

CRITICAL-THINKING SKILLS AND DECISION-MAKING ABILITIES
OF INVESTIGATORS OF CHILD ABUSE CASES

by

DAVID ROBERT MANDEL

B.A., Concordia University, 1989

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

Department of Psychology

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

August 1992

© David Robert Mandel, 1992

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Psychology

The University of British Columbia
Vancouver, Canada

Date September 22, 1992

Abstract

Dramatic increases in the number of reported cases of child abuse have necessitated attempts to improve front-line workers' investigative skills. This study examined the relation between investigators' critical-thinking skills and decision-making abilities in the context of a hypothetical scenario of a reported case of child abuse, and explored the relations between these two investigative abilities and general reasoning skills, job-related beliefs, and personality dimensions. As part of a 3-day interviewing and credibility assessment workshop, 150 respondents completed a survey, providing their reasons for their disagreement/agreement with a premature intervention made in a scenario of a reported case of child abuse. Accounting for 48% of the variability in investigators' decisions, higher levels of critical thinking and lower levels of perceived responsibility for ensuring a child's safety in a reported case of child abuse significantly predicted stronger disagreement with the premature intervention. Accounting for 17% of the variability in investigators' levels of critical thinking, (a) sex, (b) beliefs about the percentage of reported cases of child abuse that, in general, are true, and (c) levels of hostility emerged as significant predictors, such that being female, estimating a relatively low percentage of true cases, and being relatively less hostile were predictive of higher levels of critical thinking. No significant differences in investigative abilities were found between investigators and

other non-investigative professionals. Implications and suggestions for future research are discussed.

Table of Contents

Abstract.....	ii
Acknowledgements.....	v
Introduction.....	1
The Present Study.....	6
Investigative Abilities and Psychological Factors.....	9
Method.....	17
Overview.....	17
Respondents.....	18
Diagnostic Survey.....	19
Section A.....	19
Section B.....	23
Section C.....	26
Section D.....	27
Section E.....	28
Results.....	28
Critical Thinking and Decision Making in the Investigative Domain.....	29
Investigators and Other Professionals.....	30
Differences Between Social-Work- and Police Investigators.....	31
Investigative Abilities, General Reasoning, Job-related Beliefs, and Personality Dimensions...	33
General Reasoning Skills.....	33
Job-related Beliefs and Activities.....	34
Personality Characteristics.....	35
Predicting Investigators' Decisions and Critical-Thinking Skills.....	36
Discussion.....	38
Predicting Investigative Decisions.....	42
Predicting Critical Thinking.....	49
Additional Relations.....	53
Future Research.....	58
Conclusion.....	61
References.....	63
Footnotes.....	68
Statistical Tables.....	70
Appendix A: Diagnostic Survey.....	73

Acknowledgements

I thank my committee members, Darrin R. Lehman, John C. Yuille, and Jerry S. Wiggins for their invaluable help throughout all phases of this project. I also thank Victoria Desroches, Sarah M. Smith, and Paul D. Trapnell for their assistance.

The dramatic rise in the number of reported cases of child abuse has drawn focus on what was once considered a less significant social problem. It is estimated that approximately 200,000 new cases are reported annually in the U.S. (Finkelhor & Hotaling, 1984). It is reported that between 15% to 38% of females and 3% to 9% of males experience unwanted sexual activity during their childhood or youth in the U.S. (De Luca, Boyes, Furer, Grayston, & Hiebert-Murphy, 1992) And, although exact figures are difficult to determine (Russell, 1983), findings in the Report on the Federal Commission on Sexual Offences Against Children and Youths (Badgley, 1984) indicated that as many as one in four girls and one in eight boys experience unwanted sexual encounters in Canada. Moreover, despite the existing debate over the psychological impact of sexual abuse (De Luca et al., 1992), and the paucity of research on this topic (Browne & Finkelhor, 1986), existing empirical findings do indicate that a substantial number of abuse victims encounter serious mental-health problems (Browne & Finkelhor, 1986).

Such findings have forcibly called into question the assumption that children's testimony is inadmissible. Recent literature on the accuracy of children's evidence generally suggests that, although the amount of information recalled is directly related to a child's age (King & Yuille, 1987), and children are more susceptible to leading questions than adults (Hughes & Grieve, 1980; Zargoza, 1987), if children are

interviewed properly then the accuracy of their testimony is comparable to that of adults' accounts (Yuille, 1988).

The evidence attesting to children's competence to provide accurate testimony has led some investigators (e.g., Melton & Thompson, 1987) to suggest that more attention ought to be directed toward other issues, such as the task of assessing the credibility of children's testimony. Although the recent increase in acceptance of children's accounts is a positive development in general, and the credibility of children's statements concerning abuse is supported by relatively high rates of offenders' admissions (Jones & Krugman, 1986), a small percentage of false disclosures nevertheless does pose a serious problem. False disclosures threaten to undermine the credibility of children's statements in general (Yuille, 1988), and exact significant costs on children, their families, and those wrongly accused (Wakefield & Underwager, 1988).

There is evidence to suggest that cases of false disclosure often involve the manipulation of children by adults who are engaged in their own disputes (Yuille, 1988). Whereas false disclosure rates of 6% or 7% were found in overall samples of reported cases of child sexual abuse (e.g., Goodwin, Sahd, & Rada, 1978; Jones & McGraw, 1987), this figure increased dramatically to 28% (Jones & Seig, 1988) or 36% (Green, 1986) in the context of custody and visitation disputes. In other cases, false allegations are attributable to children who wish to "punish" a parent (Green, 1986;

Klajner-Diamond, Wehrspann, & Steinhauer, 1987; Yuille, 1988) or to psychiatrically-ill parents who launch false reports (Green, 1986; Klajner-Diamond et al., 1987). Nevertheless, as Coolbear (1992) reminds us, all allegations of child abuse must be carefully investigated regardless of the specific context of its occurrence.

Establishing whether children's disclosures are credible naturally involves a number of investigative skills on the part of front-line workers. Yet, this responsibility usually falls on social workers and police. Given social workers' traditional role as child advocates, their professional training would most probably have emphasized intervention and therapeutic techniques, but little in the way of interviewing and assessment procedures (Hunter, Yuille, & Harvey, 1990). In contrast, police officers primarily are concerned with potential criminal charges, and their mandate is to gain as much uncontaminated information as possible (Hunter et al., 1990). Nevertheless, police officers also receive only superficial training in interviewing skills that are appropriate for use with children (Yuille, 1984).

Attempts to improve upon the current state of investigations of reported child abuse cases have prompted the development of structured interviewing and credibility assessment techniques, such as the Step-Wise Interview (Raskin & Yuille, 1989; Yuille, Hunter, Joffe, & Zaparniuk, 1992) and Statement Validity Analysis (SVA; Raskin & Yuille, 1989; Steller & Koehnken, 1989). Although there are a number of

field studies suggesting that SVA is a useful technique for assessing the credibility of children's statements, much of this empirical work is subject to a long list of methodological shortcomings (Raskin & Yuille, 1989; Steller & Koehnken, 1989). Moreover, training in techniques such as the Step-Wise Interview and SVA also is currently limited in its availability, as it is generally offered to trainees as a workshop.

A number of disconcerting findings concerning the acquisition and application of procedural knowledge of current credibility assessment techniques recently have been reported. Coolbear (1992) suggested that even when professionals do undergo such training in interviewing and credibility assessment techniques they may have difficulty putting these skills into practice. For instance, she reported that almost none (4%) of a group of legal- and human-service professionals having attended workshops on child sexual abuse offered the basic Step-Wise Interview sequence of allowing the child to provide free narrative of the event, followed by open-ended questions, and then specific questions. And, Coolbear found that of the 19 content criteria in SVA, 16 received little or no mention leading her to conclude that "professionals [even with training in structured interviewing and assessment techniques] seem to consider only a few aspects of the actual content of the child's statements in establishing credibility" (p. 162).

Similar results have been reported by Yuille (1992a). In

a major field study testing the utility of both the Step-Wise Interview and the SVA procedure, he found that, while the Step-Wise Interview procedure was effective in improving the quality of interviews, front-line workers had difficulty learning and applying SVA. These latter findings, which are consistent with results by Joffe and Yuille (1991), support Undeutsch's (1982) assertion that the successful application of criteria-based analysis (i.e., Statement Reality Analysis or SVA) requires extensive training and experience, characteristic of clinical training.

More recently, Yuille (personal conversation, May, 1992) has suggested that the ineffective application of SVA by front-line workers might be, at least, partly a result of their lack of critical-thinking skills. This notion is consistent with Coolbear's (1992) observation that investigators often selectively attend to a limited set of information in a case, thus undermining Undeutsch's (1984) suggestion that all of the individual pieces of information in a case should be combined into a "specific and non-interchangeable configuration of facts" (p. 64). Along similar lines, a number of researchers (e.g., Klajner-Diamond et al., 1987; Wakefield & Underwager, 1988; Wehrspann, Steinhauer, & Klajner-Diamond, 1987; Yuille, 1988) have suggested that investigators who believe that virtually all children's statements are true are relatively more likely to conduct misleading interviews that elicit responses supporting their own assumptions about cases of child abuse. These

investigators are relatively more likely to selectively attend to data confirming their own hypotheses, distort facts that are inconsistent with their beliefs, and fail to examine alternative explanations for information presented to them in such cases.

In short, despite the important social benefits that would be accrued by helping investigators to deal more effectively with the increasing number of child abuse cases, and early signs of promise for techniques such as the Step-Wise Interview and SVA, research focused on the improvement of front-line workers' investigative abilities is still in its infancy. Additional controlled research is sorely needed. Moreover, the suggestions that front-line workers' critical-thinking skills are prerequisites for effective, investigative decision-making need to be examined empirically.

The primary foci of the present study were to examine, using a large, representative sample of front-line workers, (a) the relation between critical-thinking skills and decision-making abilities in the context of a hypothetical scenario of a reported case of child abuse, and (b) the relations between these two investigative abilities and a variety of other psychological characteristics, including general reasoning, job-related beliefs, and personality dimensions.

The Present Study

In collaboration with Drs. Yuille and Lehman, I have had a rare opportunity to examine a sample of 150 front-line

workers and other professionals, 91 of whom are responsible for investigating reported cases of child abuse. Participants in this unique sample took part in one of four, 3-day, interviewing and credibility assessment workshops that were designed and conducted by Dr. Yuille. During these workshops, that focused on the Step-Wise Interview- and SVA procedures, respondents completed a brief diagnostic survey designed to tap a broad range of information on their psychological characteristics and their investigative abilities.

A hypothetical scenario of a reported case of child abuse and neglect was employed as a criterion measure of respondents' critical-thinking skills and decision-making abilities. In the scenario, a social worker named Chris responds to an anonymous caller who states that he is aware of a case of child abuse and neglect. Chris conducts a relatively superficial investigation and, although stereotypic pieces of information which might prime people to believe that abuse and neglect were occurring emerge, in fact, little in the way of evidence that would substantiate the caller's allegations is detected. Nevertheless, Chris intervenes by taking the child into care. After reading the scenario, respondents are asked to indicate the extent of their disagreement/agreement with Chris's intervention, and to list the features of the scenario that led them to their decision.

Respondents' open-ended comments were coded in terms of their critical and uncritical features, and then were combined to form a single critical-thinking index that reflected

whether respondents' statements were solely uncritical (score of 1), were both uncritical and critical in certain respects (score of 2), or were solely critical (score of 3).

Indicators of critical thought included requests for additional information, concerns over the lack of thorough investigation by Chris, analytical statements in which alternative interpretations of case information were given, and other skeptical remarks that would indicate an awareness of the need to withhold judgment until more needed information was obtained. Indicators of uncritical thought included simple lists that merely reiterated features presented in the case without explicitly stating the role or weight that these features had in terms of their agreement/disagreement with Chris's intervention, and statements reflecting unwarranted assumptions that were held as facts.

Following the aforementioned suggestions in the literature that poor critical-thinking skills might undermine effective decision making, I anticipated that investigators providing only critical statements would provide the strongest disagreement with the premature intervention, followed by investigators providing a mix of critical and uncritical statements. Investigators providing only uncritical comments were expected to indicate the highest level of agreement with Chris's premature decision to remove the child.

Because respondents' occupations and primary duties were recorded, it also was possible to examine the group of investigators separately from those respondents involved in

other activities, and to contrast even finer subsamples, such as social-work investigators and police investigators.

Despite the suggestions that some investigators are prone to a number of inferential shortcomings, it was expected that, on the whole, investigators would be more critical and make more skeptical decisions than would non-investigators.

Investigators must deal with the problems of gathering and interpreting case information on a regular basis, and they are ultimately accountable for the important decisions that they make. It is plausible that their experience in dealing with such cases, especially those in which information is scant, would make them more aware than other professionals of the difficulties and uncertainties inherent in interpreting evidence and making decisions regarding appropriate actions. Presumably, then, compared with other respondents, investigators should be more likely to be aware of the weak evidence in the case scenario, and the fact that Chris could have easily gathered much more useful information, and subsequently less supportive of Chris's premature intervention. An exploratory examination of the differences in investigative abilities between social-work investigators and police investigators also was conducted.

Investigative abilities and psychological factors. Given the seriousness of child abuse and the importance that thorough investigation has in establishing the veracity of allegations, many avenues of research aimed at improving the quality of front-line workers' investigations and assessments

need to be explored. One such area that has not received attention, but that might very well be profitably pursued, involves examining normative psychological characteristics of investigators in an attempt to isolate features that predict effective investigative abilities. There are a number of psychological dimensions that represent probable candidates.

An abundance of literature in social psychology indicates that people, in general, are "cognitive misers" who, in over-utilizing simple, intuitive heuristics and under-utilizing more complex, formal decision rules, conform to a principle of least effort, rendering themselves vulnerable to a long list of inferential errors (e.g., Abelson & Levi, 1985; Einhorn & Hogarth, 1981; Fiske & Taylor, 1991; Kahneman, Slovic, & Tversky, 1982; Nisbett & Ross, 1980). Investigators often must work with incomplete information and are forced to make inferences and judgments that have important consequences. Perhaps, then, investigators who are equipped with relatively strong general reasoning abilities are better able to conduct a thorough assessment and draw reasonable conclusions from available evidence. Inductive reasoning abilities that reflect statistical and methodological sophistication are associated with a healthy skepticism towards premature conclusions, and these skills can increase the likelihood that alternative explanations for events will be generated early on (Mill, Gray, & Mandel, 1992), perhaps also resulting in a more thorough information search. Deductive skills, such as a clear understanding of material conditional and biconditional

implications (Wason, 1966), might also be useful to front-line investigators who must deal with the syntactic relations between pieces of evidence in assessing the logical consistency and credibility of allegations and other statements.

In this study, respondents answered four general reasoning questions (i.e., three that tapped inductive reasoning skills, and one that tapped deductive abilities; see section D in methods). In contrast to the measure of critical thinking employed in this study, these reasoning questions were not staged within an investigative context, and each of these items tapped a specific reasoning skill (viz., awareness of [a] the importance of comparative information, [b] the potentially confounding effects of extraneous variables, [c] the law of large numbers, and [d] conditional implications). It was hypothesized that high performance across these items, reflecting an understanding of general reasoning principles, would be associated with higher levels of critical thinking and skeptical decision-making in the investigative domain. Such a finding would be encouraging, especially in light of recent evidence suggesting that even brief training in inductive reasoning principles is effective in increasing the frequency and appropriate use of those skills in subsequent problem-solving tasks (e.g., Fong, Krantz, & Nisbett, 1986).

In addition to general reasoning skills, a number of job-related beliefs might be meaningfully related to investigative abilities. As mentioned above, some researchers (e.g.,

Klajner-Diamond et al., 1987; Wakefield & Underwager, 1988; Wehrspann et al., 1987; Yuille, 1988) have suggested that investigators who believe that virtually all children's statements are truthful might be uncritical interviewers and decision makers. Along similar lines, one might expect that investigators who believe that virtually all reported cases of child abuse are true would be less critical than their more suspicious counterparts. One plausible hypothesis is that the more frequently that investigators expect such cases to be true the less critical they will be in scrutinizing the evidence for alternative explanations that would suggest that allegations made are false. In such a case, it would be expected that investigators' beliefs about the percentage of reported cases of child abuse which are true, in general, (see question E-9 in Appendix A) would be inversely related to their levels of critical thinking, and would be directly related to their levels of agreement with the premature intervention.

Another hypothesis is that the degree to which investigators perceive themselves as being responsible for questioning the truthfulness of reported claims of abuse (see question C-8 in Appendix A) will be directly related to the quality of their investigative abilities. It is expected that the more that investigators think that they are responsible for carrying out a thorough investigation, the more critical they should be, and the less likely they should be to support the decision to remove the child, since that decision is

unwarranted in the present study. In contrast, it also is hypothesized that investigators' perceived responsibility for ensuring children's safety (see question C-7 in Appendix A) will be directly related to their willingness to support the decision to remove the child in the scenario. Investigators with relatively high levels of perceived responsibility for ensuring child safety might be inclined to intervene by removing a child when there is any indication at all that a child might be in danger, even in situations where the evidence in support of abuse is far from conclusive.

It was predicted that respondents who find it useful to develop an hypothesis early on in an investigation (see question C-6 in Appendix A) would be less critical in their statements concerning the case intervention. The premature formation of hypotheses should limit information searches and the generation of alternative explanations and increase the likelihood of selective attention, misinterpretations, and other confirmation biases on the part of investigators (Klajner-Diamond et al., 1987; Wakefield & Underwager, 1988; Wehrspann et al., 1987; Yuille, 1988).

Along somewhat similar lines, it was hypothesized that investigators' level of decisional uncertainty vis à vis the investigation of reported cases of child abuse would be directly related to the quality of their investigative abilities. A 3-item scale, employed to tap this uncertainty, included questions concerning their beliefs about the unavailability of important information on which to base a

decision (see question C-1 in Appendix A), the ambiguity of case information (see question C-9), and the resulting uncertainty concerning the quality of investigative decisions (see question C-5). It is expected that investigators who, in their everyday professional experience, are aware that information is often limited and that multiple interpretations of available case evidence often exist, and who, in retrospect, have questioned whether some of their decisions were premature, would have higher levels of critical thinking, and would express stronger disagreement with the case intervention, than those investigators who feel more certain.

Personality factors of investigators might also be predictive of good investigative abilities. Positive findings might suggest important information to consider when screenings for new investigators take place. Fletcher, Danilovics, Fernandez, Peterson, and Reeder (1986) have developed an individual difference scale designed to measure attributional complexity. Fletcher et al. reported that attributionally-complex individuals spontaneously produce more causal attributions for personality dispositions and generate more complex causal attributions for behavioral events compared with their attributionally-simpler counterparts. Since investigating allegations of child abuse not surprisingly involves having to make a number of attributions about the motives and intentions of people involved in such cases, one might predict that attributional complexity will be directly related to the quality of investigative abilities.

Considering the scenario employed in this study, one might posit that attributionally complex respondents would be reasonably more skeptical of the anonymous caller's motives than their less attributionally-complex counterparts. By the same token, however, attributionally-complex investigators might also generate more unwarranted causal attributions (e.g., "the social worker was responding to his or her own prejudices"). To test these possibilities, the relation between criterion measures and a 14-item (half-length) version of Fletcher et al.'s (1986) attributional complexity scale will be examined.

Factors II, III, and V of the five-factor model of personality, sometimes referred to as the "Big Five" (John, 1990; McCrae & Costa, 1989; McCrae & John, 1992; Trapnell & Wiggins, 1990; Wiggins & Trapnell, in press), also were assessed in this study. Factor II reflects the trait of agreeableness. A measure of agreeableness was included in this study because the wording of certain, negatively-keyed items denoting this factor are suggestive of skepticism (e.g., I tend to be cynical and skeptical of others' intentions). In contrast to those who are more trusting, respondents who are wary of others' intentions might be less likely to accept information as "fact" solely on the basis of its face validity.

Factor III reflects the trait of conscientiousness. At an intuitive level, it is plausible to suggest that investigators who are high in conscientiousness will make

better investigative decisions since they strive more for excellence and work harder to accomplish their goals compared with those who score low on Factor II. This prediction, however, must be tempered by the fact that not all professionals who are involved with reported cases of child abuse have the same mandates (Hunter et al., 1990; Yuille et al., 1992). Social workers, for example, might perceive themselves as being more responsible for child protection and less responsible for actual investigations than police, who might be relatively more concerned with determining the truth in reported cases of child abuse. To the extent that this is so, the relation between conscientiousness and decision making might be moderated by investigators' professions. It was thus expected that conscientiousness would be directly related to levels of agreement with the premature intervention (a decision which was consistent with child safety concerns in this study) among social-work investigators, but inverse related to levels of agreement among police investigators.

Factor V marks the trait of openness. Trapnell (1992b) has argued in favor of a three-factor conception of this broad trait consisting of absorption (i.e., aesthetic sensitivity), intellectance (i.e., intellectual curiosity; positively related to absorption), and traditionalism (i.e., conventionalism, religiosity, prudery, propriety, and authoritarianism markers; negatively related to intellectance and absorption). In this study, a short scale consisting of three intellectance items and one absorption item was employed

to test the hypothesis that openness is related to investigative abilities. Investigators who have a lot of intellectual curiosity might be more inclined to treat a case as a "puzzle" to be solved, and hence be more motivated to explore a case more fully than less curious individuals.

In summary, the primary purpose of this study was to test the hypothesis that higher levels of critical thinking would be associated with better investigative decision-making. As well, an exploratory analysis of the relations between these two investigative abilities and a number of other psychological features, including general reasoning skills, job-related beliefs, and personality factors, was undertaken.

Method

Overview

As part of four, 3-day, interviewing and credibility assessment workshops (see Yuille, 1992b, for a detailed description), 150 respondents in Canada and the U.S. spent approximately 20 minutes anonymously completing a diagnostic survey that probed their responses to (a) a hypothetical scenario of a reported case of child abuse and its investigation, (b) a 29-item personality inventory, (c) 10 items tapping job-related beliefs and activities, (d) a 4-item reasoning test, and (e) various questions concerning demographics and professional characteristics. A yoked design was employed such that respondents from two cities (viz., Goderich, Ontario and Vancouver, British Columbia), constituting roughly half of the total sample, completed the

survey in the first half of the seminar while the others (viz., Hamilton, Ontario and Salamanca, New York) completed it near the end of training. Respondents were surveyed within a two-month period during May and June, 1992.

Respondents

One hundred and fifty professionals enrolled in interviewing and credibility assessment workshops voluntarily participated in the study. Data from three respondents were dropped because they failed to provide sufficient information enabling their encoding on important grouping variables (e.g., profession, duties, etc.). The remaining 147 (105 female and 42 male) respondents attended a workshop in either Salamanca, New York ($n = 36$), Hamilton, Ontario ($n = 43$), Goderich, Ontario ($n = 42$), or Vancouver, British Columbia ($n = 26$). The vast majority of the sample were social workers ($n = 106$); other professional groups included police officers ($n = 19$), administrators ($n = 11$), therapists or counselors ($n = 4$), and others ($n = 7$). Respondents ranged in age from 22-57 years ($M = 36$ years, $SD = 9$ years), and their professional experience varied widely from less than one year up to 29 years ($M = 9$ years, $SD = 8$ years). Similarly, of the 129 respondents who reported having dealt professionally with reported cases of child abuse, the length of their experience in dealing with such cases was highly variable, ranging from 1-23 years ($M = 6$ years, $SD = 5$ years). Of the 143 respondents who indicated their primary duties, 91 were involved as investigators, 30 worked in treatment and counseling, 21 were intake workers, 21

worked in child protection, 14 were supervisors, 2 were responsible for referrals, and 9 were uncodable.¹ Amongst the 143 respondents who provided information on their highest level of education achieved, 11 were high-school graduates, 7 had college diplomas, 93 had one or more Bachelor's degrees (of these respondents, 39 had Bachelor's degrees in Social Work), 29 had Master's degrees (of these respondents, 21 has Master's degrees in Social Work), and 4 were uncodable.

Diagnostic Survey

This package (see Appendix A) consists of five major sections, each described below.

Section A. This section provided criterion measures of respondents' critical-thinking skills and decision-making abilities. Respondents were asked to read and carefully consider a hypothetical scenario of a reported case of child abuse and neglect, and its subsequent investigation by an "experienced" social worker named Chris, who decided to take the child into care (see Appendix A for the entire scenario).

The information presented in the case was designed to constitute very weak evidence in support of the allegations. Nevertheless, this information might appear to indicate potential abuse or neglect to a naive investigator. For example, in the scenario, Chris finds that the child is dirty, that there are bruises on both of the child's knees, and that there is a scrape on the child's elbow. While these features might appear to indicate abuse or neglect they actually are highly consistent with more mundane, and more probable,

explanations. That is, it is quite common for children to get scrapes or bruises on their knees and elbows, and to get "dirty," just from playing outdoors. The location of bruises should be a critical factor to consider in assessing the veracity of the allegations. If the child had bruises on the back of his or her knees, or in some other location that would normally be improbable, then such information rightly should alert an investigator to at least monitor the situation more carefully. Similarly, Chris's initial impression of the residence is that it is "messy." Chris also finds "little food in the refrigerator." While such "information" might be indicators of neglect, it might just as well reflect Chris's own personal biases. The scenario states that the residence is located in "a very poor area of the city..." Although this piece of information does not preclude the possibility of neglect, it does suggest that the conditions described might be the result of a poorer standard of living than Chris is accustomed to experiencing. That is, the household conditions described in the scenario might very well be characteristic of most of the households in that neighbourhood.

Chris's interviewing of the child also was very superficial. For example, Chris only asked the child two questions (viz., Where are your parents? How long have you been alone?) to which the child responded with cursory answers. There is no indication that Chris asked for clarification regarding these responses despite the fact that many possible interpretations exist--interpretations which

could probably have been clarified with further questioning. For example, the child might or might not have had an accurate sense of time or might have been afraid to answer a stranger. If it was a weekday, the child might have returned home from school fairly recently since it was 4:15 p.m. when Chris arrived at the residence. A parent may have just stepped out momentarily, etcetera. Moreover, there is no information in the scenario suggesting that Chris conducted any other follow-up investigations, such as checking for instructions or telephone numbers left for the child in case of an emergency or inquiring with neighbours. In short, no single feature or combination of features in the case scenario provides compelling evidence in support of the allegation of child abuse or neglect, and some information should have raised suspicions about the veracity of the child-abuse claim (e.g., the fact that the caller who reported the case was anonymous).

After reading the scenario, respondents were asked (a) to indicate their level of disagreement or agreement on a 1 (strongly disagree) to 7 (strongly agree) scale with Chris's actions, and (b) to list the features of the case that led them to their own decision. Their open-ended responses were content coded in terms of two major categories, reflecting critical and uncritical thought, respectively. A response was coded as being critical if it included (a) one or more requests for additional information concerning the case; requests could be general (e.g., "I need more information to make a clear judgment") or specific (e.g., "were there any

relatives periodically checking on the child?"); (b) any comments concerning the inadequacy of the investigation (e.g., "Chris should have checked with neighbours" or simply "needs more investigation"); (c) reasonable alternative explanations for "facts" in the case (e.g., "the child's scrapes might have come from normal playing" or "the parents might have just stepped out momentarily"); and/or (d) skeptical and/or analytical statements (e.g., "the caller was anonymous," "the developmental levels of 8 year old children can vary," or "home conditions might reflect systemic factors; i.e., impoverished community standards").

A response was coded as uncritical if it included (a) a simple list of features already provided in the scenario (e.g., "child alone, bruises on knees, dirty, etc.") or (b) an unwarranted assumption about the case (e.g., "the child has probably experienced nonaccidental trauma"). It is important to note that while the presence of critical statements are likely to be associated with skeptical decisions, and uncritical statements with decisions in support of Chris's actions, they need not necessarily be so. An individual who indicates support for Chris's action, and subsequently states "I think more information should be gathered, but I nevertheless have to agree with the social worker's decision," is still coded as being critical, despite the fact that some other factor that was not explicitly addressed (e.g., concerns over the child's safety) appears to have been more heavily weighted than the need to examine more information. In a

similar vein, a respondent who strongly disagrees with Chris's action, and who also provides many critical comments might be coded as uncritical as well for making an unwarranted comment like "the social worker's decision was based purely on prejudice."

Responses were coded dichotomously for the presence or absence of statements indicative of each of the two categories. A response could be coded in one, both, or neither of the categories; however, none of the responses were uncodable. Inter-rater reliabilities, calculated between the author and a blind, independent coder, and based upon a randomly selected sample of 40 responses, were 100% for critical statements and 95% for uncritical statements. Inter-rater reliabilities corrected for the frequency of positive category occurrences were calculated by taking the percent agreement over only those cases where at least one of the coders indicated category presence; this conservative approach yielded similar findings: 100% for critical statements and 88% for uncritical statements. Discrepancies were resolved by the author, who also coded the remaining responses. Content codes for these two dichotomous variables were combined to form a critical-thinking index: responses that were coded only as uncritical were scored as 1; responses including a mix of uncritical and critical statements were scored as 2; and responses that were coded only as critical were scored as 3.

Section B. Consisting of a 29-item inventory, this section was designed to tap various individual difference

constructs. Respondents were instructed to indicate on a 1 (strongly disagree) to 5 (strongly agree) scale the option which best represented their opinion concerning the statement in question.

Fourteen items (viz., items 1, 4, 7, 11, 12, 13, 17, 21, 22, 23, 24, 26, and 29) were selected from Fletcher et al.'s (1986) 28-item attributional complexity measure. Two items were selected from each of the measure's seven subscales². Eight of the 14 items were worded to reflect attributional simplicity and subsequently were scored in the reverse direction. The factor loadings for this subset of items in the derivation sample ranged from .33 to .64 ($M = .50$), and the mean corrected item-total correlation ranged from .31 to .56 ($M = .43$; see Fletcher et al., 1986 for additional information on this measure). In the present study, the corrected item-total correlations for the 14-item attributional complexity scale ranged from .26 to .52 ($M = .36$). The internal reliability coefficient (Cronbach's alpha) was .75; this value was virtually identical to the estimate of .74 forecasted by the Spearman-Brown formula (see Ghiselli, 1964) for a test of half length given the original alpha coefficient of .85 obtained by Fletcher et al. (1986) for the full, 28-item measure.

Twelve of the remaining 15 items in Section B were selected to mark factors II, III, and V of the five-factor model using three, short scales. Items 2, 8, 15, and 20, were selected from Costa and McCrae's (1989) NEO Five Factor

Inventory (FFI) to mark Factor II (Agreeableness). Factor loadings for these four items ranged from .47 to .61 ($M = .54$) in the Trapnell and Wiggins (1992) study, and were among the six highest loading items.³ Unfortunately, all of the positively keyed items had both substantially lower loadings on Factor II and less preferable wording for the purposes of the present study. Thus, the four, negatively-keyed agreeableness scale items in this study, reflecting the arrogant/calculating (BC) octant of Wiggins' (1982) circumplex model of interpersonal behavior, were reverse scored. The corrected item-total correlations for this scale obtained in this study ranged from .24 to .42 ($M = .34$). Cronbach's alpha was .55; this value meets the criterion of .50 for alpha reliability employed by some researchers in constructing short scales (e.g., Hogan, 1985; Trapnell, 1992).

Items 5, 14, 19, and 25, also selected from the FFI, were employed to mark Factor III (Conscientiousness). Trapnell and Wiggins (1992), administering the FFI to a large sample ($N > 600$) of university undergraduates, found that these four items yielded the highest loadings on Factor III, ranging from .64 to .70 ($M = .66$). Three of these four items were worded to reflect high conscientiousness; the remaining item was scored in the reverse direction. In the present study the corrected item-total correlations for this 4-item conscientiousness scale ranged from .30 to .51 ($M = .43$). Cronbach's alpha was .63.

Items 3, 6, 9, 10, 16, 18, and 28 of Section B were

combined to form a 7-item openness scale marking Factor V. Items 16 and 28 (both are intellectance markers) were selected from the FFI, and had respective factor loadings of .44 and .65 on Factor V (Trapnell & Wiggins, 1992). Items 9 (intellectance) and 18 (absorption) were taken from the reflection subscale of Trapnell's (1992a) Reflection-Rumination Questionnaire (RRQ). Reflection correlates highly ($r > .60$) with Factor V (Trapnell, 1992), and Trapnell and Wiggins (1990) report that trait adjective items that make reference to philosophical and/or abstract thinking (see item 9) have the highest factor loadings on Factor V. Items 3 and 6 were taken from the openness to values- and openness to actions subscales of the NEO Personality Inventory (Costa & McCrae, 1985). Item 10 was selected from an 8-item scale measuring attitudinal conformity (Trapnell, 1991), which is negatively correlated with openness (Trapnell, 1991). Three of these four items were worded to reflect high openness (viz., items 16, 18, and 28); the remaining four items were scored in the reverse direction. The corrected item-total correlations for this scale ranged from .16 to .54 ($M = .36$). Cronbach's alpha was .64.

Section C. Ten questions specifically designed for this study probed a number of respondents' job-related beliefs and activities vis à vis reported cases of child abuse. A 3-item scale measuring decisional uncertainty included questions 1, 5, and 9. Relatively high scores on this scale are interpreted as reflecting respondents' beliefs that, in

investigations of reported cases of child abuse, information is often incomplete (see C-1) and/or ambiguous (see C-9), and that the appropriateness of decisions made under such circumstances is often uncertain (see C-5). The corrected item-total correlations among items 1, 5, and 9 were .21, .32, and .36, respectively. Cronbach's alpha was .47. Each of the remaining items (viz., 2, 3, 4, 6, and 10) were treated separately in subsequent analyses.

Section D. This section consists of a 4-item reasoning test. Questions 1, 2, and 3, selected from a battery of items developed by Lehman, Lempert, and Nisbett (1988) and slightly modified in the present study, were employed to tap respondents' inductive reasoning abilities. Specifically, correct responses to questions 1, 2, and 3 respectively indicated respondents' awareness of (a) the potential effects of confounding variables, (b) the importance of having comparative information for evaluating certain claims (i.e., knowledge of appropriate control groups), and (c) the law of large numbers which essentially states that, as a sample of observations increases, fluctuations around the mean also increase but the relative error introduced by such variability tends toward zero (Prigogine & Stengers, 1984). Question 4 was selected from the Cornell Critical Thinking Test (Level 2; Ennis & Millman, 1982) and was slightly modified for this study. This deductive reasoning item, which involves identifying the unstated assumption underlying a given claim, taps respondents' abilities to realize the correct syntactic

relations between the elements of the statement.

The correct responses to items 1-4 are C, A, B, and A, respectively. Responses to these items were coded dichotomously as either correct (1) or incorrect (0). If a respondent left an item blank or circled more than one option for a single item their "response" was coded as incorrect. In addition, an index of the number of correct responses, with a possible range of 0-4, was calculated.

Section E. Respondents provided information on their demographic and professional characteristics in this section. Levels of education achieved by respondents were coded as follows: 1, high-school diploma; 2, college diploma; 3, Bachelor's Degree; and 4, Master's Degree. Respondents also indicated the percentage of reported cases of child abuse that they believe are generally likely to be true.

Results⁴

Because this study primarily is concerned with the relations among investigators' critical-thinking skills and decision-making abilities, and their psychological characteristics, unless otherwise specified, the analyses reported below are based upon this subsample of respondents ($n = 91$). Investigators surveyed earlier on during the workshop and those surveyed near its end did not significantly differ either in their levels of disagreement/agreement with the hypothetical social worker's intervention or in their levels of critical thinking, $t_{s(86)} < 1$. Similarly, no significant differences either in levels of disagreement/agreement with

the intervention or in levels of critical thinking were found between investigators surveyed in each of the four cities in which the workshops were held, $F_s < 1$. Therefore, these groups were collapsed in subsequent analyses.

Critical Thinking and Decision Making in the Investigative Domain

Investigators' indications of disagreement/agreement with Chris's intervention were distributed fairly evenly over the possible range of response options ($n_s = 17, 17, 13, 9, 12, 13$, and 10 for options 1-7, respectively), $\chi^2(6, N = 91) = 4.46, p > .50$, with a mean of 3.67 ($SD = 2.05$), indicative of slight, overall disagreement. Of the 89 investigators who provided reasons for their decision to disagree/agree with Chris's intervention, 21% provided only uncritical comments, 17% provided a combination of both uncritical and critical statements, and 62% provided only critical comments. The presence of critical comments was significantly associated with the absence of uncritical statements, $\chi^2(1, N = 89) = 39.08, p < .0001$.

As predicted, investigators' views of the premature intervention became increasingly skeptical as their statements went from being only uncritical ($M = 5.95, SD = 1.43$), to having both critical and uncritical elements ($M = 4.80, SD = 1.93$), to being only critical ($M = 2.62, SD = 1.41$), $F(2, 86) = 39.09, p < .0001$. A series of a priori comparisons, using multiple t tests, revealed that, as expected, the group of respondents who provided only critical statements were

significantly more skeptical of the decision to remove the child than the group providing only uncritical statements, $t(86) = -8.28$, $p < .001$, or the group providing a mix of critical and uncritical statements, $t(86) = -4.96$, $p < .001$. And, the group providing a mix was more skeptical than the group providing only uncritical statements, $t(86) = -2.12$, $p = .05$.

None of the demographic factors (i.e., sex and age) or, more interestingly, the professional characteristics of investigators (i.e., educational level, profession [viz., social workers vs. police]⁵, years of employment, years of involvement with child abuse cases, or the estimated number of reported child abuse cases investigated) were significantly associated with levels of disagreement/agreement.

A slightly different pattern emerged with the open-ended data. Females ($M = 2.60$, $SD = .67$) were significantly more critical than were males ($M = 2.03$, $SD = .95$), $t'(47) = 2.98$, $p = .005$. No other demographic/professional characteristic of investigators was significantly associated with the presence or absence of either their critical or uncritical thoughts.

Investigators and Other Professionals

Contrary to predictions, no significant difference in levels of disagreement/agreement with the case intervention was found between investigators ($M = 3.67$, $SD = 2.05$) and non-investigators ($M = 3.77$, $SD = 1.98$), $t(141) = .28$. Nor did levels of critical thinking underlying investigators' statements ($M = 2.40$, $SD = .82$) differ from that of other

professionals ($M = 2.35$, $SD = .80$), $t(138) = -.36$.

In fact, comparisons of investigators and other professionals resulted in startlingly few significant differences across any of the variables in this study. These may be summarized as follows: First, investigators were more likely to see themselves as responsible for getting to the truth in reported cases of child abuse ($M = 5.67$, $SD = 1.18$) than non-investigators ($M = 5.00$, $SD = 1.66$), $t'(73) = -2.49$, $p < .05$. Second, investigators rated their investigative abilities ($M = 5.29$, $SD = .88$) more positively than non-investigators ($M = 4.52$, $SD = 1.13$), $t(134) = -4.37$, $p < .001$.

Differences Between Social-Work- and Police Investigators.

As mentioned earlier, although both social workers and police often are directly involved in investigations of reported child abuse cases, they tend to have different educational and professional backgrounds and different agency mandates to fulfill. Table 1 summarizes all of the significant mean differences between social-work investigators and police investigators found in this study (i.e., with the exception of the difference in critical thinking that has already been discussed). As can be seen in Table 1, social-work investigators, in contrast to police investigators: (a) had less professional experience; (b) had achieved higher levels of education; (c) reported that they had removed a child from a home/laid a charge more frequently, despite suspicions that the child's statements were erroneous; (d) perceived themselves as more responsible for ensuring a

potentially abused child's safety and less responsible for questioning the truthfulness of reported claims of abuse; (e) estimated that a smaller percentage of reported cases of child abuse were true both in their own investigative experiences and in general; and (f) had higher levels of attributional complexity and openness and lower levels of hostility. In addition, 65% of police investigators were male, and 72% of social-work investigators were female, $\chi^2(1, N = 89) = 8.26, p < .005$.

Not surprisingly, given that social-work investigators constituted the vast majority of the overall sample of investigators, a significant difference in levels of critical thinking was found between female social-work investigators ($M = 2.61, SD = .67$) and male social-work investigators ($M = 2.15, SD = .93$), $t(69) = 2.32, p < .05$. No other significant relations between criterion measures (i.e., disagreement/agreement and critical thinking) and any of the other demographic or professional characteristics reported above were found within this subsample. Among the group of police investigators, there was a significant correlation between the estimated number of cases investigated and the extent to which police agreed with the intervention in the case scenario, $r(15) = .58, p < .01$. Interestingly, the more cases that police officers had investigated, the less skeptical they were in the case scenario. No other significant relations between criterion measures and demographic or professional characteristics were found amongst

this subsample.

Investigative Abilities, General Reasoning, Job-related Beliefs, and Personality Characteristics

General reasoning skills. It was hypothesized that investigators' general reasoning skills would be related to their investigative abilities, such that better performance across the reasoning problems would be associated with higher levels of critical thinking and stronger disagreement with the case intervention. As predicted, numbers of correct reasoning responses were significantly correlated with critical-thinking index scores, $r(87) = .22$, $p < .05$, and with levels of disagreement/agreement, $r(89) = -.21$, $p < .05$.

Two series of four independent-groups t tests were calculated to examine more specifically which of the four dichotomously coded reasoning items accounted for these two significant relations. The first series of tests revealed only one significant difference in levels of disagreement/agreement: the group of 27 investigators who correctly answered the deductive reasoning question (i.e., question 4 in section D) were significantly less in agreement with Chris's intervention ($M = 4.02$, $SD = 2.02$) than were the 62 investigators who answered incorrectly ($M = 2.85$, $SD = 1.92$), $t(87) = 2.55$, $p = .01$; t values for the remaining three tests were less than 1. The second series of tests again revealed that the group who correctly solved the deductive reasoning problem had significantly higher critical-thinking scores ($M = 2.67$, $SD = .68$) than those who gave incorrect

responses ($M = 2.29$, $SD = .86$), $t(87) = -2.02$, $p < .05$; t values for the remaining three tests were nonsignificant.

Job-related beliefs and activities. It was hypothesized that the higher the percentage of reported cases of child abuse that investigators' believed to be true in general, the less critical they would be, and the more inclined they would be to agree with the case intervention. As expected, a significant inverse relation between the percentage of cases that investigators believed to be true in general (%TRUE) and their levels of critical thinking was found, $r(77) = -.22$, $p = .05$. However, the relation between %TRUE and levels of disagreement/agreement with the intervention was nonsignificant, $r(79) = .16$, $p > .15$.

The hypothesis that perceived responsibility for uncovering the truth in reported cases of child abuse would be directly related to levels of disagreement with the case intervention was not borne out, $r = .02$. Nor was perceived investigative responsibility related to levels of critical thinking, $r = -.02$. However, as predicted, perceived responsibility for ensuring a child's safety in a reported case of child abuse was directly related to levels of agreement with the intervention to remove the child, $r(88) = .21$, $p = .05$. Perceived responsibility for children's safety, however, was not significantly related to levels of critical thinking, $r = -.08$.

As predicted, investigators who were high in decisional uncertainty (i.e., who were [a] uncertain about the

completeness of the information they obtained in investigating a case, [b] uncertain about how such information should best be interpreted, and [c] uncertain, in retrospect, about their final decisions and actions in handling some of such cases) were less inclined to agree with the case intervention, $r(86) = -.18$, $p < .05$, one tailed, and were more critical in their thoughts about the case, $r(85) = .21$, $p = .05$.

A significant correlation between the extent to which investigators reported using the same assessment procedure across all the cases they have investigated (see item C-2 in Appendix A) and their levels of disagreement/agreement was found, $r(88) = .20$, $p = .05$. That is, the more that investigators reported using the same assessment procedure, the less likely they were to agree with the case intervention.

Investigators' self-rated investigative abilities were not significantly related to either their levels of disagreement/agreement with the case intervention, $r(88) = .04$, nor with their levels of critical thinking, $r(87) = -.13$, $p > .20$. Similarly, investigators' beliefs about how useful it is to form an hypothesis about the nature of case early on in its investigation were not significantly related to their levels of disagreement/agreement with the case intervention, $r(88) = .14$, $p > .15$, or to their levels of critical thinking, $r = -.03$.

Personality characteristics. The level of investigators' disagreement/agreement with the case intervention was not significantly correlated with attributional complexity, r

= .03, openness, $r = -.05$, agreeableness, $r = -.16$, $p > .10$, or conscientiousness, $r(89) = .18$, $p = .09$. Because it was hypothesized that the relation between conscientiousness and investigative decision-making might be moderated by investigators' professions, Pearson correlations were calculated separately for the subsamples of social-work investigators and police investigators. As expected, the correlation between social-work investigators' levels of conscientiousness and their levels of disagreement/agreement was positive and significant, $r(70) = .29$, $p = .01$. Although this correlation was not significant for the small sample of police, it was in the expected direction, $r(15) = -.15$, $p > .25$.

Investigators' levels of critical thinking were not significantly correlated with attributional complexity, $r = .04$, openness, $r = .12$, or conscientiousness, $r(87) = -.20$, $p = .06$. This last correlation was in the same direction, and of similar magnitude, for social-work investigators, $r(69) = -.23$, $p = .06$, and police investigators, $r(15) = -.22$, $p > .15$. Investigators' levels of agreeableness were directly related to their levels of critical thinking, $r(86) = .26$, $p = .01$,

Predicting Investigators' Critical-Thinking Skills and Decisions

A stepwise multiple regression approach was used to calculate the most parsimonious predictive model of investigators' levels of disagreement/agreement with the case

intervention.⁶ The four significant zero-order correlates of critical thinking (i.e., critical-thinking index scores [CRIT], responses to the deductive reasoning item [REASON],⁷ perceived levels of responsibility for ensuring children's safety [RESP], levels of decisional uncertainty [UNCERTAIN], and the degree to which investigators approach cases in a standardized manner [STANDARD]) were entered as potential predictors. As can be seen in Table 2, CRIT and RESP were retained as significant predictors of levels of disagreement/agreement. This model, which, after adjustments, accounted for 48% of the variability in investigators' decisions regarding the appropriateness of the case intervention, was highly significant, $F(2, 84) = 41.33$, $p < .0001$. Tolerances for CRIT and RESP were .99 ($ps < .05$), and no outliers (i.e., residuals greater than 3 in absolute value) were detected. All assumptions were met.

Next, a stepwise multiple regression was calculated to predict investigators' levels of critical thinking. Once again, significant zero-order correlates were entered as potential predictors (i.e., sex, REASON, %TRUE, and levels of agreeableness [AGREE]). As can be seen in Table 3, sex, %TRUE, and AGREE were retained as significant predictors of critical thinking. This model, which, after adjustment, accounted for 17% of the variability in investigators' critical-thinking index scores, was highly significant, $F(2, 74) = 6.33$, $p < .001$. Tolerances for each of the retained predictors were all above .96 ($ps < .05$), and no outliers were

detected. Assumptions of normality and independence of error were met, however, a scatterplot of standardized residuals plotted against standardized predicted scores revealed an inverse linear pattern. Thus, caution must be taken in interpreting this model as the assumptions of linearity and homogeneity of variance are suspect.

Discussion

It is striking that, given the spotty evidence presented in the scenario, almost 40% of investigators were willing to go along with the removal of the child, and less than 20% strongly disagreed with the intervention. As well, investigators were no more skeptical in their decision making and no more critical in their underlying reasoning about the case than were other professionals who were not responsible for investigating or assessing reported cases of child abuse. Yet, in spite of these facts, investigators perceived themselves as more responsible for getting to the truth in such cases, and thought of themselves as better investigators, than did non-investigators.

How are these findings to be accounted for? The fact that there were no significant differences in mean levels of disagreement/agreement across the four geographic subsamples surveyed, which covered three, distinct investigative systems (viz., Ontario, British Columbia, and New York state; J. C. Yuille, personal communication, September 22, 1992) suggests that investigators' responses to the case question prompting them to indicate their stance on the intervention were quite

stable. And, more generally, while it is true that a number of other systems for investigating reported cases of child abuse exist, it certainly is fair to say that the generalizability of findings obtained in this study did not suffer from the use of an unrepresentative sample of investigators.

One might argue that the higher-than-expected percentage of investigators in agreement with the case intervention and the null difference between investigators and non-investigators are artifacts of the criterion measure employed in the study. Unlike in actual case investigations, investigators' decisions in case scenarios have no dire consequences for others, and investigators basically are unaccountable for their decisions. This difference, however, does not necessarily, or even probably, mean that investigators would make better decisions in more accountable situations. Accountability pressures are known to yield inconsistent effects upon reasoning, sometimes making people process information more rigorously, but other times making them more defensive and even less critical in their judgments (Tetlock, 1983; Tetlock, Skitka, & Boettger, 1989). Though difficult to ascertain, it also does not appear that respondents simply didn't try to consider the scenario carefully. All but three respondents both answered the question concerning their disagreement/agreement with the intervention and provided open-ended justifications for their decisions.

That a single item was employed as a criterion measure of investigative decision-making might be a more significant problem. If respondents were asked to make their own decisions about the case, if forced to make a yes/no decision regarding an intervention, or if asked to describe, in an open-ended manner, what they thought would be an appropriate course of action, rather than disagreeing/agreeing with Chris's decision, a different pattern of results might have emerged. Although this uncertainty only can be dispelled by future research employing and comparing multiple criterion measures, there doesn't appear to be any clear reason to expect that a different pattern of findings would emerge.

Despite these possible drawbacks, the criterion measure employed in this study has a number of positive features. First, the use of a scenario of a reported case of child abuse allowed criterion data to be collected that could be unambiguously compared across groups of interest. Second, the hypothetical scenario allowed for precise control over the nature of the "case." In this study, this meant that "evidence" could be constructed so that its stereotypical features might prime respondents to believe that abuse was occurring, and yet, at the same time, actually constitute very weak evidence in support of the allegations. Third, and perhaps because of the control gained through constructing the case, a nice even distribution across the response options was obtained using this measure. This finding ruled out the possibility that the null differences in decision-making

abilities between investigators and other professionals merely were attributable to a restricted range.

In short, it does not appear likely that either the higher-than-expected percentage of respondents in agreement with the intervention or the null differences between investigators and non-investigators merely are artifacts of measurement. Nor can it be reasonably argued that investigators merely were unexperienced since this group had medians of 7 years of professional experience and 4 years of experience specifically dealing with child abuse cases.

One possible explanation for the null differences between investigators and other professionals is that the latter group might have had investigative experience despite the fact that their primary duties presently did not include investigation. Social-work and police administrators, for example, are likely to have been primarily responsible for investigations at earlier points in their careers, and still might be involved in investigative casework, albeit less directly. Similarly, others who report that their primary duties include child protection or treatment are also likely to be, or have been, involved more or less directly in investigative work (J. C. Yuille, personal conversation, September 9, 1992). Thus, it is plausible that the distinction between investigators and non-investigators in this study simply might not be that meaningful. The validity of this interpretation, however, cannot be verified given the lack of more detailed information on respondents' professional histories. Therefore, it is also

possible that the distinction between investigators and non-investigators is valid, and that the null differences between these groups, in fact, are representative of these samples. In future research examining such differences, it would be important to collect data on both respondents' present- and past duties.

Predicting Investigative Decisions

It has been suggested recently by a number of researchers (e.g., Coolbear, 1992; Klajner-Diamond et al., 1987; Wakefield & Underwager, 1988; Wehrspann et al., 1987; J. C. Yuille, personal conversation, May 1992) that the difficulty that investigators have in effectively assessing the credibility of allegations in reported cases of child abuse might stem from poor critical-thinking skills. Examples are selectively attending to information that confirms one's own hypothesis about a case, overlooking alternative explanations for information presented in a case, making unwarranted assumptions that might bias one's interpretation of case evidence, etcetera. Findings from this study support these researchers' contentions. By far, the single best predictor of investigators' decisions to disagree/agree with the case intervention was their levels of critical thinking, accounting for 46% of the variability in investigators' decisions. And, investigators whose statements reflected only critical thoughts, on the whole, were quite skeptical of Chris's intervention.

Investigators who provided a mix of critical and

uncritical statements were both significantly less skeptical about the intervention than those providing only critical comments and significantly more skeptical about the intervention than those providing only uncritical comments. These findings are interesting because they suggest that the beneficial effects of critical thought on investigative decision making are additive, and they support the notion that critical thinking is a highly multifaceted endeavor consisting of a number of finer skills, dispositions, and abilities (Ennis, 1981). Indeed, the aspects of critical thought that were tapped in this study overlap a subset of the features of critical thinking outlined by Ennis (1981).

It could be argued that the magnitude of the relation between levels of critical thinking and the extent of disagreement/agreement is inflated, since both indices are based upon the same case scenario. In fact, it is likely that the magnitude of this relation would be lower than that obtained in this study if a separate critical-thinking test were employed. This is to be expected for two reasons. First, the present coding scheme was designed a posteriori, in an attempt to cover fairly exhaustively the content of respondents' statements. This resulted in a coding scheme consisting of two major categories that were sufficiently broad enough to incorporate almost all of the statements that were made, and yet the criteria for inclusion in these categories were elaborated in enough detail, enabling a high percentage of agreement between coders to be reached. This

"best-fit" approach necessarily increases the magnitude of the observed relation, since it focuses the definition of critical thought on only those aspects of the construct that were explicitly found to vary in respondents' statements. Second, the approach taken in this study avoids the potentially weakening effects of domain independence. That is, even if the same features of critical thinking were tapped in a separate test, it is likely that the relation between test scores and levels of disagreement/agreement with the case intervention would be lower than that obtained in this study, since the aspects of critical thought triggered in the course of evaluating Chris's intervention (e.g., need for more information) might not be triggered in other contexts.

What are the interpretive implications of this measurement approach? The consequences of focusing upon those critical-thinking skills that are found to be relevant to investigative decision-making are not necessarily detrimental to either the prediction of investigative decisions or to the understanding of their underlying cognitive components. Insofar as a subset of critical-thinking features are consistently found to be related to investigative decision-making, and these features can be measured independent of ultimate criteria (Wiggins, 1973), then it should make little difference for the purpose of prediction if this pool of features exhausts a larger definition of critical thought. Indeed, if a more exact subset of features could be isolated that also are associated consistently with investigative

decision-making, then the prediction of investigative decisions should only be more effective. Similarly, our understanding of the cognitive underpinnings of investigative decision-making should only be strengthened by knowing the specific aspects of critical thought that are prerequisites for making effective decisions. Presently, it is premature to accept or reject the practical significance of the observed relation. Future empirical research is needed to adequately clarify whether the obtained findings are stable, or whether the approach taken disallows for sufficient generalizability even within the investigative domain.

The question of why some investigators were prone to uncritically accept the information presented in the case as evidence of abuse or neglect also must be raised. As mentioned above, the evidence presented in the case scenario was designed to prime investigators to believe that abuse and neglect were occurring, despite any clear evidence to support such allegations. For example, the description of the case stated that "the address that was given is located in a very poor area of the city predominantly populated by Blacks." Although it was not actually stated that the child was from a poor family, it is likely that a number of investigators inferred that this was the case. Zellman (1992) has found that, in most instances, professionals who work with children (e.g., clinical psychologists, social workers, child psychiatrists, etc.) are more likely to report cases, and label them as serious and abusive, when the family's

socioeconomic status (SES) is low. Zellman (1992) also found that incidents alleged to have happened to younger children were rated as relatively more deserving of a report. In the hypothetical case, the child was approximately 8 or 9 years old. Consistent with Zellman's findings, even though the developmental level of the child was unknown, listing the child's age as a reason for supporting Chris's intervention was highly frequent.

While certain case characteristics rightly should alert an investigator to the need for further investigation or perhaps removal of a child (e.g., a history of previous abuse, the severity and nature of injuries), the appropriate weight of other case characteristics, such as the child's age, on investigative judgments are less clear cut. While it is clear that an infant should not be left alone, it is less clear, in the absence of information on a child's developmental level, how problematic it might be to leave an 8- or 9-year-old alone (as is the case in the scenario). Still, the influence of other case characteristics, such as family SES and race, are more problematic since they are likely to reflect the introduction of personal biases into the decision-making process. In psychological terms, an investigator who is influenced by the SES of the alleged perpetrator is at least partly relying on a representativeness heuristic (Kahneman & Tversky, 1972) to arrive at an investigative decision concerning the appropriate course of action to be taken. And, it is well-known that decisions based upon intuitive

heuristics often contradict those arrived at through the use of correct base-rate information and/or more formal reasoning principles.

Even when investigators are critical, and do realize that information substantiating allegations of abuse or neglect is weak or even nonexistent, their abilities to resist heavily weighting features of a case that evoke stereotypic images of abuse or neglect are not guaranteed. One social worker's justification for her slight agreement with Chris's intervention illustrates this point:

All circumstantial evidence--bruises, scrapes--kid could have gotten them from play--some 8-9 year olds are left alone while parents work--maybe a parent or caretaker out shopping for food. But total situation all added up looks neglectful--by taking child into care, may bring in some services that could improve situation and enable child to return--at least get situation checked out more thoroughly.

Responses such as these also provided examples illustrating why critical statements are not always associated with disagreement with the case scenario.

Individual responses also suggest that the benefits of critical thinking on investigative decision-making are additive. It seems unlikely that the social worker just quoted would have arrived at the same decision if she also had considered the macroscopic factors that might be contributing to the reported situation, as did this social worker, who

strongly disagreed with the intervention:

It appears to me that this 8 or 9 year old child could provide a lot more information to an investigator on which to base a decision than essentially two questions. I do assume that the community standards of care are low here and this social worker may well apprehend the neighbourhood if Chris did drop-in visits down the street.

Investigators' levels of perceived responsibility for ensuring a child's safety in a reported case of child abuse (RESP) also significantly predicted their investigative decisions, such that those with relatively high levels of RESP were more inclined to accept Chris's intervention. Though the strength of this predictor is very modest in contrast to that of critical thinking, it is an interesting finding since it suggests that when decisions based on child-safety concerns are at odds with decisions based on the strength of the facts in a case (as in the present case), some investigators nevertheless will choose to intervene by removing the child.

The fact that RESP is not related to levels of critical thinking ($r = -.08$) also is interesting, since it suggests that the willingness to endorse the premature removal of the child by investigators with relatively high levels of RESP is not based upon a lack of critical thinking. This finding suggests that these investigators are as able as others who think of themselves as relatively less responsible for child safety to realize when there is little evidence in a case that supports allegations of child abuse, but they feel compelled

not to take any chances.

We know that social-work investigators reported significantly higher levels of perceived responsibility for ensuring child safety than did police investigators. We also know that police investigators perceived themselves as significantly more responsible for discovering the truth in a reported case of child abuse than did social-work investigators. One interesting question, then, is whether investigators with relatively high levels of RESP generally thought of themselves as being less responsible for discovering the truth in reported cases of child abuse? In fact, no relation between RESP and investigators' perceived responsibility for finding the truth was observed, $r = -.02$, suggesting that those who perceive themselves as relatively highly responsible for ensuring children's safety generally do not feel more or less responsible for finding out the truth in reported cases of child abuse. Moreover, that investigators' perceived responsibility for discovering the truth in reported cases of child abuse was not significantly related to their disagreement/agreement with the intervention, or to their levels of critical thinking, suggests that merely accepting responsibility for an investigative task does not ensure that information will be processed vigilantly, or that appropriate investigative decisions will actually be made.

Predicting Critical Thinking

Investigators' levels of critical thinking were, by far, the best predictors of their decisions concerning the case

intervention. Thus, it is important to examine, in turn, the factors that significantly predict critical thinking. Three variables--investigators' sex, their levels of hostility, and the percentage of reported cases of child abuse that they thought were true, in general--emerged as significant predictors, accounting for 17% of the variability in their levels of critical thinking. Roughly 10% of this variability was explained by investigators' sex. Female investigators' statements reflected significantly higher levels of critical thinking than did male investigators' statements. A closer examination of this relation revealed that 90% of females versus only 58% of males provided critical statements, $\chi^2(1, N = 89) = 12.00, p = .001$, and only 29% of females compared to 55% of males provided statements indicative of uncritical thought, $\chi^2(1, N = 89) = 5.58, p < .05$.

Interestingly, although female investigators were significantly more critical than male investigators, females' self-ratings of their investigative abilities ($M = 5.10, SD = .89$) were significantly lower than that of male investigators ($M = 5.65, SD = .76$), $t(88) = 2.91, p < .005$. As well, female investigators had significantly less professional experience ($M = 7.48, SD = 6.03$) than male investigators ($M = 11.00, SD = 7.98$), $t(89) = -2.35, p < .05$.

As hypothesized, the percentage of reported cases of child abuse that investigators thought were true in general (%TRUE) also predicted their levels of critical thinking, such that higher estimates of %TRUE were associated with lower

levels of critical thought. The interpretation of this finding, however, is complicated by the fact that there are two types of reports that investigators might have considered when providing their estimates of %TRUE (J. C. Yuille, personal conversation, September 9, 1992). That is, of the total number of reports that are received by social-work agencies, roughly 50% are followed up initially by social workers. The remaining 50% include reports that do not seem potentially serious enough to require further investigation. Of the reports that are followed up, roughly 90% are found to involve true allegations of child abuse (Jones & McGraw, 1987). Social workers, who receive the initial reports, and who constituted the vast majority of the sample of investigators in this study, might have based their estimates on either the total number of reports that are received or on the fraction of those reports that were actually followed up with investigations. Police, on the other hand, generally are informed of only the smaller proportion of cases that are thought to warrant further investigation, and would most probably have based their estimates of %TRUE on this restricted number of cases. Unfortunately, this unforeseen ambiguity makes the relation between %TRUE and critical thinking difficult to interpret.

Compared with police investigators, social-work investigators had significantly lower estimates of both %TRUE and the percentage of cases that they thought were true in their own investigative experiences. This finding supports

the notion that a substantial proportion of social workers, but not of police, included the 50% of cases (i.e., roughly) that do not receive further investigations in their reported estimates. Because of the uncertainty concerning social-work investigators' interpretations of the question prompting their estimates of %TRUE, it is difficult to assess the accuracy of these figures. Police investigators, however, who are likely to unambiguously base their estimates on the restricted number of cases (of which roughly 90% are found to be true), generally did have quite accurate estimates of %TRUE.

A third predictor of critical thinking was investigators' levels of agreeableness. Counter to expectations, investigators who were relatively more agreeable were relatively more critical in their statements concerning the case intervention. In an attempt to shed light on this relation, correlations between levels of critical thinking and each of the agreeableness scale items were calculated. Findings revealed that the more investigators thought of themselves as cynical and skeptical of others' intentions (item B-2) and as hard-headed and tough minded in their attitudes (item B-15) the less critical they were, $r(86) = -.26$, $p < .05$ and $r(87) = -.23$, $p < .05$ for items B-2 and B-15, respectively.

In retrospect, the preceding findings are not that surprising. As mentioned earlier, the short scale employed in this study tapped the arrogant-calculating (BC) octant of the interpersonal circumplex (Wiggins, 1982). As Wiggins

(personal communication, September 22, 1992) pointed out, individuals who lie within this region of the interpersonal circle are wary of others' intentions and view the world as an uncertain and dangerous place. In support of this notion, investigators levels of uncertainty were, in fact, inversely related to their levels of agreeableness, $r(85) = .31$, $p < .01$. Contrary to the healthy skepticism I had originally thought might be associated with a BC interpersonal style, it actually is more likely that investigators with such a disposition would be prone to cynically distrust others regardless of the sufficiency of the case evidence presented to them. Investigators with a cold and calculating disposition might then be relatively more inclined to prematurely conclude that allegations of abuse or neglect were true since they generally assume that the world is fundamentally a rotten place and that their job is to "clean it up." Given the fact that only four items were used to tap agreeableness in this study, it might be worthwhile to examine the relation between agreeableness and critical thinking using a larger pool of such items in future research.

Additional Relations

Although investigators' general reasoning skills did not significantly predict either their critical-thinking levels or their decisions concerning the case intervention, it was found that the number of reasoning items that investigators correctly answered was directly related to their levels of critical thinking and to their levels of disagreement with the

premature intervention. More importantly, it was found that both of these relations were primarily accounted for by the single deductive reasoning question. That is, neither investigators' levels of critical thinking nor their decisions regarding Chris's intervention significantly differed as a function of their responses to any of the three inductive reasoning questions.

Despite the fact that the deductive reasoning item did not emerge as a significant predictor in the analyses reported, it is interesting to note that if levels of critical thinking were removed from the regression analysis, then this single item would emerge as the only significant predictor of investigators' levels of disagreement/agreement. This is not surprising given the significantly higher levels of critical thinking found amongst investigators who correctly answered this reasoning question, and the expected overlap between this item (which was selected from Ennis & Millman's [1982] critical-thinking test) and critical-thinking index scores.

Deduction involves the ability to analyze the syntactic relations between pieces of information, and to draw valid conclusions from such information. It might be the case that deductive skills are particularly important for investigators of reported cases of child abuse, who must organize and draw conclusions from case evidence. Given that only a single deductive reasoning item was employed in this study, it might be worthwhile to examine in future research how the number of correct responses summed across an entire scale of deductive

reasoning items would predict investigative decision-making.

As predicted, investigators' levels of decisional uncertainty also were directly related to their levels of critical thinking and to their levels of disagreement with the case intervention. Since decisional uncertainty reflects the fact that investigators are aware that information is often limited, that multiple interpretations of available case evidence often exist, and that investigative decisions, in retrospect, are sometimes premature, it is not surprising to find that investigators displaying relatively high levels of decisional uncertainty would be more likely to be aware of the lack of adequate information and the multiple (and more probable) interpretations of information presented in the case scenario. Investigators who felt relatively more uncertain perceived themselves as being more responsible for uncovering the truth in reported cases of child abuse, $r(85) = .23$, $p < .05$.

It was expected that investigators who thought that it was useful to develop an hypothesis early on in the investigation of reported cases of child abuse would be less critical of the intervention made in the case scenario. It was reasoned that investigators who were willing to develop an hypothesis early on would be relatively more likely to selectively attend to information that supported their hunches, and to end their information searches prematurely, since a "more probable" interpretation already was arrived at. Since the features in the case scenario are designed to appear

suggestive of abuse and neglect, it was anticipated that such investigators would be more likely to be primed by this information and accept it as fact. It is difficult to determine why this relation was not obtained. Only a single item was used to measure this tendency. And, this item also was ambiguous in that it did not clarify what "an hypothesis" was directed at.

Neither attributional complexity nor openness were related to investigators' levels of critical thinking or their decisions concerning Chris's intervention. As was suggested earlier, attributionally complex investigators might have generated more unwarranted attributions, in addition to generating reasonable ones. While the exploratory analyses of relations between personality constructs and investigative abilities generally were unrevealing, one interesting finding did emerge. The prediction that the relation between investigators' levels of conscientiousness and the nature of their investigative decisions would be moderated by their professions (viz., social workers or police) was borne out. As expected, conscientiousness was directly related to social-work investigators' levels of agreement with the intervention, but inversely related to that of police investigators. The isolation of successful moderator variables has been infrequent (Wiggins, 1973), and the present findings lend further support to the notion that, when found, the best moderators are often dichotomous, categorical variables (J. S. Wiggins, personal communication, September 22, 1992).

Substantively, these findings support the notion that investigators working under different agencies perceive themselves as having different investigative objectives (Hunter et al., 1990). These findings also suggest that environmental factors, such as agency mandates, which are likely to influence investigators' perceptions of what they are specifically responsible for achieving, also interact with investigators' levels of conscientiousness to influence the quality of investigative decision-making. Social-work investigators, in fact, did have significantly higher levels of perceived responsibility for ensuring child safety than did police investigators. And, police investigators had significantly higher levels of perceived responsibility for getting to the truth in reported cases of child abuse than did social-work investigators. Social-work investigators might be more concerned about their rate of false negatives than their rate of false positives, since the onus is on them to ensure that potentially victimized children, in fact, are safe.

Finally, it was found that police investigators' professional experience was significantly related to their investigative decision-making, such that investigators with relatively more professional experience were more likely to support the case intervention. Though seemingly counter-intuitive, this finding actually is not that surprising. Dent (1982), for example, has found that experienced police investigators elicited very inaccurate descriptive information in interviewing children. Yuille (personal conversation,

September 9, 1992) has suggested two explanations for the finding obtained in this study. First, police training in investigative procedures has improved through the years. Second, police are routinely switched from one duty to another over the course of their careers. However, police officers who are found to be ineffective at their jobs are actually less likely to be routinely switched from one duty to another. Thus, longer experience in one area might be a sign that an officer is not meeting expected standards. Though both of these facts are consistent with the obtained finding, one must also be cautious in interpreting this result, given the small sample size on which it is based.

Future Research

The study of investigators' critical-thinking skills and decision-making abilities represents new territory to be explored. Social workers, police, and other professionals, such as the respondents examined in this study, represent highly important samples of individuals who make crucial decisions in their everyday duties. In addition to the potential this research offers for extending beyond typical university samples what is by now a tradition of cognitive social psychological research on human information processing, it has important applied ramifications for the investigation of child abuse. At this very early stage, it is important not only to test hypotheses and formulate theories, but to examine methods as well.

In the present study, time constraints did not permit the

latter of these objectives to be pursued. Contrarily, an attempt was made to survey a broad range of psychological information, most often with short-scales or single items. Although this approach yielded useful information, providing a first stab at what factors should be pursued further, it did not allow for a number of methodological issues to be cleared. Justifications for the measures employed in this study have been provided. Nevertheless, it is realized that these techniques must be empirically validated in this domain of research. As I have already suggested, it would be interesting to know if comparable results would be obtained using even simple variations of the presently employed criterion measure. Would investigators think about their decisions more carefully if they had to specify their own decision about the case, rather than commenting on another investigators' intervention? Would it make an important difference if respondents had to indicate either a "yes" or a "no," regarding a decision, or choose from an array of decisional options, rather than checking a scale?

In addition to the preceding variations, it would be important to examine the effects of systematically manipulating the contents of case scenarios on investigative decision-making (as Zellman, 1992, has begun to do), and also on critical thinking. In this study, there was not enough information to make a clear indication that abuse or neglect was or was not occurring. But, given the facts presented, removing the child surely was an inappropriate intervention.

It would be interesting to examine how critical thinking relates to decision making in other investigative situations. What would happen, for instance, if the situation were turned around and stereotypic information would suggest that abuse (or neglect) was not occurring (e.g., accusations of child abuse are launched against the child's father who is a prominent, wealthy member of society in the midst of a political campaign [thus, there are motives for making up a smear campaign]), but subtle pieces of evidence would suggest that allegations, in fact, were true? The case scenario in this study provided little information to respondents. It would be important to examine investigators' decisions and critical-thinking skills when a great deal of detailed information is provided. Under such circumstances, requests for additional information would probably diminish, and other analytical skills might play a more important role in the prediction of investigators' decisions. Although time constraints did not permit multiple case scenarios to be used in the present study, it would be interesting to examine how open-ended responses from one case scenario that were coded for levels of critical thinking would predict investigative decisions concerning other case scenarios.

A comparison of alternative methods for measuring critical thinking should also be pursued. Coding open-ended responses allows researchers to examine the types of critical and uncritical thoughts that are spontaneously generated by front-line workers making decisions in investigative

situations. While this type of data is revealing of the aspects of critical thought that are actually triggered in investigative situations, the fact that all respondents do not focus on the same features of the case, and are often critical or uncritical in different respects, makes it difficult to examine how specific critical-thinking skills predict decision-making abilities. Thus, it might be useful to examine how scores from a critical-thinking test that uniformly probes a number of specific skills across an entire sample would predict decision-making abilities in an investigative context. We know that a single item from Ennis and Millman's (1982) critical-thinking test was significantly correlated with investigative decision-making in this study. It might be profitable to examine how investigators' responses to the subsections of this test, each of which taps distinct skills, would predict their investigative decision-making abilities.

Conclusion

Despite the increasing number of reported cases of child abuse, and the fact that approximately 90% of these cases are true, there was no indication in this study that police or social workers who are responsible for investigating such cases are more able to make effective decisions concerning appropriate actions than other professionals who are not responsible for carrying out such investigations. The fact that roughly 40% of investigators agreed to some extent with the intervention is disturbing given that the case scenario

had virtually no compelling evidence to support allegations of abuse or neglect. Clearly, the findings in this study support the notion that critical-thinking skills are essential prerequisites for effective decision-making. It remains to be seen, however, whether a consistent set of these skills can be adequately defined and measured. If this objective can be achieved, it will also be important to examine how investigators' critical-thinking skills could be improved through training. These questions could be profitably addressed in future research.

References

- Badgley, R. (1984). Report of the federal commission on sexual offences against children and youths. Ottawa: Federal Department of Justice and Health and Welfare.
- Browne, A., & Finkelhor, D. (1986). Impact of child sexual abuse: A review of the research. Psychological Bulletin, 99, 66-77.
- Costa, P. T., Jr., & McCrae, R. R. (1985). The NEO Personality Inventory manual. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., Jr., & McCrae, R. R. (1989). The NEO/FFI manual supplement. Odessa, FL: Psychological Assessment Resources.
- Dent, H. R. (1982). The effects of interviewing strategies on the results of interviews with child witnesses. In A. Trankell (Ed.), Reconstructing the past. Stockholm: Norstedt & Soners.
- Ennis, R. H. (1981). Rational thinking and educational practice. In J. F. Soltis (Ed.), Philosophy and education (Eightieth yearbook of the National Society for the Study of Education, Part 1). Chicago: NSSE.
- Ennis, R. H., & Millman, J. (1982). Cornell critical thinking test, level 2. Champaign, IL: Illinois Critical Thinking Project, University of Illinois.
- Finkelhor, D., & Hotaling, G. T. (1984). Sexual abuse in the national incidence study of child abuse and neglect: An appraisal. Child Abuse & Neglect, 8, 23-33.
- Fletcher, G. J. O., Danilovics, P., Fernandez, G., Peterson, D., & Reeder, G. D. (1986). Attributional complexity: An individual differences measure. Journal of Personality and Social Psychology, 51, 875-884.
- Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. Cognitive Psychology, 18, 253-292.
- Ghiselli, E. E. (1965). Theory of psychological measurement. McGraw-Hill.
- Goodwin, J., Sahd, D., & Rada, R. (1978). Incest hoax: False allegations, false denials. Bulletin of the American Academy of Psychiatry and Law, 6, 269-276.
- Green, A. H. (1986). True and false allegations of sexual abuse in child custody disputes. Journal of the American

- Academy of Child Psychiatry, 25, 449-456.
- Hogan, R. (1985). Manual for the Hogan Personality Inventory. Minneapolis, MN: Nation Computer Systems.
- Howell, D. C. (1987). Statistical methods for psychology (2nd ed.). Boston, MA: PWS-Kent.
- Hughes, M., & Grieve, R. (1980). On asking children bizarre questions. In M. Donaldson, R. Grieve, & C. Pratt (Eds.), Early childhood development and education. Oxford: Basil Blackwell.
- Hunter, R., Yuille, J. C., & Harvey, W. (1990). A coordinated approach to interviewing in child sexual abuse investigations. Canada's Mental Health, 38, 14-18.
- Joffe, R., & Yuille, J. C. (1991). Assessing the credibility of a child's recollection. Unpublished manuscript, University of British Columbia, Vancouver, B.C.
- John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural languages and in questionnaires. In L. A. Pervin (Ed.), Handbook of personality: Theory and research. New York: Guilford.
- Jones, D. P., & Krugman, R. D. (1986). Can a three-year-old bear witness to her sexual assault and attempted murder? Child Abuse and Neglect, 10, 253-258.
- Jones, D. P. H., & McGraw, J. M. (1987). Reliable and fictitious accounts of sexual abuse to children. Journal of Interpersonal Violence, 2(1), 27-45.
- Jones, D. P. H., & Seig, A. (1988). Child sexual abuse in custody and visitations disputes. In E. B. Nicholson & J. Bulkley (Eds.), Sexual abuse allegations in custody and visitation cases: A resource book for judges and court personnel. Washington, DC: American Bar Association.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). Judgment under uncertainty: Heuristics and biases. New York: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. Cognitive Psychology, 3, 430-454.
- King, M. A., & Yuille, J. C. (1987). Suggestibility and the child witness. In S. J. Ceci, D. F. Ross, & M. P. Toglia (Eds.), Children's eyewitness memory. New York: Springer-Verlag.

- Klajner-Diamond, H., Wehrspann, W., & Steinhauer, P. (1987). Assessing the credibility of young children's allegations of sexual abuse: Clinical Issues. Canadian Journal of Psychiatry, 32, 610-614.
- Lehman, D. R., Lempert, R. O., & Nisbett, R. E. (1988). The effects of graduate training on reasoning: Formal discipline and thinking about everyday-life events. American Psychologist, 43, 431-442.
- McCrae, R. R., & Costa, P. T., Jr. (1989). The structure of interpersonal traits: Wiggins' circumplex and the five-factor model. Journal of Personality and Social Psychology, 56, 586-595.
- McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. In R. R. McCrae (Guest Ed.), Journal of Personality, (Special Issue), 60, No. 2.
- Melton, G. B., & Thompson, R. A. (1987). Getting out of a rut: Detours to less travelled paths in child witness research. In S. J. Ceci, D. F. Ross, & M. P. Toglia (Eds.), Children's eyewitness memory. New York: Springer-Verlag.
- Mill, D., Gray, T., & Mandel, D. R. (1992). Influence of research methods and statistics courses on everyday reasoning and belief in unsubstantiated phenomena. Manuscript under review.
- Nisbett, R. E., & Ross, L. (1980). Human inference: Strategies and shortcomings of social judgment. Englewood Cliffs, NJ: Prentice-Hall.
- Prigogine, I., & Stengers, I. (1984). Order out of chaos. New York: Bantam Books.
- Raskin, D. C., & Yuille, J. C. (1989). Problems in evaluating interviews of children in sexual abuse cases. In M. P. Toglia, S. J. Ceci, & D. F. Ross (Eds.), Adult perceptions of child testimony. New York: Springer-Verlag.
- Russell, D. E. H. (1983). The incidence and prevalence of intrafamilial and extrafamilial sexual abuse of female children. Child Abuse and Neglect, 7, 133-146.
- Steller, M., & Koehnken, G. (1989). Criteria-based statement analysis. In D. C. Raskin (Ed.), Psychological methods in criminal investigation and evidence. New York: Springer-Verlag.
- Tetlock, P. E. (1983). Accountability and the complexity of

- thought. Journal of Personality and Social Psychology, 45, 74-83.
- Tetlock, P. E., Skitka, L., & Boettger, R. (1989). Social and cognitive strategies for coping with accountability: Conformity, complexity, and bolstering. Journal of Personality and Social Psychology, 57, 632-640.
- Trapnell, P. D. (1991). [Factor analysis derived to measure attitudinal conformity]. Unpublished raw data.
- Trapnell, P. D. (1992a). Private self-consciousness: Reflection or Rumination? Unpublished manuscript, University of British Columbia, Vancouver, B.C.
- Trapnell, P. D. (1992b). Vocational interests and the facet structure of Factor V. Paper presented as part of the symposium Intellectance and openness in the "Big Five" personality structure (John F. Kihlstrom, Chair), at the Centennial Convention of the American Psychological Association, Washington, D. C.
- Trapnell, P. D., & Wiggins, J. S. (1990). Extension of the Interpersonal Adjective Scales to include the Big Five dimensions of personality. Journal of Personality and Social Psychology, 59, 781-790.
- Trapnell, P. D., & Wiggins, J. S. (1992). [Factor analysis of the Five Factor Inventory]. Unpublished raw data.
- Undeutsch, U. (1982). Statement reality analysis. In A. Trankell (Ed.), Reconstructing the past. Stockholm: Norstedt & Soners.
- Undeutsch, U. (1984). Courtroom evaluation of eyewitness testimony. International Review of Applied Psychology, 33, 51-67.
- Wakefield, H., & Underwager, R. (1988). Accusations of child sexual abuse. Springfield, IL: Charles C. Thomas.
- Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), New horizons in psychology (Vol. 1). Hammondsorth, Middlesex, England: Penguin.
- Wehrspann, W. H., Steinhauer, P. D., & Klajner-Diamond, H. (1987). Criteria and methodology for assessing credibility of sexual abuse allegation. Canadian Journal of Psychiatry, 32, 615-623.
- Wiggins, J. S. (1973). Personality and prediction: Principles of personality assessment. Reading, MA: Addison-Wesley.
- Wiggins, J. S. (1982). Circumplex models of interpersonal

- behavior in clinical psychology. In P. C. Kendall & J. N. Butcher (Eds.), Handbook of research methods in clinical psychology. New York: Wiley.
- Wiggins, J. S., & Trapnell, P. D. (in press). Personality structure: The return of the Big Five. In S. R. Briggs, R. Hogan, & W. J. Jones (Eds.), Handbook of personality psychology. Orlando, FL: Academic Press.
- Yuille, J. C. (1984). Research and teaching with police: A Canadian example. International Review of Applied Psychology, 33, 5-23.
- Yuille, J. C. (1988). The systematic assessment of children's testimony. Canadian Psychology, 29, 247-262.
- Yuille, J. C. (1992a). Investigative interviewing of children and the systematic assessment of their statements. Unpublished manuscript, University of British Columbia, Vancouver, B.C.
- Yuille, J. C. (1992b). Statement Validity Analysis: Interviewing and assessing children in sexual abuse cases. Unpublished manuscript, University of British Columbia, Vancouver B.C.
- Yuille, J. C., Hunter, R., Joffe, R., & Zaparniuk, J. (1992). Interviewing children in sexual abuse cases. Manuscript submitted for publication.
- Zargoza, M. S. (1987). Memory, suggestibility, and eyewitness testimony in children and adults. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), Children's eyewitness memory. New York: Springer-Verlag.
- Zellman, G. L. (1992). The impact of case characteristics on child abuse reporting decisions. Child Abuse & Neglect, 16, 57-74.

Footnotes

¹Frequencies sum to more than 143 because 54 respondents reported more than one primary duty. Only the first two duties reported were coded.

²The basis for these subscales is strictly theoretical since only one major factor was found by Fletcher et al. (1986).

³The decision to select the fifth and sixth highest loading items (items 2 and 15, respectively) over the second and fourth highest loading items was based on a preference for the wording of the former items.

⁴Because of the exploratory nature of this study, a decision was made to strictly enforce per-comparison $\alpha \leq .05$ as a limit (i.e., probabilities slightly higher than .05 were not reported as marginal), but not to adjust the error rate per experiment. This approach was adopted with a clear understanding that revealing findings would have to be examined more precisely in future research. All probabilities are two tailed, unless otherwise specified. One-tailed probabilities were reported only in situations where specific predictions were made, and where two-tailed probability levels were nonsignificant.

In some situations where t tests were employed, the assumption of homogeneity of variances was found to be violated (i.e., using the F test for heterogeneity of variances, where the larger of the two variance estimates, \underline{s}_L^2 , is divided by the smaller variance estimate, \underline{s}_S^2 , and where the

F statistic is distributed as the F distribution on $n_1 - 1$ and $n_2 - 1$ df). In cases where heterogeneity of variances was detected, t' was calculated based on the separate variance estimates. The Welch-Satterthwaite solution (Howell, 1987) was used to calculate df' , which was rounded to the nearest integer.

⁵Since 89 of the 91 investigators were either social workers ($n = 72$) or police ($n = 17$), only these two professions were considered.

⁶Missing data were treated by listwise deletion. The criterion for variable inclusion in forward analysis was set at $\alpha = .05$. The criterion for variable exclusion in backward analysis was set at $\alpha = .10$.

⁷Since the significance of the index of reasoning items correctly answered derives from this single item, only the latter was entered as a potential predictor.

Table 1

Summary of significant mean differences between social-work- and police investigators.

Variable	Social Work		Police		df	t
	M	SD	M	SD		
<u>Demographic/Professional</u>						
Prof. experience	7.13	5.88	14.29	6.72	87	-4.40***
Educational level	3.14	.43	1.64	.84	81	9.89***
<u>Beliefs/Actions</u>						
C-4: Remove child despite suspicions	2.67	1.49	1.88	1.17	87	2.02*
C-7: Responsibility for child's safety	6.17	1.25	5.53	1.33	87	1.87
C-8: Responsibility to find truth	5.42	1.16	6.65	.61	48	-6.09*** ^a
E-8b: % true cases in own experience	57.33	26.04	83.94	12.85	54	-5.88*** ^a
E-9: % true cases in general	70.27	21.78	85.41	10.79	54	-3.98*** ^a
<u>Personality</u>						
Attrib. Complexity	4.14	.38	3.71	.29	87	4.16***
Openness	3.85	.48	3.46	.38	86	3.03**
Agreeableness	3.65	.54	3.28	.54	86	2.53**

* $p < .05$, ** $p < .01$, *** $p < .001$; $p < .05$, one tailed.

^abased on t' and df' .

Table 2

Summary table of stepwise regression predicting levels of disagreement/agreement with the case intervention.

Step	Variable	B	SE B	β	\underline{R}	Adj. \underline{R}^2	\underline{F} change
<hr/>							
IN EQN							
<hr/>							
1	CRIT	-1.72	.20	-.69	.69	.46	75.20**
2	RESP	.26	.12	.16	.70	.48	4.43*
NOT IN EQN							
<hr/>							
	REASON			-.13			2.55
	UNCERTAIN			-.05			.32
	STANDARD			.06			.80
<hr/>							

* $p < .0001$, ** $p < .05$.

Table 3

Summary table of stepwise regression predicting critical-thinking index scores.

Step	Variable	B	SE B	β	R	Adj.R ²	Fchange
IN EQN							
1	SEX	-.54	.19	-.31	.31	.09	8.22**
2	%TRUE	-.01	.00	-.23	.39	.13	4.52*
3	AGREE	.35	.16	.24	.45	.17	5.16*
NOT IN EQN							
	REASON			.16			2.44

* $p < .05$, ** $p \leq .005$.

Appendix A
Diagnostic Survey

The purpose of this survey is to gather important information from front-line workers who deal with reported cases of child abuse. In order for quality information to be collected, it is important that you read each question carefully and answer as honestly as you can. Your responses are completely anonymous and are being gathered solely for research purposes. Thank you very much for participating.

A. Below is a hypothetical case scenario. Please read and consider it carefully. Then answer the following questions.

A social work agency receives an anonymous call on May 28th from a caller who states that he is aware of a case of child neglect and abuse. The caller gives the name of the family and the address.

Chris, a social worker with many years experience, responds to the call. The address that was given is located in a very poor area of the city that is predominantly populated by Blacks. At 4:15 p.m., the social worker knocks on the door. It is answered by a young child approximately 8 or 9 years old.

At first glance Chris sees that the child is dirty. There are bruises on both the child's knees and there is a scrape on the child's right elbow. Looking past the child, Chris views the residence. It seems quite messy. Upon closer inspection Chris finds that there is little food in the refrigerator.

Chris questions the child on the whereabouts of the parents, and the child's response is "I don't know." Chris also asks how long the child has been alone. The child responds, "All day." After considering the situation, Chris decides to take the child into care.

1. To what extent do you agree or disagree with the social worker's intervention? Circle the one number that best represents your answer.

**STRONGLY
DISAGREE**

**STRONGLY
AGREE**

1 2 3 4 5 6 7

2. List the features of the case that led to your decision in question 1?

B. For each of the following statements please indicate the response which best represents your opinion according to the following scale:

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Circle "1" if the statement is definitely false or you strongly disagree.

Circle "2" if the statement is mostly false or you disagree.

Circle "3" if the statement is about equally true or false, or if you cannot decide, or if you are neutral on the statement.

Circle "4" if the statement is mostly true or you agree.

Circle "5" if the statement is definitely true or you strongly agree.

- | | | | | | |
|--|---|---|---|---|---|
| 1. I really enjoy analyzing the reasons or causes for people's behavior. | 1 | 2 | 3 | 4 | 5 |
| 2. I tend to be cynical and skeptical of others' intentions. | 1 | 2 | 3 | 4 | 5 |
| 3. I believe that loyalty to one's ideals and principles is more important than "openmindedness." | 1 | 2 | 3 | 4 | 5 |
| 4. Once I have figured out a single cause for a person's behavior I usually don't go any further. | 1 | 2 | 3 | 4 | 5 |
| 5. I am a productive person who always gets the job done. | 1 | 2 | 3 | 4 | 5 |
| 6. I'm pretty set in my ways. | 1 | 2 | 3 | 4 | 5 |
| 7. I believe it is important to analyze and understand our own thinking processes. | 1 | 2 | 3 | 4 | 5 |
| 8. If I don't like people, I let them know it. | 1 | 2 | 3 | 4 | 5 |
| 9. Philosophical or abstract thinking doesn't appeal to me that much. | 1 | 2 | 3 | 4 | 5 |
| 10. I like to adopt the opinions of my friends. | 1 | 2 | 3 | 4 | 5 |
| 11. I have thought very little about my own family background and history in order to understand why I am the sort of person I am. | 1 | 2 | 3 | 4 | 5 |
| 12. I think a lot about the influence that I have on other people's behavior. | 1 | 2 | 3 | 4 | 5 |
| 13. I tend to take people's behavior at face value and not worry about the inner causes for their behavior (e.g., attitudes, beliefs, etc.). | 1 | 2 | 3 | 4 | 5 |
| 14. I strive for excellence in everything I do. | 1 | 2 | 3 | 4 | 5 |

- | | | | | | |
|--|---|---|---|---|---|
| 15. I'm hard-headed and tough minded in my attitudes. | 1 | 2 | 3 | 4 | 5 |
| 16. I have a lot of intellectual curiosity. | 1 | 2 | 3 | 4 | 5 |
| 17. I think a lot about the influence that society has on my behavior and personality. | 1 | 2 | 3 | 4 | 5 |
| 18. I love exploring my inner self. | 1 | 2 | 3 | 4 | 5 |
| 19. I never seem to be able to get organized. | 1 | 2 | 3 | 4 | 5 |
| 20. Some people think of me as cold and calculating. | 1 | 2 | 3 | 4 | 5 |
| 21. I am not really curious about human behavior. | 1 | 2 | 3 | 4 | 5 |
| 22. I prefer simple rather than complex explanations for people's behavior. | 1 | 2 | 3 | 4 | 5 |
| 23. When I analyze a person's behavior I often find the causes form a chain that goes back in time, sometimes for years. | 1 | 2 | 3 | 4 | 5 |
| 24. I give little thought to how my thinking works in the process of understanding or explaining people's behavior. | 1 | 2 | 3 | 4 | 5 |
| 25. I work hard to accomplish my goals. | 1 | 2 | 3 | 4 | 5 |
| 26. I think very little about the different ways that people influence each other. | 1 | 2 | 3 | 4 | 5 |
| 27. To understand a person's personality/behavior I have found it is important to know how that person's attitude, beliefs, and character traits fit together. | 1 | 2 | 3 | 4 | 5 |
| 28. I often enjoy playing with theories or abstract ideas. | 1 | 2 | 3 | 4 | 5 |
| 29. When I try to explain other people's behavior I concentrate on the person and don't worry too much about all the existing external factors that might be affecting them. | 1 | 2 | 3 | 4 | 5 |

C. Please indicate your response to each of the following questions by circling the appropriate number.

1. In assessing the credibility of a report of child abuse, how often must your decision be based upon what you consider limited evidence?

NEVER								QUITE OFTEN
	1	2	3	4	5	6	7	

2. To what extent do you evaluate each reported case of child abuse in a similar manner?

USUALLY USE SAME ASSESSMENT PROCEDURE FOR EACH CASE								ASSESSMENT PROCEDURE OFTEN DEPENDS UPON THE CASE
	1	2	3	4	5	6	7	

3. Given the consequences of your decision, what degree of certainty are you likely to reach before you think action must be taken in reported cases of child abuse?

HIGH UNCERTAINTY								HIGH CERTAINTY
	1	2	3	4	5	6	7	

4. Have you ever removed a child from a home, or laid a charge, despite suspicions that the child's statements were erroneous?

NO, NEVER								YES, QUITE OFTEN
	1	2	3	4	5	6	7	

5. In retrospect, have you ever thought that a decision you made about a reported case of child abuse was premature?

NO, NEVER								YES, QUITE OFTEN
	1	2	3	4	5	6	7	

6. Do you find it useful to develop an hypothesis about a reported case of child abuse early on in your investigation?

NO, NOT AT ALL USEFUL

YES, VERY USEFUL

	1	2	3	4	5	6	7	
--	---	---	---	---	---	---	---	--

7. In the course of dealing with the various aspects of a particular reported case of child abuse a number of different professionals are likely to become involved. To what extent do you think that it is your responsibility to ensure that the potentially abused child is safe?

PRIMARYLY OTHERS'
RESPONSIBILITY

PRIMARYLY MY
RESPONSIBILITY

	1	2	3	4	5	6	7	
--	---	---	---	---	---	---	---	--

8. Again, given that a number of other professionals are likely to become involved in a reported case of child abuse, to what extent do you think that it is your responsibility to question the truthfulness of the reported claims?

PRIMARYLY OTHERS'
RESPONSIBILITY

PRIMARYLY MY
RESPONSIBILITY

	1	2	3	4	5	6	7	
--	---	---	---	---	---	---	---	--

9. Do you ever find yourself left with more than one interpretation of the information provided to you in a reported case of child abuse?

NO, NEVER

YES, QUITE OFTEN

	1	2	3	4	5	6	7	
--	---	---	---	---	---	---	---	--

10. How would you rate your investigative abilities?

POOR

EXCELLENT

	1	2	3	4	5	6	7	
--	---	---	---	---	---	---	---	--

D. Please answer the following questions by circling the appropriate letter.

1. "New! Grapefruit tablets for the most effective diet ever. Watch pounds roll off! (Warning: Dieters must not restrict their food to the tablets alone, but must eat no more than three balanced meals a day.)"

Which of the following is the strongest criticism of the suggestion made in the advertisement above that the tablets will cause weight loss?

- A. Weight loss is normally achieved primarily by a combination of exercise, proper nutrition and dieting.
 - B. Although the advertisement says, "Watch pounds roll off!" it does not specify when weight loss will occur.
 - C. If the tablets are used as directed and weight loss occurs, the restriction of food intake to three balanced meals a day could be the cause.
 - D. The amount of weight lost on such a diet depends entirely on the amount the dieter is overweight at the beginning of the diet.
2. The city of Middleopolis has had an unpopular police chief for a year and a half. He is a political appointee who is a crony of the mayor, and who had little previous experience in police administration when appointed. The mayor has recently defended the chief in public, announcing that in the time since he took office, crime rates decreased by 12%.

Which of the following pieces of evidence would most deflate the mayor's claim that his chief is competent?

- A. The crime rates of the two cities closest to Middleopolis in location and size have decreased by 18% in the same period.
- B. An independent survey of the citizens of Middleopolis shows that 40% more crime is reported by respondents in the survey than is reported in police records.
- C. Common sense indicates that there is little a police chief can do to lower crime rates. These are mostly due to social and economic factors beyond the control of officials.
- D. The police chief has been discovered to have business contacts with people who are known to be involved in organized crime.

3. A talent scout attends a musical competition with the intention of observing the talent and skill of a particular violinist. In each of the first six performances, he repeatedly plays difficult passages with a fluency worthy of the best professional performers. However, in the final performance of the competition, as one of the two semi-finalists, the player stumbles over a key solo passage, stops playing, and tries again. The other semi-finalist performs her concerto flawlessly, and goes on to win the competition.

The scout reports that the player in question "has excellent skills, and should be recruited. He has a tendency to misplay under extreme pressure, but this will probably disappear with more experience and training."

The scout's report is:

- A. Probably accurate both in assessing the player's general level of ability and his tendency to misplay under pressure.
 - B. Probably accurate in assessing the player's general level of ability, but perhaps inaccurate in assessing his tendency to misplay under pressure.
 - C. Perhaps inaccurate in assessing the player's general level of ability, but probably accurate in assessing his tendency to misplay under pressure.
 - D. Probably inaccurate in assessing both the player's general level of ability and in assessing his tendency to misplay under pressure.
4. "The explanation for the misbehavior of Mr. Smith's children is a simple one. These children have been severely punished in the past. That's the trouble."

Which of the following is most probably the unstated assumption of the preceding statement?

- A. Children who have been severely punished in the past misbehave.
- B. Children who misbehave have been severely punished in the past.
- C. Children who misbehave probably have been severely punished in the past.
- D. Children who haven't been severely punished in the past behave properly.

E. Please answer each of the following questions. Remember that your responses are completely anonymous.

1. What is your age? _____ years.
2. What is your sex? _____ Female Male _____
3. What is your current level of education? Please specify the degree(s) or type of training you have received?

4. What is your profession? _____
5. How many years have you been employed in this profession?
_____ years.
6. How many years have you been involved specifically with child abuse cases?
_____ years.
7. What is/are your primary duty/duties in dealing with reported cases of child abuse?

- 8a. Estimate the number of reported child abuse cases that you have investigated.
I have investigated approximately _____ such cases.
- 8b. In what percentage of these cases did you conclude that, by and large, the report was substantiated?
_____ %
9. In general, what percentage of reported child abuse cases do you think are likely to be true?
_____ %
10. In what city is this workshop being held? _____

THANK YOU VERY MUCH FOR PARTICIPATING