0.41	0.61	
m36 2x3				0.35	0.36	-0	0.689	0.3			0.48	0.24	0.61	
M37 1x2		-0		0.15	0.15		0.352	0		0.5		0		0.55
m37 1x3		1		0.62	0.22		0.532	0.52		0.7		0		-0.1
m37 2x3		-0		0.16	0.66		0.316	0		0.8		-0		0.42
M38 1x2	0.55	0.5	0.9	0.84	0.56		0.679	0.43		0.7	-0	-0	1	0.6
m38 1x3	0.59	0.5	0.9	0.76	0.56		0.816	0.44		0.3	-0	0	1	0.72
m38 2x3	0.7	1	1	0.84	0.74		0.706	0.49		0.4	-0	0	1	0.83
M39 1x2		0.3		0.13	0.4	1	0.572	0.67	1		0.53	0.28	0.75	
M39 1x3		0.7		-0	0.77	1	0.609	0.87	1		0.64	0.45	0.89	
m39 2x3		0.7		-0	0.42	1	0.402	0.47	1		0.47	0.35	0.64	
M41 1x2		0.3		0.41	0.66		0.078	0.37		0	0.18	0.43	0	
m41 1x3		0.6		0.79	1		-0.04	0.54		-0	-0	0.66	0	
m41 2x3		0.6		0.35	0.66		0.301	0.31		0	0.27	0.43	-0	
M42 1x2	0.82			1	0.85		0.809	0.88	0.64		1	1		
m42 1x3	1			1	0.85		0.809	0.88	0.64		1	1		
m42 2x3	0.82			1	1		1	1	1		1	1		

TABLE 4.2, Continued

	2	3	5	6	7	8	9	10	11	12	14	15	16	17
	neutr al	ang er	disg ust	cont emp t	sadn ess	pain /hur t	tense	anxie ty	fear	em barr ass	whin e	cran ky	with draw	inter est
M43 1x2							0.576	0.6	0.66	0.5	0.65		0.6	0.71
M43 1x3							0.607	0.44	0.79	0.8	0.3		0.56	0.5
m43 2x3							0.448	0.44	0.85	0.8	0.34		0.66	0.46
M44 1x2				-0			0.254	0.49			0.47	0	-0	
m44 1x3				0.32			0.373	0.46			0.55	-0	1	
m44 2x3				0.39			0.399	0.52			0.33	0	-0	
M45 1x2					0.49		0.443	0.52		0.5	0.53	0.66		
m45 1x3					0.32		0.607	0.07		0.4	0.64	0.65		
m45 2x3					0.74		0.462	0.17		0.4	0.46	0.39		
M49 1x2				0.57			0.785	0.33				-0		
m49 1x3				0.86			0.821	0.33				-0		
m49 2x3				0.66			0.867	0.23				0.67		
M50 1x2					0		0.181	0.11		0.5	0	0.56		
m50 1x3					0.66		0.611	0.54		0.4	0	0.66		
m50 2x3					0		0.341	0.36		0.7	0.38	0.65		
Mean	0.44	0.6	0.4	0.47	0.5	0.7	0.513	0.46	0.48	0.4	0.39	0.38	0.53	0.37

TABLE 4.3, Specific interrater reliability for interviewer VITAL categories

(null sets (0) removed from analyses)

(where proportion of time spent in question/statement category was at least .01%)

	General ?s						Personal ?s									Legal
	Non e	Con nect or	Rapp ort	Spe cific	Clo sed	Reca p	Ope n	Spec ific	A	Close d	A	Leadi ng	A	Rec ap	A	
	1	5	7	11	13	17	19	21	22	23	24	25	26	27	28	30
M2	0.7	0.79	0.84				1	0.82	0.88	0.51	0.24	-0.01	-0.01	0.51	-.01	
M3	1	1	1				1	0.48	0	0		0		0.66	0.66	
M4	1	0.97	0.83	0	0.38			0.71	0.94	0.56	0.79	0.66	1	0.66	0.88	
M5	1	0.96	0.87				1	0.85	0.53	0.14	0.26	0	-0.01	-.02	0.66	
M6	0.92	0.87	0.78	0.8	0.8			0.49	0.59	0.48	0.28	0.44	0.49	0	0	
M7	1	0.96	0.98	1	0.79		1	0.71	0.73	0.24	0.22		0	0.49	0.66	
M9	0.97	0.93	0.95				0.87	0.56	0.65	0.59	0.54	0.56	0.21	0.66	1.51	
M12	0.92	0.85	0.83	0.65	0.72	0.7	0	0.02	0.72	0.27	0.52	1	0.8	0.49		
M13	1	0.78	0.94	0.91		1	0.65	0.7	0.69		1			0.66		
M14	1	1	1	0.47		1		0.7	0.65	0.47			0			
M19	1	0.85	0.7	0.49	0	-0	1	0.9	0.71	0.65	0.67	-0.01	0.43	0.62	0.35	
M21	0.98	0.94	0.97				0.49	0.75	0.67	0.62	0.6	0.54	0.23	0.61	0.29	0.86
M23	0.95	0.92	0.92	0.8	0.91	0.89	0.4	0.78	0.66	0.88	0.4	1		0.93	0	
M24	0.99	0.91	0.84	0.66	0.8		0.5	0.8	0.72	0.4	0.34	0.66	0.66	0.73	0.2	
M26	1	0.96	0.88	0.85		1	0.76	0	0.58	0	0			1	1	
M28		1	1		1					1						
M29	1	0.97	1	1				0.83	0.94	0.71	0.76	0.49		0.55	0	
M34	1	0.9	0.93	1	1	0.65	0.79	0.48	0.49	0.66	-0	0.66	0	-.01	0	
M35	1	0.97	0.97	0.43	0.7			0.52	0	0.68	1	1	0	0.66	1	
M36	1	0.9	0.9		0		0.8	0.94	0	0.69	0.6	0.33	0	0.48	0.39	
M37	1	0.92	0.93				1	-0.03	0.43	0.26	-0	1		0.75	0.39	
M38	1	0.87	0.86	0.8	1			0.79	0.76	0.65	0.66	0.43	0.49	0.83		
M39	1	0.81	0.91					0.78	0.79	0.56	0.39	0.55	0.8	0.66	0	
M41	1	0.98	0.98					0.91	0.61	0.6	0	0.18	1	0.32	0	1
M42	1	1	1				1	1	1	1	1	1	1	1		
M43	1	0.51	0.8		0.66		0	0.48		0.56	0	0.52		0.8		
M44	0.98	0.76	0.88	0.66	-.01			0.74	0.8	0	0.8	0.49	1	1		
M45	1	0.84	0.87					0.75	0.79	0.57	0.66			0.78	1	
M49	1	0.97	0.89				1	0.83	0	0.44		-0	0	0	0.67	0
M50	0.92	0.95	0.9					0.85	0.59	0.55	0.85			0.22	0.66	
Mean	0.98	0.9	0.9	0.7	0.63	0.75	0.74	0.66	0.6	0.51	0.48	0.5	0.4	0.57	0.47	0.62

TABLE 5.1, Criterion validity results

<u>Hypotheses</u>	<u>Statistical Significance</u>	<u>Trend</u>	<u>Direction</u>
OVERALL:			
High uncomfortable	Yes*		45% of sample
Low positive	Yes*		5% of sample
Most specific and closed ?s	Yes*		17% & 10%
Least open and leading ?s	Yes*		2% & 5%
More specific, closed, and leading questions with younger children	No	Yes	Fewer closed ?s w/younger children
Fewer open ?s w/ young children	No	No	
Rapport longer with young children	Yes		Longest for youngest children
	<i>*signif. not tested</i>		
FOUNDED v. UNFOUNDED:			
Little to no emotion difference	Yes (null)		No differences
Slightly higher on anger for founded cases	No	Yes	Higher in founded
Slightly higher on scared for founded cases	No	Yes	Lower in founded
Slightly higher on depressed for founded cases	No	Yes	Higher in founded
Slightly higher on positive for unfounded cases	No	No	Higher in founded
No diff. on uncomfortable	Yes (null)		No difference
Similar times in rapport	Yes (null)		Similar times
More time in personal ?s w/founded cases	Yes		More in founded
More time in alleged abuser ?s w/founded cases	Yes		More in founded
More anger in founded physical cases	Yes		More in founded physical than unfound physical & found/unfounded sexual
MARYLAND v. CRAIG:			
High distress w/personal questions and founded cases	Yes		Higher externalizing
High distress w/alleged abuser questions and founded cases	No	No	Higher internalizing
Low distress w/general questions and founded	No	Yes	Lower externalizing when > general ?s

TABLE 5.2, Risk and protective factors results

Hypotheses	Significance	Trend	Direction
Gender little effect	Yes (null)		
Boys higher externalizing		Yes	Higher for girls
Girls higher internalizing		No	Oldest girls higher internalize than oldest boys
Age little effect	Yes (null)		
[Not hypothesized: Older children more angry]		Yes	Oldest group most angry; [oldest boys > oldest girls]
[NH: Younger more positive]		Yes	Older boys > older girls
Greater frequency little effect	Yes (null)		[more externalize & less internalize]
Longer duration little effect	Yes (null)	Yes	
Longer worse externalize		Yes	Highest at longest duration
Longer worse internalize		Yes	Highest internalize, highest scared, and lowest uncomfortable at 7-31 days
Threat/Use of Force little effect	Yes (null)	Yes	Lower externalizing
Invasive sexual abuse little effect	Yes (null)		
Closer relationship little effect	Yes (null)		
More investigative interviews higher child VITAL	No	Yes	Higher internalizing; lower externalizing
More rapport, less uncomfortable	No	No	
Family dysfunction/unstable	No	No	
Social support (Risk Factor)	Yes		Higher internalizing
Social support (Protective factor)	Yes		Lower internalize, & higher externalize & uncomfortable
Presence of non-offending parental support, lower VITAL ratings	Yes	Yes	<i>Trend:</i> Lower internalize & depression & higher anger; <i>Sig:</i> higher uncomfortable
FOUNDED v. UNFOUNDED:			
Demographic (gender &/or age)	Yes		Oldest more in founded
Frequency, Duration & Relation	No		
Threat or Use of Force	Yes		More threat/force in found
Invasive sexual abuse	Yes		More invasive in founded
Systemic risk factors	No		
Protective factors	No		

REFERENCES

- Achenbach, T. M. (1986). Child behavior checklist – direct observation form. Burlington, VT: University of Vermont.
- Ackerman, P. T., Newton, J. E. O., McPherson, W. B., Jones, J. G., & Dykman, R. A. (1998). Prevalence of post traumatic stress disorder and other psychiatric diagnoses in three groups of abused children (sexual, physical, and both). Child Abuse & Neglect, *22*, 759-774.
- Allen, S. N. (1994). Psychological assessment of post-traumatic stress disorder: Psychometrics, current trends, and future directions. Psychiatric Clinics of North America, *17* (2), 327-349.
- Amaya-Jackson, L., & March, J. S. (1995). Posttraumatic stress disorder. In J. S. March (Ed.), Anxiety disorders in children and adolescents (pp. 276-300). New York: Guilford.
- Ambrosini, P. J., Metz, C., Prabucki, K., & Lee, J. (1989). Videotape reliability of the third revised edition of the K-SADS. Journal of the American Academy of Child and Adolescent Psychiatry, *28* (5), 723-728.
- American Professional Society on the Abuse of Children (1990). Guidelines for psychosocial evaluation of suspected sexual abuse in young children.
- American Professional Society on the Abuse of Children (1998). Practice guidelines: Psychosocial evaluation of suspected sexual abuse in children (2nd ed.).
- American Academy of Child and Adolescent Psychiatry (1997). Practice parameters for the forensic evaluation of children and adolescents who may have been physically or sexually abused. Journal of the American Academy of Child and Adolescent Psychiatry, *36*(10 Suppl.), 37s-56s.
- Anastasi, A. (1988). Psychological testing (6th Ed.). New York: Macmillan.

- Arata, C. M. (1998). To tell or not to tell: Current functioning of child sexual abuse survivors who disclosed their victimization. Child Maltreatment, 3 (1), 63-71.
- Arroyo, W., & Eth, S. (1995). Assessment following violence-witnessing trauma. In E. Peled, P. G. Jaffe, & J. L. Edleson (Eds.), Ending the cycle of violence: Community responses to children of battered women (pp. 27-42). Thousand Oaks, CA: Sage.
- Bagley, C., & Thurston, W. E. (1996). Understanding and preventing child sexual abuse: Critical summaries of 500 key studies: Vol. 1. Children: Assessment, social work and clinical issues, and prevention education. Hants, England: Arena.
- Bakeman, R., & Gottman, J. M. (1997). Observing interaction (2nd ed.). Cambridge, United Kingdom: Cambridge University Press.
- Bala, N. (in press). Child witnesses in the Canadian criminal courts: Recognising their capacities and needs. Psychology, Public Policy, and Law.
- Barker-Collo, S. L., Melnyk, W. T., & McDonald-Miszczak, L. (2000). A cognitive-behavioral model of post-traumatic stress for sexually abused females. Journal of Interpersonal Violence, 15 (4), 375-392.
- Becker, J. V., Alpert, J. L., BigFoot, D. S., Bonner, B. L., Geddie, L. F., Henggeler, S. W., Kaufman, K. L., & Walker, C. E. (1995). Empirical research on child abuse treatment: Report by the child abuse and neglect treatment working group, American Psychological Association, Journal of Clinical Child Psychology, 24 (Suppl.), 23-46.
- Berliner, L., & Conte, J. R. (1995). The effects of disclosure and intervention on sexually abused children. Child Abuse & Neglect, 19 (3), 371-384.
- Bray, J. H., & Maxwell, S. E. (1985). Multivariate analysis of variance. Newbury Park, CA: Sage Publications.
- Briere, J. N. (1992). Child abuse trauma: Theory and treatment of the lasting effects. J. R. Conte (Series ed.), Interpersonal Violence: The practice series. Newbury Park, CA: Sage.

Briere, J. (1996). Trauma symptom checklist for children (TSCC): Professional manual. Psychological Assessment Resources.

Briere, J., Elliott, D. M., Harris, K., & Cotman, D. (1995). Trauma symptom inventory: Psychometrics and association with childhood and adult victimization in clinical samples. Journal of Interpersonal Violence, 10 (4), 387-401.

Brown, J., Cohen, P., Johnson, J. G., & Salzinger, S. (1998). A longitudinal analysis of risk factors for child maltreatment: Findings of a 17-year prospective study of officially recorded and self-reported child abuse and neglect. Child Abuse & Neglect, 22 (11), 1065-1078.

Bruck, M., Ceci, S. J., & Melnyk, L. (1997). External and internal sources of variation in the creation of false reports in children. Learning and Individual Differences, 9 (4), 289-316.

Bugental, D. B., Blue, J., Cortez, V., Fleck, K., & Rodriguez, A. (1992). Influences of witnessed affect on information processing in children. Child Development, 63, 774-786.

Burleson, B. R., Albrecht, T. L., Goldsmith, D. J., & Sarason, I. G., (1994). Introduction: The communication of social support. In B. R. Burleson, T. L. Albrecht, & I. G. Sarason (Eds.), Communication of social support: Messages, interactions, relationships, and community (pp. xi-xxx). Thousand Oaks, CA: Sage.

Carlson, E. B. (1997). Contemporary issues and innovations in the assessment of adult survivors of traumatic abuse, The APSAC Advisor, 10 (3), 12-15.

Carlson, E. B., Furby, L., Armstrong, J., & Shales, J. (1997). A conceptual framework for long-term psychological effects of traumatic childhood abuse. Child Maltreatment, 2, 272-295.

Carter, C. A., Bottoms, B. L., & Levine, M. (1996). Linguistic and socioemotional influences on the accuracy of children's reports. Law and Human Behavior, 20, 335-358.

Cecchetti-Whaley (1992), cited in Walton, E. (1994). The confrontation clause and the child victim of sexual abuse, Child & Adolescent Social Work Journal, 11(3), 2009.

Ceci, S. J., & Bruck, M. (1993). The suggestibility of the child witness: A historical review and synthesis. Psychological Bulletin, 113, 403-439.

Chadwick, D. L. (1999, Winter). A national call to action: Approaching the elimination of child maltreatment. APSAC Advisor, 12 (4), 2-5.

Chaffin, M., Wherry, J. N., & Dykman, R. (1997). School age children's coping with sexual abuse: Abuse stresses and symptoms associated with four coping strategies. Child Abuse & Neglect, 21 (2), 227-240.

Chandler, S. M. (1982). Knowns and unknowns in sexual abuse of children. Journal of Social Work and Human Sexuality, 1 (1/2), 51-58.

Chandy, J. M., Blum, R. W., & Resnick, M. D. (1996). Female adolescents with a history of sexual abuse: Risk outcome and protective factors. Journal of Interpersonal Violence, 11 (4), 503-518.

Children's Institute International & United Way, Inc. (1996). Response Child Sexual Abuse: The Clinical Interview [Videotape recording].

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. Psychological Bulletin, 98 (2), 310-357.

Coolbaugh, K., & Hansel, C. J. (2000, March). The comprehensive strategy: Lessons learned from the pilot sites. OJJDP Juvenile Justice Bulletin, NCJ 178258, Washington, DC.

Coster, W., Tickle-Degnen, L., & Armenta, L. (1995). Therapist-child interaction during sensory integration treatment: Development and testing of a research tool. The Occupational Therapy Journal of Research, 15 (1), 17-35.

Courtois, C. A. (1979). The incest experience and its aftermath. Victimology: An International Journal, 4, 337-347.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281-302.

Curle, C. E., & Williams, C. (1996). Post-traumatic stress reactions in children: Gender differences in the incidence of trauma reactions at two years and examination of factors influencing adjustment. British Journal of Clinical Psychology, 35 (2), 297-309.

Davenport, C., Browne, K., & Palmer, R. (1994). Opinions on the traumatizing effects of child sexual abuse: Evidence for consensus. Child Abuse & Neglect, 18 (9), 725-738.

Davies, G., & Noon, E. (1993). Video links: Their impact on child witness trials. Issues in Criminological & Legal Psychology, 20, 22-26.

Davies, E., & Seymour, F. W. (1997). Child witnesses in the criminal courts: Furthering New Zealand's commitment to the United Nations convention on the rights of the child. Psychiatry, Psychology and Law, 4 (1), 13-24.

Davis, S. L., & Bottoms, B. L. (in press). Social support and children's eyewitness reports. Chapter to appear in M. L. Eisen, G. S. Goodman, & J. A. Quas (Eds.), Children's eyewitness testimony in the forensic interview.

Dorsey, B. L., Nelson, R. O., & Hayes, S. C. (1986). The effects of code complexity and of behavioral frequency on observer accuracy and interobserver agreement. Behavioral Assessment, 8, 349-363.

Dubowitz, H., Black, M., Harrington, D., & Verschoore, A. (1993). A follow-up study of behaviour problems associated with child sexual abuse. Child Abuse and Neglect, 17, 743-754.

Ebert, L., & Fairbank, J. A. (Fall, 1996). The trauma symptom inventory: Assessing the spectrum of symptoms associated with traumatic stress. AP-LS News, 16 (3), 21-23.

Eckman, P., & Friesen, W. V. (1978). Facial action coding system. Palo Alto, CA: Consulting Psychological Press, Inc.

Eth, S. (1988). The child victim as witness in sexual abuse proceedings. Psychiatry, *51*, 221-232.

Faller, K. (1988). Criteria for judging the credibility of children's statements about their sexual abuse. Child Welfare, *5*, 389-401.

Famularo, R., Fenton, T., Kinscherff, R., Ayoub, C., & Barnum, R. (1994). Maternal and child posttraumatic stress disorder in cases of child maltreatment. Child Abuse & Neglect, *18*, 27-36.

Farrell, A. D., Curran, J. P., Zwick, W. R., & Monti, P. M. (1983). Generalizability and discriminant validity of anxiety and social skills ratings in two populations. Behavioral Assessment, *6*, 1-14.

Feiring, C., Taska, L. S., & Lewis, M. (1998). Social support and children's and adolescents' adaptation to sexual abuse. Journal of Interpersonal Violence, *13* (2), 240-260.

Foy, D. W., Madvig, B. T., Pynoos, R. S., & Camilleri, A. J. (1996). Etiologic factors in the development of posttraumatic stress disorder in children and adolescents. Journal of School Psychology, *34* (2), 133-145.

Friedrich, W. (1995). Evaluation and treatment: The clinical use of the Child Sexual Behavior Inventory: Commonly asked questions. American Professional Society on the Abuse of Children (APSAC) Advisor, *8* (1), 1, 17-20.

Galatzer-Levy, R. M. (1994). Children, bad happenings, and meanings. Journal of the American Psychoanalytic Association, *42* (4), 997-1000.

Garvin, B. J., Kennedy, C. W., & Cissna, K. N. (1988). Reliability in category coding systems. Nursing Research, *37* (1), 52-55.

Goodman, G. S., Bottoms, B. L., Rudy, L., & Schwartz-Kenney, B. M. (1991). Children's testimony about a stressful event: Improving children's reports. Journal of Narrative and Life History, 1, 66-69.

Goodman, G. S., Levine, M., & Melton, G. B. (1992). The best evidence produces the best law. Law and Human Behavior, 16 (2), 244-251.

Goodman, G. S., Levine, M., Melton, G. B., & Ogden, D. W. (1991). Child witnesses and the confrontation clause: The American psychological association brief in Maryland v. Craig. Law and Human Behavior, 15 (1), 13-29.

Goodman, G. S., Quas, J. A., Bulkley, J., & Shapiro, C. (in press). Innovations for child witnesses: A national survey. Psychology, Public Policy, and Law.

Goodman, G., Taub, E., Jones, D., England, P., Port, L., Rudy, L. & Prado, L. (1992). Testifying in criminal court. Monographs of the Society for Research in Child Development, 57 (5), x.

Gottman, J. M., McCoy, K., Coan, J., & Collier, H. (1996). The specific affect coding system (SPAFF) for observing emotional communication in marital and family interaction. In Gottman, J. M. (Ed.), What Predicts Divorce? The Measures. Mahwah, NJ: Lawrence Erlbaum.

Grisso, T. (1996). Society's retributive response to juvenile violence: A developmental perspective. Law and Human Behavior, 20 (3), 229-247.

Grisso, T. (2000). Forensic assessment of violent juvenile offenders. Workshop presentation for American Academy of Forensic Psychology, Tucson, AZ.

Hall, S. R. (1995). A multidisciplinary manual for the use of televised alternative procedures (videotaped interviews and videotaped and closed-circuit testimony) with victims and non-victim child witnesses. Phoenix, AZ: Governor's Division for Children.

Hall, S. R. (1997). Evaluation of multidisciplinary investigations of child abuse: A literature review. Unpublished manuscript submitted to GCS Consulting, Tucson, AZ.

Hall, S. R. (1998). A VITAL Development: The Development and Interrater Reliability of the Videotaped Interview Trauma Assessment List. Unpublished master's thesis, University of Arizona, Tucson.

Hall, S. R. (1999). Protection of abused children as witnesses. Unpublished preliminary paper, University of Arizona, Tucson.

Heller, S. S., Larrieu, J. A., D'Imperio, R., & Boris, N. W. (1999). Research on resilience to child maltreatment: Empirical considerations. Child Abuse & Neglect, 23 (4), 321-338.

Henry, J. (1997). System intervention trauma to child sexual abuse victims following disclosure. Journal of Interpersonal Violence, 12 (4), 499-512.

Herbert, J., & Attridge, C. (1975). A guide for developers and users of observation systems and manuals. American Educational Research Journal, 12 (1), 1- 20.

Hops, H., Davis, B., & Longoria, N. (1995). Methodological issues in direct observation: Illustrations with the living in familial environments (LIFE) coding system. Journal of Clinical Child Psychology, 24 (2), 193-203.

Hsia, H. M., & Beyer, M. (2000). System change through state challenge activities: Approaches and products. OJJDP Juvenile Justice Bulletin, NCJ 177625, Washington, DC.

Iacobucci, D., & Wasserman, S. (1988). A general framework for the statistical analysis of sequential dyadic interaction data. Psychological Bulletin, 103(3), 379-390.

Imhoff, M. C., & Baker-Ward, L. (1999). Preschoolers' suggestibility: Effects of developmentally appropriate language and interviewer supportiveness. Journal of Applied Developmental Psychology, 20, 407-429.

Inderbitzen-Pisaruk, H., Shawchuck, C. R., & Hoier, T. S. (1992). Behavioral characteristics of child victims of sexual abuse: A comparison study. Journal of Clinical Child Psychology, 21 (1), 14-19.

Iverson, G. R., & Norpoth, H. (1987). Analysis of Variance (2nd ed.). Newbury Park, CA: Sage.

Jones, D. P. H. (1991). Professional and clinical challenges to protection of children, Child Abuse & Neglect, 15 (Suppl. 1), 57-66.

Jones, D., & McQuiston, M. (1986). Interviewing the sexually abused child (2nd ed.). Denver, CO: The C. Henry Kempe National Center for the Prevention and Treatment of Child Abuse and Neglect.

Jung, S., & Bavelas, J. B. (2000, March). Inductive analysis of police interviews with child witnesses. Poster session presented at the biannual conference of the American Psychology-Law Society, New Orleans, LA.

Kaminer, Y., Feinstein, C., Seifer, R., Stevens, L., & Barrett, R. P. (1990). An observationally based rating scale for affective symptomology in child psychiatry. The Journal of Nervous and Mental Disease, 178 (12), 750-754.

Kaplan, S. J., Pelcovitz, D., & Labruna, V. (1999). Child and adolescent abuse and neglect research: A review of the past 10 years. Part 1: Physical and emotional abuse and neglect. Journal of the American Academy of Child and Adolescent Psychiatry, 38 (10), 1214-1222.

Kelley, B., Thornberry, T. P., & Smith, C. A. (1997 August). In the wake of child maltreatment. OJJDP Juvenile Justice Bulletin. Washington, DC: U.S. Department of Justice.

Kendall-Tackett, K. A., Williams, L. M., & Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. Psychological Bulletin, 113 (1), 164-180.

Kolko, D. J. (1992). Characteristics of child victims of physical violence: Research findings and clinical implications. Journal of Interpersonal Violence, 7 (2), 244-276.

Kovera, M. B., & Borgida, E. (1996). Children on the witness stand: The use of expert testimony and other procedural innovations in U.S. child sexual abuse trials. In B. L. Bottoms & G. S. Goodman (Eds.). International perspectives on child abuse and children's testimony: Psychological research and law (pp. 201- 220). Thousand Oaks, CA: Sage.

Kratochwill, T. R. (1996). Posttraumatic stress disorder in children and adolescents: Commentary and recommendations. Journal of School Psychology, 34 (2), 185-188.

Kvalseth, T. O. (1989). Not on Cohen's Kappa. Psychological Reports, 65, 223-226.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33, 159-174.

Lipovsky, J. A. (1994). The impact of court on children: Research findings and practical recommendations. Journal of Interpersonal Violence, 9 (2), 238-257.

Lipovsky, J., & Stern, P. (1997). Preparing children for court: An interdisciplinary view. Child Maltreatment, 2 (2), 150-163.

Livingston, R., Lawson, L., & Jones, J. G. (1993). Predictors of self-reported psychopathology in children abused repeatedly by a parent. Journal of the American Academy of Child & Adolescent Psychiatry, 32, 948-953.

Maryland v. Craig, 497 U.S. 836 (1990).

McAuliff, B. D., & Kovera, M. B. (2000, March). Accommodating children in court: How do jurors view alternative testimonial procedures? Paper presented at the biannual conference of the American Psychology-Law Society Conference, New Orleans, LA.

- McClellan, J., Julie Adams, Donna Douglas, Chris McCurry & Mick Storck, (1995). Clinical characteristics related to severity of sexual abuse: A study of seriously mentally ill youth, Child Abuse & Neglect 19 (10), 1245-1246.
- McGee, R. A., Wolfe, D. A., & Wilson, S. K. (1997). Multiple maltreatment experiences and adolescent behavior problems: Adolescents' perspectives. Development and Psychopathology, 9, 131-149.
- McMillen, C., & Zuravin, S. (1997). Attributions of blame and responsibility for child sexual abuse and adult adjustment. Journal of Interpersonal Violence, 12 (1), 30- 48.
- McNally, R. J. (1996). Assessment of posttraumatic stress disorder in children and adolescents. Journal of School Psychology, 34 (2), 147-161.
- Melton, G. B. (1992). Children as partners for justice: Next steps for developmentalists, Monographs of the Society for Research in Child Development, 57 (5)(229), 153.
- Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C. (1997). Psychological evaluations for the courts: A handbook for mental health professionals and lawyers (2nd ed.). New York: Guilford.
- Melton, G. B., & Thompson, R. A., (1987). Getting out of a rut: Detours to less traveled paths in child-witness research. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), Children's eyewitness memory. NY: Springer-Verlag.
- Miller, T. W., & Veltkamp, L. J. (1995). Assessment of sexual abuse and trauma: Clinical measures. Child Psychiatry and Human Development, 26 (1), 3-10.
- Moffitt, T. (in press). Juvenile delinquency: Seed of a career in violent crime, just sowing wild oats – or both? Washington, DC: Federation of behavioral, psychological and cognitive sciences.
- Morrow, K. (1991). Attributions of female adolescent incest victims regarding their molestation. Child Abuse & Neglect, 15, 477-483.

- Mudford, O. C., Hogg, J., & Roberts, J. (1997). Interobserver agreement and disagreement in continuous recording exemplified by measurement of behavior state. American Journal on Mental Retardation, 102 (1), 54-66.
- Murray, K. (1995). Live television link: An evaluation of its use by child witnesses in Scottish criminal trials. The Scottish Office Central Research Unit.
- Myers, J. E. B. (1992). Steps toward forensically relevant research. Monographs of the Society for Research in Child Development, 57 (5)(229), 143-152.
- Nader, K., Stuber, M., & Pynoos, R. (1991). Posttraumatic stress reactions in preschool children with catastrophic illness: Assessment needs. Comprehensive Mental Health Care. Vol. 1 (3), 223-239.
- Nader, K. O. (1997). Assessing traumatic experiences in children. In J. P. Wilson & T. M. Keane (Eds.), Assessing psychological trauma and PTSD (pp. 291-348). New York: Guilford.
- Oberlander, L. B. (1995). Psycholegal issues in child sexual abuse evaluations: A survey of forensic mental health professionals. Child Abuse & Neglect, 19 (4), 475-490.
- O'Donohue, W., & Elliott, A. N. (1991). A model for the clinical assessment of the sexually abused child. Behavioral Assessment, 13, 325-339.
- Orsak, C. L. (1993). An empirical examination of child abuse assessment interviews: Factors related to interview credibility. Unpublished doctoral dissertation, Washington State University.
- Pence, D., & Wilson, C. (1994). Team investigation of child sexual abuse: The uneasy alliance. Thousand Oaks, CA: Sage.
- Peterson, C., & Biggs, M. (1997). Interviewing children about trauma: Problems with "specific" questions. Journal of Traumatic Stress, 10, 279-290.

Rao, K., DiClemente, R. J., & Ponton, L. E. (1992). Child sexual abuse of Asians compared with other populations. Journal of the American Academy of Child and Adolescent Psychiatry, 31 (5), 880-885.

Ray, K., & Jackson, J. (1997). Family environment and childhood sexual victimization. Journal of Interpersonal Violence, 12 (1), 3-17.

Runyan, D. K., Everson, M. D., Edelsohn, G. A., Hunter, W. M., & Coulter, M. L. (1988). Impact of legal intervention on sexually abused children. Pediatrics, 113, 647-653.

Sas, L. D., Hurley, P., Hatch, A. Malla, S., & Dick, T. (1993). Three years after the verdict: A longitudinal study of the social and psychological adjustment of child witnesses referred to the child witness project. Ottawa, Canada: Health and Welfare Canada.

Sattler, J. S. (1998). Clinical and forensic interviewing of children and families: Guidelines for the mental health, education, pediatric, and child maltreatment fields. San Diego, CA: Jerome M. Sattler, Publisher, Inc.

Saywitz, K. J. (1989). Children's conceptions of the legal system: 'Court is a place to play basketball.' In S. J. Ceci, D. F. Ross, & M. P. Togli (Eds.), Perspectives on Children's Testimony (pp. 131-157). New York: Springer-Verlag.

Saywitz, K. (1995). Improving children's testimony: The question, the answer and the environment. In M. Zaragoza, G. Graham, G. N. Hall, R. Hirschmann, & Y. Ben-Porath (Eds.), Memory and testimony in the child witness (pp. 113-140). Thousand Oaks, CA: Sage.

Saywitz, K. J., & Camparo, L. (1998). Interviewing child witnesses: A developmental perspective. Child Abuse & Neglect, 22 (8), 825-843.

Saywitz, K. J., & Nathanson, R. (1993). Children's testimony and their perceptions of stress in and out of the courtroom. Child Abuse & Neglect, 17 (5), 613-622.

Saywitz, K. J., & Snyder, L. (1993). Improving children's memory with preparation. In G. S. Goodman & B. L. Bottoms (Eds.), Child victims, child witnesses: Understanding and improving testimony (pp. 71-93). New York: Guilford.

Seifer, R., Sameroff, A. J., Baldwin, C. P., & Baldwin, A. (1992). Child and family factors that ameliorate risk between 4 and 13 years of age. Journal of the American Academy of Child and Adolescent Psychiatry, 31 (5), 893-903.

Shields, A., & Cicchetti, D. (1998). Reactive aggression among maltreated children: The contributions of attention and emotion dysregulation. Journal of Clinical Child Psychology, 27, 381-395.

Small, M. A., & Melton, G. B. (1994). Evaluation of child witnesses for confrontation by criminal defendants. Professional Psychology: Research and Practice, 25 (3), 228-233.

Smith, B., Elstein, S., Trost, T., & Bulkley, J. (1993). The prosecution of child sexual and physical abuse cases. Final report to the National Center on Child Abuse and Neglect, Washington, DC.

Spaccarelli, S. (1994). Stress, appraisal, and coping in child sexual abuse: A theoretical and empirical review. Psychological Bulletin, 116 (2), 340-362.

Sternberg, K. J., Lamb, M. E., Esplin, P. W., & Baradaran, L. P. (1999). Using a scripted protocol in investigative interviews: A pilot study. Applied Developmental Science, 3 (2), 70-76.

Steward, M. S., & Steward, D. S. (1996). Interviewing young children about body touch and handling. Monographs of the Society for Research in Child Development, 61, 4-5.

Suen, H. K., & Ary, D. (1989). Analyzing quantitative behavioral observation data. Hillsdale, NJ: Lawrence Erlbaum.

Swim, J. K., Borgida, E. & McCoy, K. (1993). Videotaped versus in-court witness testimony: Does protecting the child witness jeopardize due process? Journal of Applied Social Psychology, 23 (8), 603-631.

Topf, M. (1988). Interrater reliability decline under covert assessment. Nursing Research, 37 (1), 47-49.

Underwager, R., & Wakefield, H. (1992). Poor psychology produces poor law. Law and Human Behavior, 16 (2), 233-243.

Urquiza, A. J. (1988). The effects of childhood sexual abuse in an adult male population. Unpublished doctoral dissertation, University of Washington, Seattle.

van der Kolk, B. A., McFarlane, A. C., & Weisaeth, L. (Eds.). (1996). Traumatic stress: The effects of overwhelming experience on mind, body, and society. New York: Guilford.

Virginia Conference (1997). A search for truth: Use of closed-circuit and videotaped testimony of children. June 26-27, 1997, Alexandria, VA.

Vogeltanz, N. D., & Drabman, R. S. (1995). A procedure for evaluating young children suspected of being sexually abused. Behavior Therapy, 26, 579-597.

Westcott, H., Davies, G., & Horan, N. (1998). Evaluating investigative interviews for suspected child sexual abuse carried out by English police officers. Poster session presented at the biannual conference of the American Psychology-Law Society, Hilton Head, SC.

Westcott, H. L., Davies, G. M., & Spencer, J. R. (in press). Children, hearsay and the courts: A perspective from the UK. Psychology, Public Policy & Law.

Whitcomb, D. (1993). Child victims as witnesses: What the research says. Newton, MA: Education Development Center, Inc.

White, S., & Edlestein, B. (1991). Behavioral assessment and investigatory interviewing. Behavioral Assessment, 13, 245-264.

Wildt, A. R., & Ahtola, O. T. (1978). Analysis of covariance. Newbury Park, CA: Sage.

Williams, M. B. (1993). Assessing the traumatic impact of child sexual abuse: What makes it more severe? Journal of Child Sexual Abuse, 2(2), 41-59.

Wolfe, D. A., Sas, L., & Wekerle, C. (1994). Factors associated with the development of posttraumatic stress disorder among child victims of sexual abuse. Child Abuse & Neglect, 18, 37-50.

Wolfe, V. V., Sas, L., & Wilson, S. K. (1987). Some issues in preparing sexually abused children for courtroom testimony. The Behavior Therapist, 10(5), 107-113.

Wood, B., Orsak, C., Murphy, M., & Cross, H. J. (1996). Semistructured child sexual abuse interviews: Interview and child characteristics related to credibility of disclosure. Child Abuse & Neglect, 20(1), 81-92.

Zajac, R., Gross, J., & Hayne, H. (2000, March). Courtroom questioning and children's testimony: Do the questions matter? Poster session presented at the biannual conference of the American Psychology-Law Society, New Orleans, LA.