

TURUN YLIOPISTON JULKAISUJA
ANNALES UNIVERSITATIS TURKUENSIS

SARJA – SER. B OSA – TOM. 319

HUMANIORA

**Expertise and decision making
among clinicians in investigations
of alleged child sexual abuse**

by

Katarina Finnilä-Tuohimaa

TURUN YLIOPISTO
Turku 2009

From the Department of Psychology, University of Turku, and the Department of Psychology, Åbo Akademi University, Turku, Finland

Supervised by:

Professor Pekka Niemi, *PhD*
Department of Psychology
University of Turku, Turku, Finland

and

Professor Pekka Santtila, *PhD*
Department of Psychology
Åbo Akademi University, Turku, Finland

Reviewed by

Professor Maggie Bruck, *PhD*
Division of Child and Adolescent Psychiatry
Johns Hopkins Medicine, Baltimore, Maryland, USA

and

Professor Pär-Anders Granhag, *PhD*
Department of Psychology
University of Gothenburg, Gothenburg, Sweden

Opponent

Professor Pär-Anders Granhag, *PhD*
Department of Psychology
University of Gothenburg, Gothenburg, Sweden

ISBN 978-951-29-3893-3 (PRINT)

ISBN 978-951-29-3894-0 (PDF)

ISSN 0082-6987

Nordprint Oy – Helsinki, Finland 2009

ABSTRACT

The purpose of the present thesis was to explore different aspects of decision making and expertise in investigations of child sexual abuse (CSA) and subsequently shed some light on the reasons for shortcomings in the investigation processes. Clinicians' subjective attitudes as well as scientifically based knowledge concerning CSA, CSA investigation and interviewing were explored. Furthermore the clinicians' own view on their expertise and what enhances this expertise was investigated. Also, the effects of scientific knowledge, experience and attitudes on the decision making in a case of CSA were explored. Finally, the effects of different kinds of feedback as well as experience on the ability to evaluate CSA in the light of children's behavior and base rates were investigated. Both explorative and experimental methods were used.

The purpose of **Study I** was to investigate whether clinicians investigating child sexual abuse (CSA) rely more on scientific knowledge or on clinical experience when evaluating their own expertise. Another goal was to check what kind of beliefs the clinicians held. The connections between these different factors were investigated. A questionnaire covering items concerning demographic data, experience, knowledge about CSA, self-evaluated expertise and beliefs about CSA was given to social workers, child psychiatrists and psychologists working with children. The results showed that the clinicians relied more on their clinical experience than on scientific knowledge when evaluating their expertise as investigators of CSA. Furthermore, social workers possessed stronger attitudes in favor of children than the other groups, while child psychiatrists had more negative attitudes towards the criminal justice system. Male participants held less strong beliefs

than female participants. The findings indicate that the education of CSA investigators should focus more on theoretical knowledge and decision making processes as well as the role of beliefs.

In **Study II** school and family counseling psychologists completed a Child Sexual Abuse Attitude and Belief Scale. Four CSA related attitude and belief subscales were identified: 1. *The Disclosure subscale* reflecting favoring a disclosure at any cost, 2. *The Pro-Child subscale* reflecting unconditional belief in children's reports, 3. *The Intuition subscale* reflecting favoring an intuitive approach to CSA investigations, and 4. *The Anti Criminal Justice System subscale* reflecting negative attitudes towards the legal system. Beliefs that were erroneous according to empirical research were analyzed separately. The results suggest that some psychologists hold extreme attitudes and many erroneous beliefs related to CSA. Some misconceptions are common. Female participants tended to hold stronger attitudes than male participants. The more training in interviewing children the participants have, the more erroneous beliefs and stronger attitudes they hold. Experience did not affect attitudes and beliefs.

In **Study III** mental health professionals' sensitivity to suggestive interviewing in CSA cases was explored. Furthermore, the effects of attitudes and beliefs related to CSA and experience with CSA investigations on the sensitivity to suggestive influences in the interview were investigated. Also, the effect of base rate estimates of CSA on decisions was examined. A questionnaire covering items concerning demographic data, different aspects of clinical experience, self-evaluated expertise, beliefs and knowledge about CSA and a set of ambiguous material based on real trial documents concerning an alleged CSA case was given to child mental health professionals. The experiment was based on a 2 x 2 x 2 x 2 (leading

questions: yes vs no) x (stereotype induction: yes vs no) x (emotional tone: pressure to respond vs no pressure to respond) x (threats and rewards: yes vs no) between-subjects factorial design, in which the suggestiveness of the methods with which the responses of the child were obtained were varied. There was an additional condition in which the material did not contain any interview transcripts. The results showed that clinicians are sensitive only to the presence of leading questions but not to the presence of other suggestive techniques. Furthermore, the clinicians were not sensitive to the possibility that suggestive techniques could have been used when no interview transcripts had been included in the trial material. Experience had an effect on the sensitivity of the clinicians only regarding leading questions. Strong beliefs related to CSA lessened the sensitivity to leading questions. Those showing strong beliefs on the belief scales used in this study were even more prone to prosecute than other participants when other suggestive influences than leading questions were present.

Controversy exists regarding effects of experience and feedback on clinical decision making. In **Study IV** the impact of the number of handled cases and of feedback on the decisions in cases of alleged CSA was investigated. One-hundred vignettes describing cases of suspected CSA were given to students with no experience with investigating CSA. The vignettes were based on statistical data about symptoms and prevalence of CSA. According to the theoretical likelihood of CSA the children described were categorized as abused or not abused. The participants were asked to decide whether abuse had occurred. They were divided into 4 groups: one received feedback on whether their decision was right or wrong, one received information about cognitive processes involved in decision making, one received both, and one did not receive feedback at all. The results showed

that participants who received feedback on their performance made more correct positive decisions and participants who got information about decision making processes made more correct negative decisions. Feedback and information combined decreased the number of correct positive decisions but increased the number of correct negative decisions. The number of read cases had in itself a positive effect on correct positive decision.

CONTENTS

ABSTRACT.....	3
ACKNOWLEDGEMENTS.....	10
LIST OF ORIGINAL PUBLICATIONS.....	12
INTRODUCTION.....	13
Definitions of sexual crimes against children.....	15
The prevalence and reporting. of child sexual abuse.....	16
Problems with investigations of child sexual abuse	20
Physical evidence.....	20
Behavioral symptoms as evidence of abuse.....	21
The relationship between the victim and the perpetrator..	23
Children’s verbal reports.....	23
Additional techniques.....	27
Decision making in CSA cases.....	30
Recommended approaches to decision making in CSA cases.....	30
Problems related to decision making and expertise in CSA cases.....	31
Factors influencing decision making in CSA cases.....	35
Summary.....	43
THE AIMS OF THE PRESENT THESIS.....	44
PARTICIPANTS.....	48
Studies I and III.....	48
Study II.....	49
Study IV.....	51

METHODS..... 52

Studies I & III..... 52

 Questionnaire..... 52

 Design, Study III..... 56

 Procedure..... 57

Study II..... 58

 Construction of the questionnaire..... 58

 Procedure..... 62

Study IV..... 62

 The questionnaire..... 62

 Procedure..... 64

STATISTICAL ANALYSES..... 65

ETHICAL CONSIDERATIONS..... 68

RESULTS 68

Study I..... 68

Study II..... 70

Study III..... 74

Study IV..... 77

DISCUSSION..... 80

Expertise among clinicians conducting CSA investigations..... 80

CSA related attitudes and beliefs among Finnish clinicians..... 87

**Factors influencing expertise and attitudes among clinicians
investigating CSA..... 90**

**The effects of experience and feedback on non-experienced evaluators’
decisions in CSA investigations..... 95**

Methodological limitations..... 99

CONCLUSIONS AND IMPLICATIONS..... 101

SUGGESTIONS FOR FUTURE RESEARCH.....	106
REFERENCES.....	107
ORIGINAL PUBLICATIONS	

ACKNOWLEDGEMENTS

First of all I am deeply indebted to my two wonderful supervisors, Pekka Santtila and Pekka Niemi. Pekka Santtila has been incredibly flexible, patient, inspiring and supportive during the time I have been writing my thesis. He has helped me with everything from problems with statistics to deep existential questions about the imperfection of the clinical world and life itself. Pekka Niemi has been an invaluable guard of the quality of all aspects of my work, from spelling and grammar to the methodology of my research (not to mention the lay out of the reference lists...). Both Pekkas have shown an incredible ability to always see the positive side of whatever comes along: I have learned a great deal from working with you, professionally as well as personally, and I have had a lot of fun in the process.

I am also grateful to all my co-authors, Jutta Sinkkonen, Niina Hakala, Linda Björnberg, Mikael Sainio and Kenneth Sandnabba for their important contributions to this thesis. Thank you for all your help!

I would like to thank Professor Maggie Bruck and Professor Pär-Anders Granhag for their reviews of the first version of this manuscript and for their valuable suggestions for improving the text.

A warm thank you also goes to my dear friend Charlotte Wikman and her husband Jason Quirk for checking my English with an insanely short notice.

For inspiring discussions, encouragement and support, as well as for tons of practical advice, I would like to thank my good friend and colleague Julia Korkman. It is a great asset to have such a close friend working in the same field and sharing common interests in various areas of both work and private life. Furthermore I wish to thank my friend Nina Mahlberg, without

whom I would not have discovered the fascinating world of forensic psychology. Also, I wish to thank my friends and co-workers at the Child Forensic Psychiatry Center for making it possible to put the scientific knowledge in this field into practice.

On a personal level I wish to thank my family and especially my father Peter Finnilä for helping me with my computer and all kinds of technical problems that are far from my area of expertise, and my dearest husband Tapani Tuohimaa for supporting me in all possible ways throughout our life together.

I would like to dedicate this work to my beautiful baby daughter Lilian, humbly wishing to be able to give her the childhood all children deserve.

This thesis was financially supported by the Finnish Ministry of Interior, the Finnish Academy, Stiftelsen för Åbo Akademis Forskningsinstitut, the Psykonet graduate school, Ella och Georg Ehrnroths stiftelse, Otto Malmbergs fond and Turun Yliopistosäätiö.

LIST OF ORIGINAL PUBLICATIONS

- I. Finnilä-Tuohimaa, K., Santtila, P., Sainio, M., Niemi, P., & Sandnabba K. (2005). Connections between experience, beliefs, scientific knowledge, and self-evaluated expertise among investigators of child sexual abuse in Finland. *Scandinavian Journal of Psychology*, 46,1-10.
- II. Finnilä-Tuohimaa, K., Santtila, P., Björnberg, L., Hakala, N., Niemi, P. & Sandnabba, K. (2008). Attitudes related to child sexual abuse: scale construction and explorative study among psychologists. *Scandinavian Journal of Psychology*, 49, 311 – 323
- III. Finnilä-Tuohimaa, K., Santtila, S., Sainio, M., Niemi, P., & Sandnabba, K (2009). Expert judgment in cases of alleged child sexual abuse: clinicians' sensitivity to suggestive influences, pre-existing beliefs and base rate estimates. *Scandinavian Journal of Psychology*, 50, 129-143
- IV. Finnilä-Tuohimaa, K., Santtila, S., Sinkkonen, J., & Niemi, P. (submitted). The effects of experience, outcome feedback and cognitive feedback on decision making in child sexual abuse cases: a simulation study

INTRODUCTION

Over the past few decades the subject of child sexual abuse (CSA) has raised media and research interest both in Finland and internationally because of strongly increased numbers of reports of alleged sex crimes against children. Through updated legislation concerning such crimes (Finnish criminal law, Decree 563/1998) and new guidelines for professionals investigating CSA (Sisäasiainministeriö [Ministry of the Interior], 2006; Taskinen, 2003) attempts have been made to improve the legal protection for both children suspected of having been sexually abused, as well as for those suspected of having committed sex crimes against children. The possibility of both underreporting and overreporting of CSA have become serious issues (Ceci & Bruck, 1995; Kauppinen, Sariola & Taskinen, 2000; Sariola, 2005), and have, therefore, been taken into consideration in the new directives for professionals working with CSA allegations.

The police frequently turn to mental health professionals (in this thesis called clinicians) for assistance with, for example, interviewing small children or evaluating the significance of behavioral symptoms in investigations of suspected CSA. Clinicians may serve both in assessment roles (e.g. as interviewers of children or adults) and in actuarial or advisory roles (e.g. when providing the legal system with data concerning base rates) and may therefore be responsible for categorizing the child as "abused" or "non-abused", and may serve as expert witnesses in court. However, despite available CSA assessment guidelines and scientific literature there is, nevertheless, controversy about how well clinicians perform as experts of CSA regardless of the role in which they serve in a specific case. For example, recent research suggests that investigative child interviews

conducted by clinicians are of alarmingly poor quality both in Finland (Korkman, Santtila, & Sandnabba, 2006; Korkman, Santtila, Westeråker, & Sandnabba, 2008) and in other countries (Cederborg, Orbach, Sternberg, & Lamb, 2000; Davies, Wescott, & Horan, 2000), and that decisions made by experienced clinicians on the basis of pre-trial material in cases of alleged CSA vary to such an extent that they cannot be considered reliable (Horner, Guyer & Kalter, 1993a, 1993b).

To improve the quality of CSA investigations it has been emphasized that extensive training should be arranged for clinicians involved in such investigations. In particular, training in interviewing children has been the focus of attention and training involving not only theoretical information but also active feedback on performance as interviewer, has been recommended (Korkman et al., 2006; Taskinen, 2003). However, in order to design training that effectively improves the clinicians' ability to investigate CSA it is essential to know what mechanisms lie behind the shortcomings of the investigations, not only when it comes to interviewing children, but also when the perception of the case as a whole, as well as the conclusions drawn from all of the available evidence, are concerned. Do clinicians lack sufficient knowledge about investigating CSA and interviewing children? Or is the knowledge they possess not based on scientific evidence? Or do clinicians have accurate knowledge, but experience difficulties putting it into practice? Or could it be that they lack the motivation or will to conduct CSA investigations according to scientific knowledge due to, for example, subjective or emotionally based attitudes related to CSA? Do the clinicians know what constitutes an expert and how expertise is enhanced so that they can actively take steps to improve their own expertise in the field of CSA? And do the clinicians understand what problems are related to judgment and

decision making *per se* so that they can avoid bias in actual cases of alleged CSA? These issues have to date been subject to empirical research to only a limited extent. The aim of the present thesis is to shed light on these questions and contribute to our knowledge concerning the mechanisms that direct the decision making of clinicians when serving as experts in cases of alleged CSA.

Definitions of sexual crimes against children

According to Finnish criminal law (chapter 20, paragraph 10) sexual intercourse is defined as penetration of the genitals or with the genitals. Sexual acts are defined as actions through which sexual arousal or satisfaction is pursued. When children are concerned also showing pornographic pictures to children or using a sexually injurious language in front of children can be considered sexually offensive acts.

Child sexual abuse is defined as intercourse with a child under the age of sixteen, as well as any other form of sexual act with a child younger than sixteen years of age, which is bound to harm a child's development (6. §). If the perpetrator is in an authority position in relation to the child, these acts with a child under the age of eighteen are considered sexual exploitation (5. §). In addition, buying sexual services from a person under the age of eighteen is criminalized (8. §). Aggravated CSA (7. §) is defined as CSA in which the abuse, due to the age and developmental stage of the child or due to the relationship between the child and the perpetrator, is bound to cause particular harm to the child, or in which the abuse is conducted in a particularly humiliating manner, and if the abuse as a whole can be considered aggravated.

What constitutes a sexual act can vary according to the age of the child. For small children, for example, touching of the child's genitals can be considered a sexual act, whereas also other forms of touching can be sexual for older children. Also, the act does not have to include touching at all, but can involve, for example, showing sexual material to children (Taskinen, 2003). Accordingly, showing pornographic material to children or placing sexual or pornographic objects or material so that children under the age of fifteen can see them, for example, in advertisements or shop windows, is also criminalized (Chapter 17, paragraphs 18b and 20). In addition, possession or distribution of pornographic material depicting children is considered a crime (18. §, 18a. §, 19. §).

For all sexual crimes against children also attempting these acts or encouraging a child to perform them is criminalized. It is worth noting that sexual acts with children are criminalized even if the child has participated seemingly without force or persuasion. On the other hand, sexual acts between children or young people at approximately the same developmental stage are not criminalized.

The prevalence and reporting of child sexual abuse

There has been a considerable increase in the number of reports of alleged CSA in Finland, as well as in many other western countries, since the beginning of the 1980's (Ceci & Bruck, 1995; Kauppinen et al., 2000). The number of CSA investigations in Finland increased from 354 during the years 1983 and 1984 to 933 during the period of one year between May 1998 and April 1999. In 1999 399 cases of alleged CSA were reported to the Finnish police, whereas the amount of reported allegations was 1025 in 2007

(Oikeuspoliittinen tutkimuslaitos [National Research Institute of Legal Policy], 2007). However, it should be noted that the police registers crimes differently in different parts of Finland, which means that a case of alleged CSA can be registered based on the report made to the police, that is, as one incident, or based on the suspected number of abuse incidents connected to the alleged case. Even so, the number of cases registered by the police can give some information about the trends in this area.

Several explanations for the increase in reports of CSA are possible. Research suggests that there are problems with both underreporting and overreporting of CSA (Ceci & Bruck, 1995; Lipian, Mills, & Brantman, 2004; London, Bruck, Ceci & Shuman, 2005; Svedin & Back, 2003). Many cases of sexual abuse never come to the authorities' knowledge and there seems to be a problem with children not telling about abuse (London et al., 2005; Svedin & Back, 2003). On the other hand, the reported cases can include false reports that may be incorrectly classified (Bruck, Ceci, & Hembrooke, 1998; Lipian et al., 2004). Internationally ratings of the proportions of reports lacking foundation vary between 5% and 35% (Ceci & Bruck, 1995). These can either be honest mistakes made in good faith or deliberate lies told by children or adults. For example, accusations of CSA can be used as a weapon by parents engaged in custody disputes (Ceci & Bruck, 1995; Lipian et al., 2004; Trocmé & Bala, 2005).

One explanation for the radically increased number of CSA allegations reported to Finnish police during recent years could be changes in guidelines for professionals working in social and health services, which include directives for reporting suspected CSA (Taskinen, 2003). According to these guidelines, contrary to previous ones, the police should be informed immediately when CSA is suspected and be the authority responsible for

investigating the allegations. This means that many of the cases that might have been investigated only by mental health professionals prior to these new guidelines are now reported to the police.

According to Sariola (2005) a popular assumption exists that the sexual abuse of children has in fact increased over the past decades. However, Sariola points to the increased media attention that CSA receives as the reason for this impression, and as one explanation for the growing number of reports. This notion seems to be supported by a recent large prevalence study conducted in Finland (Ellonen, Kääriäinen, Salmi & Sariola, 2008) in which new data were compared to representative prevalence studies carried out in the 1980's (Sariola & Uutela, 1994; 1996). This study suggests that the prevalence of CSA in Finland is in fact lower today than 20 years ago and that the context and nature of sexual experiences among 15 year olds have changed. Ellonen and colleagues (2008) report that sexual acts including physical touch are less common today than in the 1980's, and that sexual harassment and sexual suggestions on the Internet are now the most common types of sexual experiences of minors.

Thirteen percent of the 15 years old participants in Sariola's and Uutela's study in 1988 reported sexual experiences with a person at least 5 years older than themselves, whereas only 8% of the 15 year-olds in the 2008 study reported such experiences. Especially for girls the first sexual experience occurs on the average later today than 20 years ago. In 2008 3% of 15 years old girls with sexual experiences reported having had the first experience earlier than at the age of 10, whereas 5% of the girls reported the same in 1988 (Ellonen et al., 2008). For boys there seems to be a slight increase in early sexual experiences, but the researchers warn that conclusions should not be drawn from the collected data since the total

number of boys with sexual experiences before the age of 10 is so low ($N = 3$). In earlier studies it has been shown that incest is a less common phenomenon than sexual experiences of minors with a person outside the family and new data show a similar trend. In 1988 only five 15 years old girls out of a 1000 reported having experienced father-daughter or stepfather-daughter incest (Sariola & Uutela, 1996), whereas only 3 out of 1000 reported this in 2008 (Ellonen et al., 2008). For boys, only one boy with sexual experiences with a stepmother and no boys with sexual experiences with a stepfather or a biological parent were found in the 1988 data. In 2008 no boys reported sexual experiences with a parent.

Many participants in the prevalence studies in Finland reported sexual experiences with a friend or a person who might have been on the same developmental level as the participant, and in 2008 as many as 32% of the girls and 67% of the boys reported having experienced the sexual activity as positive immediately after the event. Therefore, Ellonen and colleagues (2008) have separately analyzed the prevalence of incidents that most probably can be defined as sexual abuse according to Finnish law. According to Ellonen et al. (2008) over the past 20 years there has been a decrease in such cases among girls: in 1988 6% of the 15 years old female participants had had experiences of sexual abuse, whereas 3% of the 15 years old girls in the 2008 study had had similar experiences. Such a decrease in abuse experiences was not found among boys (2-3% in both studies), but the researchers point out that the total number of boys with such experiences is so low that the effect of chance is greater than for girls.

A similar prevalence study was recently carried out in Denmark (Helweg Larsen & Bøving Larsen, 2002). In this study the prevalence rates were slightly lower than in the Finnish study from 1988: 5% of the girls and

1% of the boys had had experiences which the children themselves considered to be CSA. Of the experiences reported in the study 11% were intra-familial abuse. In other words, this study also suggests that there have not been any radical changes in the prevalence of CSA in northern Europe. It is worth noting that the CSA incidence rates from Scandinavia are lower than in many studies conducted in the USA (e.g. Finkelhor, Hotaling, Smith & Lewis, 1990). A possible explanation is that different criteria for defining CSA have been used in these studies (Sariola & Uutela, 1994; 1996).

In conclusion, observing only reported cases does not give reliable information about the prevalence of CSA. Because of the problems with overreporting and underreporting it is not possible to determine whether the increased number of reports of alleged CSA means that more genuine cases come to the authorities' attention than before, or whether the greatest increase has been in the number of false alarms. However, some research suggests that the number of reported cases of actual CSA remains low, even though the number of suspected cases grows (London et al., 2005; Sariola, 2005).

Problems with investigations of child sexual abuse

Several characteristics of cases of alleged CSA have been identified that make investigations of such crimes particularly challenging for professionals involved. These problems are described below.

Physical evidence

Only a minority of CSA suspicions can be validated on the basis of physical evidence, such as pregnancy, sexually transmitted diseases or the presence of semen (Bays & Chadwick, 1993). In other words, the lack of

such evidence does not rule out the possibility of CSA. Furthermore, it has been established that approximately 70% of CSA cases lack strong corroborating evidence, that is, either medical or other material evidence (e.g. photographs) (Herman, 2005).

Behavioral symptoms as evidence of abuse

In 1983 Summit described what was called the “child sexual abuse accommodation syndrome”. This “syndrome” included a set of five symptoms (secrecy, helplessness, accommodation, delayed unconvincing disclosure and retraction), which were thought to occur consistently together in sexually abused children and thus could be seen as diagnostic criteria. Accordingly, some clinicians in the field of CSA have seen being a victim of CSA as a psychiatric diagnosis that can be identified on the basis of certain behavioral symptoms. However, both research on symptoms related to CSA (Kendall-Tackett, Meyer Williams, & Finkelhor, 1993) as well as on the way children tell about abuse (London et al., 2005) has shown that no discrete clinical syndrome related to CSA seems to exist. It thus follows that sexual abuse cannot be diagnosed in the same way as a psychiatric disorder. In fact, no specific behavioral symptom or combination of symptoms in children can be considered valid enough indicators of sexual abuse for identification of an abused individual (Kendall-Tackett et al., 1993). On the contrary, the interaction of sexual abuse with complex factors, such as personality, family and abuse characteristics, will produce a myriad of responses rather than the same set of symptoms in all abused children even though some behaviors have been associated with identifying abused children as a group (Kendall-Tackett et al., 1993). Kuehnle (1998), who recommends strongly relying on empirical data when evaluating sexual abuse, warns that seeing sexual abuse

as a trigger for a predictable set of symptoms may misguide practitioners and lead them to inappropriately interpret data or identify features to support their identification and placement of a child in a fictional homogenous group labeled “sexually abused children”. The appropriate manner in which a child’s behavior should be viewed when investigating CSA allegations is by seeing CSA as a life event or series of events that will produce a broad range of behaviors in children (Kuehnle, 1998).

The most common behavioral problems, which according to guidelines require further investigation to rule out possible CSA (Taskinen, 2003) are sexualized behavior and post traumatic stress disorder (PTSD) (Friedrich, 1998; Kendall-Tackett et al., 1993). Some researchers have warned that sexual abuse can not be identified solely on the basis of these behavioral symptoms either, since considerable variability in them have been noted and since there seems to be no clear connection between behavioral symptoms and sexual abuse having happened (Drach, Wientzen & Ricci, 2001). However, Kuehnle (1998) presents a somewhat more complicated picture of the use of behavioral indicators of CSA. She points out that some sexual behaviors in children should be seen as strong indicators of abuse, whereas others should not be taken into consideration. For example, according to Kuehnle (1998) touching private parts, undressing or attempting to touch women’s breasts are such common behaviors (present in at least 40% of non-abused children) that no conclusions regarding sexual abuse can be drawn from them. On the other hand, some studies have shown that, for example, no non-abused children attempted to touch others’ private parts with their mouth or initiated intercourse with others, which makes these behaviors very strong indicators of abuse (Kuehnle, 1998). It follows that children’s behavior can give additional information when evaluated in the light of base rates (Dammeyer,

1998; Kuehnle, 1998; Wood & Wright, 1995), but does not alone provide conclusive evidence about the alleged event.

Few studies exist concerning clinicians' ability to estimate the probability of CSA by evaluating children's behavior using base rates, but those that do exist suggest that this ability is limited (e.g. Wood & Wright, 1995). In the present thesis one purpose was to investigate whether this ability among non-experienced evaluators of CSA could be improved through experience and different kinds of feedback.

The relationship between the victim and the perpetrator

The victim and the perpetrator are often the only ones who know about the crime. In order to avoid serious penalties, the offender may remain silent or lie about the alleged crime and is, for the same reason, likely to challenge the credibility of the child (Lamb, Sternberg, & Esplin 1998). Hence, getting a detailed witness statement from the children involved by interviewing them becomes a central part of the CSA investigations (Ceci & Bruck, 1995; Lamb, 1994). However, it has been found that children are reluctant to tell about abuse experiences and that one reason for this might be feelings of loyalty towards the perpetrator, fear of the perpetrator and feelings of shame and guilt, which might be deliberately induced by the perpetrator (e.g. Svedin & Back, 2003).

Children's verbal reports

Many professionals believe that the child's report is one of the best indicators when deciding whether abuse has taken place or not (Berliner & Conte, 1993; Lamb, 1994; Poole & Lamb, 1998). The rationale is that false allegations made by children are quite rare and are more likely to emanate

from adults or to be related to misinterpretations rather than to deliberate fabrication. However, heavy reliance on children's reports and interviews with children has also been criticized on several grounds (e.g. Ceci & Bruck, 1995; Dammeyer, 1998). First, it has been argued that the rates and characteristics of false allegations may not be known, since investigators in ambiguous cases can never be completely certain in their decision about the accuracy of CSA allegations (Dammeyer, 1998; Horowitz, Lamb, Esplin, Boychuk, Reiter-Lavery & Krispin, 1995). Second, especially pre-school children are highly susceptible to suggestions and while most experts agree that even young children can indeed provide accurate reports, there is no doubt that a child's account can be distorted by improper questioning (e.g. Ceci & Bruck, 1995). It is, for example, possible that a well-meaning parent might inadvertently lead or coach a child or that a parent engaged in a custody dispute might manipulate a child to make false allegations of abuse without it becoming obvious (Dammeyer, 1998; Lipian et al., 2004). Moreover, children's limited linguistic and memory skills may reduce the possibilities of obtaining accurate information concerning forensically important events, especially in cases involving young children and when there is a considerable delay between the event and the interview (Ceci & Bruck, 1995). For example, in Finland units conducting CSA investigations have been criticized for long waiting lists and for investigations which are too lengthy, leading to long delays between the alleged event and the interview with the child.

In recent years the possibility that professionals conducting forensic interviews with children in fact do not perform very well at this task has raised much concern. Thus, it is possible that investigators themselves compromise the forensic value of children's reports. Recent research in

Finland suggests that despite recommendations and tools for interviewing children properly in cases of alleged CSA, forensic interviews with children are of poor quality (e.g. Korkman et al., 2006). It has been argued that interviewers use language which is too complicated, as well as different kinds of suggestive techniques. (Bruck & Ceci, 1995; Ceci & Bruck, 1995; Garven, Wood, Malpass, & Shaw, 1998; Poole & Lamb, 1998; Korkman, Santtila, Drzewiecki, & Sandnabba, (2008); Korkman et al., 2006). Garven and colleagues (1998) identified four different categories of suggestive interviewing techniques used by professionals in a highly publicized day care abuse case:

1. *Suggestive questions*: This method consists of introducing new information into an interview when the child has not yet mentioned it in the same interview (Garven et al., 1998). An example would be an interviewer asking, "How did it feel when he touched you?", if a child had not mentioned having been touched in the first place. A considerable amount of research has confirmed that children are more likely than adults to give in to leading questions (for a review, see Ceci & Bruck, 1993).

2. *Social influence*: This category embraces a wide range of suggestive techniques. In the case studied by Garven et al. (1998) the most commonly used technique was telling the child about the statements of other people, thus inducing social conformity. Other techniques included intimidation, telling the child the interviewer's opinions or inducing stereotypes. Inducing a negative stereotype concerning the alleged perpetrator may increase the likelihood of both assent to leading questions and false details being inserted into a statement (Bruck et al., 1998; Leichtman & Ceci, 1995; Lepore & SESCO, 1994; Thompson, Clarke-Stewart, & Lepore, 1997). These effects are observable in both single and multiple interviews and both when the

stereotype induction takes place before and when it takes place after the event (Ceci & Bruck, 1995). This means that stereotypes may lead children both to misinterpret events when they are first observed or change their understanding of their memories concerning an event at a later date.

3. *Reinforcement*: Interviewers may sometimes use supportive statements such as, "You'll feel better once you tell", or "Don't be afraid to tell" to help a child reveal details of an allegation, thus, showing they believe that an event has in fact taken place (Bruck & Ceci, 1995, Garven et al., 1998). Several studies have suggested that children may be more likely to make false statements when such utterances are used (for a review see Ceci & Bruck, 1993). Learning theory and research suggest that promises of praise, approval or other rewards (such as presents or food) to a child for making certain kinds of statement or, alternatively, criticism or disagreement with what the child has said or otherwise indicating that the statement is not acceptable, are likely to affect a child's statement in an interview setting (Garven et al., 1998).

4. *Removal from direct experience*: The last category identified by Garven et al. (1998) consisted of asking the child to speculate and conversing indirectly through puppets. Such techniques make it difficult to determine whether children tell about real experiences or use their imagination.

In the study by Garven and colleagues (1998) it was found that when using this wide range of suggestive techniques, children could be led to give false reports of abuse even during a single interview. In a more recent Finnish study (Finnilä, Mahlberg, Santtila, Sandnabba, & Niemi, 2003) these techniques were also studied in an experimental setting and were shown to influence children's answers greatly even during a single suggestive interview.

One suggested reason for problems in child interviews has been *a priori* over- or underestimations of the occurrence of certain events on the part of the interviewer. False beliefs and strong attitudes concerning CSA may also mold linguistic interactions in an interview and lead to the use of suggestive interviewing methods and therefore to false disclosure about sexual abuse (Ceci & Bruck, 1993; 1995). Some studies exist concerning attitudes among professionals working with CSA (Brigham & Spier, 1992; Davey & Hill, 1995; Hartman, Karlson, & Hibbard, 1994; Hetherington & Beardsall; 1998), but more research is needed to fully understand what kind of attitudes and beliefs, especially among clinicians, might have an impact on the interviews. There is also anecdotal evidence that the reason for suggestive interviewing might be that clinicians do not know what kind of questions are in fact suggestive. However, this question has not yet been subjected to empirical research and one purpose of the present thesis was to shed light on it.

Dammeyer (1998) argues that the factors reducing the credibility of children's reports make it difficult to operationalize the usefulness of such reports as indicators of abuse. Under ideal conditions the child's statement provides moderate to strong evidence for deciding the truthfulness of allegations. On the other hand, it is difficult to determine whether such conditions are present and cases may vary between "ideal" and "highly compromised" (Dammeyer, 1998).

Additional techniques

Apart from interviewing children, several alternative techniques for determining if a child has been sexually abused have been used in clinical practice. These involve the use of psychological tests, screening instruments

and anatomically detailed (AD) dolls, drawings, puppets, dollhouses or the use of specific props, for example, the so called Touch Survey, in which the interviewer directly asks the child whether someone has touched different parts of the child's body and what feelings the touching evoked (Ceci & Bruck, 1995; Dammeyer, 1998; Hewitt, 1999). There is very little empirical research on some of these techniques (e.g. the Touch Survey) (Cronch, Viljoen & Hansen, 2005; Hewitt, 1999), whereas others have been subjected to extensive scientific exploration (e.g. AD dolls) (Bruck, Ceci, Francouer & Renick, 1995; DeLoache & Marzolf, 1995; Everson & Boat, 1997; Koocher, Goodman, White, Friedrich, Sivan & Reynolds, 1995; Santtila, Korkman, & Sandnabba, 2004; Skinner, 1996; Thierry, Lamb, Orbach, & Pipe, 2005).

Generally, psychological tests have not been considered reliable tools for identifying sexually abused children, as it has not been established what responses would be valid indicators of abuse (Dammeyer, 1998; Kuehnle, 1998). On the contrary, responses supposedly indicative of abuse do not consistently occur with higher frequency among abused children than among non-abused children. Another problem is that even though some responses occur more often among sexually abused than among non-abused children, the base rate of these responses is so low for all groups of children that their utility for investigating abuse is limited (Dammeyer, 1998).

The use of AD dolls in CSA investigations has been intensively debated in the literature for over 20 years and opinions on their usefulness still vary considerably. Some researchers recommend continued use of AD dolls, for example when discussing anatomy with a child or when introducing topics related to the genital area, as long as the interviewer has received proper training (Everson & Boat, 1997; Koocher et al., 1995). Others claim that AD dolls should not be used in forensic interviews with children, since

the dolls in themselves are suggestive (Bruck et al., 1995), since young children do not understand the symbolic meaning of the dolls (DeLoache & Marzolf, 1995), and since the dolls decrease the quality of the questions asked (Santtila et al., 2004). It has also been pointed out that no normative data on children's play with AD dolls exists. Therefore, it is not possible to determine whether a child has been abused or not based on how the child plays with AD dolls (e.g. Ceci & Bruck, 1995; Skinner, 1996). On the same grounds the use of other toys or drawings has been criticized (Dammeyer, 1998; Garven et al., 1998). However, most experts agree that AD dolls should not be used as psychological "tests" to which certain responses can be judged as indicative of abuse, and that further research is needed to determine the best use of AD dolls.

The so called Touch Survey that has gained popularity in recent years was developed by psychologist Sandra Hewitt in the early 1980's (Cronch et al., 2005; Hewitt, 1999). The technique is based on the idea that touches fall along a continuum, ranging from good to neutral to bad, and the survey consists of a discussion about different kinds of touches using a drawing or dolls (Hewitt, 1999). The Touch Survey is recommended for children aged 4-8. The empirical data on this technique is limited to one pilot study by Hewitt and Arrowood (1994), which suggested that there might be problems with children not disclosing abuse through this procedure. Another problem with the Touch Survey is that it involves direct questioning about touching of the genital area. It has been well established in numerous studies (Ceci & Bruck, 1995; Poole & Lamb, 1998) that such questions should not be used with pre-school children, since these questions, due to their suggestiveness, may elicit false reports of abuse.

All in all it is difficult to assess the effects of these techniques on children's answers and decision making in CSA cases, since they can be used both as indicators and communicators (Dammeyer, 1998). The information elicited through these techniques should generally be viewed with caution, but it has also been argued that when taking individual differences into account the usefulness of such techniques may vary so that they in fact can be important tools in interviewing some children (Dammeyer, 1998; Poole & Dickinson, 2005).

Decision making in CSA cases

Recommended approaches to decision making in CSA cases

Since there is no specific indicator of CSA that in the absence of indisputable material evidence could be used to determine whether a child has been sexually abused or not, most experts recommend relying on comprehensive assessments rather than on separate indicators (e.g. APSAC, 1990; Kuehnle, 1998; Lamb, 1994). Furthermore, the guidelines currently followed in Finland recommend that a CSA investigation should consist of the gathering of background information from parents, day care and authorities (e.g. through interviews and screening tools, such as the Child Behavior Check List), a medical examination of the child, a psychological evaluation of the child to get information about behavioral symptoms and developmental stage (e.g. using common psychological tests and play observation), as well as the interview with the child (Taskinen, 2003).

Problems related to decision making and expertise in CSA cases

Research on judgment and decision making has suggested different kinds of definitions on what constitutes an expert and expert judgment. Einhorn (1974) proposes that experts should be able to identify relevant cues (i.e. those containing information relevant for the final decision) from the multidimensional stimulus they encounter and furthermore be able to combine these, so that their judgment is reasonably correct, that is, better than chance. Cronbach and Meehl (1955) emphasize that when measuring cues and relevant information, intrajudge reliability, construct validity and judgmental bias need to be taken into consideration. In other words, experts should be able to reproduce their measurements of the cues, that is, reach the same decision when confronted with the same problem twice. Furthermore, the cues being measured should have the status of explanatory concepts, and well known cognitive illusions and biases should not contaminate the judgment of an expert (Cronbach & Meehl, 1955). Faust and Ziskin (1988), who examined the decision making of expert witnesses, concluded that experts should make decisions with reasonable certainty, that is, better than chance; should be reasonably certain about the issues pertaining to the case at hand (i.e. not merely about their field in general or some specific issues); and should be able to help the court reach a more valid conclusion than would be possible without expert testimony. However, in light of these definitions several problems have been identified in clinical decision making and decision making in CSA cases.

Even if it is logical to prefer looking at multiple indicators over single ones when making decisions about alleged CSA, there are problems with this approach (Dammeyer, 1998). There are no recommendations about which indicators should be taken into consideration and what number of indicators

assure sufficient material for making a decision about whether a child has been abused or not. More importantly, however, no guidance is given to clinicians regarding how to integrate the complex information they encounter when conducting a multimodal assessment (e.g. APSAC, 1990; Taskinen, 2003). Dammeyer (1998) argues that this might be one reason for the wide range of assessment approaches among clinicians working with abuse allegations, and warns that without guidance clinicians may selectively choose indicators consistent with personal biases. It is also possible that clinicians collect a massive amount of data, some of which is not at all relevant for investigating the alleged abuse, and consequently draw conclusions from factors that in reality give no additional information concerning the suspicion. These risks are clearly evidenced in research on judgment and decision making in general, which shows that human judgment under uncertainty, that is, the ability to integrate complex data and draw correct conclusions from multiple sources of information, is highly flawed (e.g. Dawes, Faust, & Meehl, 1989). However, less research has been conducted into how these shortcomings in judgment and decision making affect CSA investigations in particular.

Judgment in clinical psychology and psychiatry has been the focus of interest for a few decades since decisions made in these areas mostly are made under high uncertainty (e.g. Dawes, 1994; Gambrill, 2006; Garb, 1989). More recently the expertise and judgment of clinicians in forensic settings and especially in CSA investigations has evoked research attention (Dammeyer, 1998; Fargason, Peralta-Carcelen, Fountain, Amaya, & Centor, 1997; Horner et al., 1993; Kuehnle, 1998). In accordance with research on judgment and decision making in general and clinical decision making in psychology and psychiatry in particular, the studies conducted into clinicians'

performance as experts in CSA cases suggest that serious problems may exist in this area and that clinicians involved in CSA examinations do not use skills consistent with good decision making or expertise in their work (e.g. Fargason et al., 1997).

Failure to understand the role of uncertainty in CSA investigations can cause decision makers to increase the number of children erroneously categorized as either “abused” or “not abused” when interpreting the information available (Dawes, 1994; Fargason et al., 1997; Horner, Guyer, & Kalter, 1993a; b; Realmuto & Wescoe, 1992). In order to understand the factors that might distort the decision making process in CSA investigations it is necessary to understand the specific features of the crime as well as the problems that lead to systematic errors in human judgment under uncertainty *per se*.

Clinical versus actuarial prediction and postdiction. CSA assessment usually involves postdiction (i.e. whether the child has been previously abused) as opposed to prediction. However, the demands on clinicians are identical in both types of decision making (i.e. the ability to examine complex and sometimes contradictory data, successfully integrate this information and arrive at accurate conclusions). Over the last forty years a vast body of research evidence has been amassed, indicating that clinical prediction in general is consistently inferior to prediction based on actuarial (statistical) data (Dawes, 1994; Dawes, Faust & Meehl, 1989; Goldberg, Faust, Kleinmutz & Dawes, 1991). It has been shown that human judgment under uncertainty usually is not based on statistical probabilities but rather on heuristic rules, which in complex tasks may lead to systematic errors (Edwards & Winterfeldt, 1986; Tversky & Kahneman, 1974) and that this also affects clinical work (Dammeyer, 1998; Dawes, 1994; Horner, Guyer &

Kalter, 1993b). Uncertainty should be represented in terms of probabilities, which should be subsequently adjusted based on new information (Baron, 1994; Dammeyer, 1998; Kuehnle, 1998). This can be achieved by using Bayes's Theorem, which has been shown to be applicable to decision making in CSA investigations (Herman, 2005; Wood, 1996). The theorem assumes that a decision maker starts with a baseline estimate of the probability of an event. The theorem then provides a framework for adjusting this baseline estimate when the decision maker is provided with new information. Appropriate clinical application of Bayes's Theorem depends on knowledge of the base rate probability in question, that is, the prevalence of CSA and knowledge of the sensitivity and specificity of assessment methods. However, there is research evidence that suggests that Bayes's Theorem is not used as the basis for clinicians' conception of uncertain data in CSA investigations (Dammeyer, 1998; Fargason et al., 1997; Horner, Guyer & Kalter, 1993a; b).

Confirmation bias. CSA is an emotionally charged crime and there is considerable pressure on clinicians to reach decisions in the best interest of the child. Therefore, clinicians may fail to notice a pre-existing assumption concerning the alleged abuse (Ceci & Bruck, 1995), and fail to take alternative hypotheses into consideration when examining the child, even though ruling out different hypotheses would be consistent with the criteria for good decision making processes described in the literature (Baron, 1994; Tversky & Kahneman, 1974). In other words, when confronted with complex data about a child, clinicians might selectively consider only pieces of information that confirm their pre-existing assumptions about the case.

Availability. The frequency of a certain event and thus the probability upon which the decision is based has been shown to be assessed on the basis

of how easily instances are recalled and not on the basis of how common the phenomenon in question actually is (Edwards & von Winterfeldt, 1986; Tversky & Kahneman, 1974). Clinicians might exaggerate the probability of certain phenomena, for example, those they encounter frequently or those that are emotionally charged. For example, those who encounter many cases of alleged CSA in their work or attend extensive training in the field might be prone to overestimate the probability of abuse.

Overconfidence. This bias in human judgment means that the estimated probability on which the decision is based is more extreme than it is in reality, because it is easier to operate with such numbers (Edwards & von Winterfeldt, 1986). Research has shown that people tend to be especially overconfident in complex tasks, which makes CSA investigations vulnerable to such bias.

Factors influencing decision making in CSA cases

Apart from the fact that problems related to human judgment *per se* may influence decision making in CSA cases, some specific factors affecting decision making both in clinical and legal settings have been identified.

Attitudes and beliefs. Research suggests that both interviews and decision making processes in CSA cases can be negatively influenced by erroneous beliefs about CSA and by strong attitudes related to the phenomenon (e.g. Ceci & Bruck, 1995). Erroneous beliefs and strongly held attitudes may decrease the ability of clinicians to detect and avoid suggestive influences and may result in molding of linguistic interactions, as well as interpretations of children's answers and behavior during an interview to the extent that false narratives of sexual abuse are obtained (Ceci & Bruck, 1993; 1995). There is reason to believe that attitudes and pre-existing beliefs may

be especially strong in cases of alleged CSA, since it represents such an emotionally charged crime.

There have been previous attempts to create questionnaires to measure CSA related attitudes. For example, Collings (1997) developed the Child Sexual Abuse Myth Scale as an instrument to measure the acceptance of CSA myths and stereotypes. Collings, however, recruited a non-probability sample, that is lay people, drawn from community. Ferrara (1999), on the other hand, studied school psychologists' attitudes when constructing the Child Sexual Abuse Attitude Scale. Ferrara's focus was on the validation of an attitude scale based on Festinger's (1957) cognitive dissonance theory. She found four attitude subscales with confirmatory factor analysis: cognition, behavior, value, and affect. These scales could be useful, for example, when customizing staff training programs and evaluating their effectiveness (Ferrara, 1999). However, the focus of Ferrara's attitude scale was on underreporting of CSA; issues concerning overreporting were not included.

Attitudes have been defined as a readiness to respond in a favorable or unfavorable manner to an object or class of objects (Gergen & Gergen, 1986). Accordingly, attitudes have a topic, are judgmental and are thought to be relatively long lasting. In addition, it has been established that attitudes have three components: a cognitive component, an affective component and a behavioral component. Attitudes are thought to be formed early in life, through modeling or later through similar processes by the influence of authorities and, for example, the media. Also, negative experiences in the form of punishment have been thought to play a role in the forming of attitudes. Furthermore, attitudes are maintained by social support. To change attitudes it is necessary to address both the emotional and the cognitive aspects of the attitudes, for example by inducing what Festinger (1957)

named cognitive dissonance. According to Festinger (1957), holding inconsistent cognitions is an unpleasant state of arousal and people are motivated to reduce such an unsettling state by changing their cognitions and behavior. It follows that attitudes are not primarily or solely formed, maintained or changed by education and information, even though one component of attitudes is cognitive, that is, consisting of factual notions about the object or phenomenon. Beliefs among experts on the other hand could be reasoned to be purely factual and obtained through education and information. In theory, false beliefs (i.e. beliefs that are not supported by scientific knowledge) among experts could be based on misunderstandings, outdated or incomplete knowledge. If this is the case it should not be necessary to address any emotional component to change such beliefs and therefore, it could be reasoned that false beliefs could be more easily changed and corrected simply through education than attitudes. It would also be reasonable to assume that professionals would have a readiness to change their beliefs when faced with new information, since it lies in the nature of being an expert to constantly change notions and opinions according to new scientific data. In previous research on CSA related attitudes and beliefs little attention has been paid to the question of whether CSA related beliefs can be considered “correct” or “false”, that is, in line with scientific knowledge or not. However, since there is, at least in theory, a clear difference between how attitudes as opposed to factual beliefs among experts are formed, maintained and changed, it is reasonable to assume that the effects of attitudes on CSA investigations and on the implications for education programs in the field of CSA may be different from those of beliefs. Therefore, it is essential to investigate them separately. In the present thesis, attitudes were defined as subjective emotions and opinions related to CSA,

the correctness of which is not possible to determine, whereas beliefs were defined as viewpoints, the correctness of which can be evaluated in the light of scientific knowledge.

Attitudes and beliefs can be strongly related to the professional role (Brigham & Spier, 1992; Hartman et al., 1994). For example, defense attorneys might be apt to hold attitudes and beliefs that favor the defendant, while prosecutors might hold attitudes and beliefs favoring the alleged victim. Clinicians on the other hand might have a treatment-centered approach to CSA. Prosecuting attorneys, child protection workers, and law enforcement personnel have been found to have strong confidence in child witnesses' cognitive capacity and memory (Brigham & Spier, 1992). These professionals thought, for example, that children are able to resist suggestion and that a professional could usually tell when a child was telling the truth. In the same study defense attorneys, on the other hand, felt that compared to adults, children usually attach more inconsistencies to their testimonies and that they are more suggestible than adults. Cheung and McNeil Boutte-Queen (2000) found that police officers were more likely to have ambivalent and revenge feelings in relation to CSA compared to social workers, who were more likely to feel empathy with the child's condition and fear of being inadequate in handling the situation.

Davey and Hill (1995) studied different professional groups who interview children while investigating suspected sexual abuse. They examined the background, training, and beliefs concerning indicators of CSA, and concluded that psychologists and medical professionals were the most cautious of all in linking behavioral and psychological signs to sexual abuse. However, Davey and Hill (1995) also concluded that training did not appear to have any effect on the professionals' beliefs about CSA indicators.

Gender has been found to influence CSA related attitudes and beliefs among both professionals and lay people (Ford Peters, 2001; Hetherington & Beardsall, 1998; Hubbartt & Singg, 2001; Trute, Adkins & MacDonald, 1992; Wellman, 1993). Hence, the gender of the investigator might have some effect on the outcome of the CSA investigations. Trute and colleagues (1992) found that women were more likely to regard sexual abuse of children as widespread. Both women and men seem to feel uneasy about sexual abuse in general, but women seem to have stronger emotional reactions to CSA, such as feelings of fright and disgust (Wellman, 1993). Also, women seem to possess more child-believing attitudes (Ford Peters, 2001; Gabora et al., 1993). Hubbartt and Singg (2001) argue that these differences might reflect the female nurturing and care-taking role in our societies. Having children of their own, however, does not seem to have an effect on clinicians' CSA attitudes (Davey & Hill, 1995).

Researchers emphasize theoretical training for improving the decision making of clinicians in CSA cases. Training could be helpful in preventing the use of heuristics and decrease pre-existing attitudes and erroneous beliefs (Dawes, 1989; 1994; Garb, 1989). However, in some studies concerning CSA related attitudes, training has been shown to have no effect on attitudes, beliefs or decision making (Davey & Hill, 1995).

To date no research has been conducted in Finland to investigate what kind of attitudes and beliefs related to CSA clinicians dealing with these questions hold or how these attitudes and beliefs influence the clinicians' decisions. Therefore, the present thesis addresses these questions.

Experience and feedback. Experience is often referred to as an important factor in clinical decision making and forming of expertise, in CSA investigations as well as in other mental health practice. For example,

experience is listed as one of the cornerstones for good clinical work in the guidelines for ethical clinical practice for Finnish and European psychologists (Finnish Psychological Association, 2002; Övreeide, 2002) and in recently published guidelines for investigating CSA (Taskinen, 2003). However, there is research evidence suggesting that clinical experience does not improve accuracy of judgment in clinical mental health practice (Dawes, 1989; 1994; Garb, 1989). In order for people to learn from experience, two conditions need to be met. They have to know what constitutes correct and incorrect answers and they have to receive feedback on whether their decisions were wrong or right (Dawes, 1994). An environment where these criteria are met is what Hogarth (2001) calls a kind learning environment, that is, an environment where not only expertise can be developed, but also an unconscious body of knowledge that can be successfully used when making decisions in complicated situations. However, neither of these conditions is usually satisfied in clinical mental health work since psychiatric diagnoses, the basic decision types in this field, do not represent natural categories in the sense that, for example, genetic diseases do and, therefore, the rules for how to categorize individuals in them are not clear. Furthermore, clinicians seldom receive systematic and non-biased feedback on their performance. Clinicians may either experience feedback that is too difficult to interpret or no feedback at all. Furthermore, the feedback that mental health professionals receive may represent so called self-fulfilling prophecies that have arisen from the professional's own decisions (Dawes, 1989; 1994; Einhorn & Hogarth, 1978). Clinicians may see this as evidence of the decision having been correct. Accordingly, diagnostic situations have been identified as what Hogarth (2001) calls wicked learning environments, that is, situations where

the right lesson is not learned and neither expertise nor useful intuitive thinking is developed (Ganbrill, 2006; Hogarth, 2001)

The problem with feedback is further complicated by clinical work itself creating an illusion of learning and of non-biased feedback, which can make it difficult for the clinicians to be aware of the problems concerning learning from experience (Dawes, 1994). The times when professionals are right in their categorizations are often remembered quite well, while the times when they are wrong in their judgments are not. An impression that the clinicians make correct decisions most of the time is thus created.

A similar impression can be created by what Tversky and Kahneman (1974) call *representative* thinking. This means that clinicians combine a certain problem with a certain characteristic, thus forgetting that it is possible to have the characteristic without having the problem. Since clinicians continuously may be in contact with people with the problem *and* the characteristic, for example depressed people with dominant mothers, they conclude that every new patient with both criteria constitutes feedback verifying that those with the problem have the characteristic. This conclusion is not logical because the clinicians then forget all those people *with* the characteristic who do *not* have the problem whom they do not meet, that is, people with dominant mothers who are *not* depressed and therefore do not seek professional help (Dawes, 1986; Melton, 1987; Murphy, 1987). Sexualized behavior in children, for example, can be overestimated as definitive proof of CSA due to this bias, since sexualized behavior in the child may be seen as *representative* of the group "sexually abused children". Only clinical work based on scientific knowledge, not on impressions, can correct this error.

In addition to the fact that the self-fulfilling prophecy phenomenon also creates an illusion of positive non-biased feedback there is yet another cognitive illusion that, for example according to Edwards and von Winterfeldt (1986), distorts the impression of feedback and learning for clinicians. This is the *hindsight* illusion. It means that when the clinician receives feedback that is not consistent with his judgment, he reconstructs his own memory of his decision making process so that it, according to his memory, leads to the correct answer. This way the decision making process will not change as a result of feedback.

In cases of CSA the conditions "not abused" as opposed to "abused" represent natural categories in a more transparent way than diagnoses in mental health work usually do, since the law offers clear criteria for what CSA is. When identifying CSA the investigator tries to find out whether an event has taken place as opposed to finding a diagnosis. Feedback could also be non-biased in some of the cases, for example, when perpetrators confess to their crimes or unambiguous physical evidence is found. Therefore, it is not self-evident that the research findings concerning the effects of experience on clinical decision making can be generalized to CSA cases. However, little research has been done to clarify this and, therefore, the effects of experience were given particular attention in the present thesis.

The importance of feedback for learning from experience has been emphasized in much of the scientific literature and the lack of feedback on the correctness of decisions has been pointed out as one of the main reasons why experience does not improve clinical judgment (e.g. Dawes 1989; 1994; Garb 1989). However, some researchers suggest that outcome feedback that is, feedback on the correctness of the decision reached, is not the only type of feedback that can influence the expertise and improvement of judgment

under uncertainty (Balzer, Doherty & O'Connor Jr, 1989). Since outcome feedback seldom is available in clinical settings, a different kind of feedback, so called cognitive feedback, could according to these researchers improve clinical judgment, to an even greater extent. Cognitive feedback means that persons making the decisions receive feedback and information through which they can determine how well they perform as decision makers, that is, how correctly they use good decision making strategies, how well they avoid cognitive illusions and biases or, as in the present research, how well they utilize information on the risk to make errors when reaching a decision.

Summary

Reports of suspected child sexual abuse have increased considerably in the past few decades. Some research evidence suggests that this reflects the increased media attention and public discussion of such problems, which may have led to a higher number of false reports, rather than an actual increase in sex crimes against children (e.g. Helweg Larsen & Bøving Larsen, 2002; London et al., 2005; Sariola, 2005). The basic definition of CSA according to Finnish criminal law is sexual acts with a child younger than 16 years of age. Since the police frequently turn to clinicians for help in the investigation of alleged CSA it is of utmost importance to investigate the ability of these professionals to investigate CSA. In spite of existing guidelines for clinicians about how to investigate CSA properly, controversy exists on how well clinicians perform when investigating such cases. It has been argued that clinicians use suggestive interviewing techniques and interpret evidence selectively according to personal pre-existing bias. This, in turn, leads to erroneous decisions when evaluating whether a child has been sexually

abused or not. However, CSA investigations are extremely demanding, since clear evidence often is not available and decisions must be made on the basis of several different uncertain indicators.

To reach a deeper understanding of why shortcomings seem to exist in investigations of CSA it is essential to understand what processes underlie such shortcomings. Several factors related to expertise and decision making have been identified as possible sources for problems in CSA investigations. One problem may be the human tendency to reach decisions through heuristics, that is "rules of thumb" that may not be in line with scientific knowledge at all, but which can overwhelm the decision making process at the expense of expert knowledge. Another problem is that clinicians may possess false beliefs or strong subjective attitudes that influence the investigations by inducing bias. In addition, even though experience has been emphasized as a factor diminishing errors in clinical decision making and in CSA investigation in particular, there is little evidence that expertise actually increases through clinical experience. The reason for this has been suggested to be the lack of unbiased feedback, a circumstance that is typical for clinical situations.

THE AIMS OF THE PRESENT THESIS

The general aim of the present thesis was to investigate different aspects of decision making processes and expertise in CSA investigations performed by mental health professionals, in this thesis called clinicians, and thus shed light on the reasons why previous research has shown that CSA investigations conducted by clinicians are of alarmingly poor quality. In order

to investigate these issues from different angles both experimental and explorative approaches were adopted.

The purpose of **Study I**, which represented explorative research, was to map important factors related to decision making in CSA investigations among clinicians in Finland: How familiar are clinicians conducting CSA investigations with scientific literature concerning CSA? Do clinicians depend more upon clinical experience or on theoretical knowledge when investigating CSA? What are the determinants of the level of expertise that the clinicians attribute to themselves? What kind of attitudes and beliefs do clinicians involved in CSA investigations hold on these issues? Are there any differences in bases for decision making between groups of professionals or between male and female clinicians?

Researchers have repeatedly suggested the need for further and more thorough assessment of CSA related attitudes with samples of professionals who work with CSA cases (e.g. Cheung & McNeil Boutte-Queen, 2000; Maynard & Wiederman, 1997; Wagner et al., 1993). Therefore, the aim of **Study II** was to take the previous research a step further and more thoroughly explore both attitudes and beliefs among Finnish psychologists who work with children, and, therefore, are potentially working with suspected CSA cases. The effects of training, different levels of clinical experience and experience with CSA, as well as gender, were explored. Because no previous CSA attitude and belief scale for professionals working with children of all ages existed, a CSA attitude and belief scale (CSAABS) was constructed first for this purpose.

In **Study III**, which was an experimental study, the aim was to investigate whether clinicians investigating CSA are sensitive to suggestive influences in an investigative interview with a child. Furthermore, the aim

was to see how experience, scientific knowledge, self-evaluated expertise and pre existing beliefs related to CSA affect this sensitivity and the evaluation of a CSA case. Finally, it was checked if the clinicians in the present study used Bayesian reasoning, that is, were sensitive to their own base rate estimates of CSA, when making a decision in a CSA case presented to them. For investigating these issues, five hypotheses were formulated:

1. The presence of suggestive techniques in interview transcripts will decrease the credibility of the information obtained in the interview, and will make clinicians less prone to decide that the abuse has occurred. There will be additive effects so that when several suggestive techniques are present, the credibility of the information will be even lower than when just one technique is present, and thus the clinicians will be even less prone to accuse.
2. The clinicians will be familiar with the research suggesting problems in interviews with children, and thus, be aware of the strong possibility of suggestive interviewing when no interview transcripts are available. Their judgment concerning the credibility of the accusations will therefore be as low as when the interview transcripts show poor interviewing practices, and lower compared to interview transcripts showing good interview practices.
3. Clinical experience in general and experience with CSA cases in particular, as well as scientific competence, will be positively related to sensitivity to specific suggestive techniques.

4. (a) Holding extreme beliefs on the two belief scales will be associated with higher estimates of the credibility of the accusations. (b) Holding extreme beliefs will be associated with decreased sensitivity to suggestive techniques.
5. Clinicians' own estimates of incest abuse of girls will be positively related to the estimate of the probability of abuse in the presented case.

Finally, in **Study IV** the focus was on investigating the effects of experience and different kinds of feedback on decision making in CSA investigations in an experimental simulation design. Based on earlier research the following hypotheses were formulated:

1. Experience in itself does not improve decision making when no feedback is given.
2. Both outcome and cognitive feedback have positive effects on decision making.
3. The more feedback the participants receive, the better they perform when assessing the cases.

A summary of the participants, methods and purpose in all four studies are displayed in Table 1.

Table 1

Participants, Design and Purpose of Studies I, II, III and IV

Study	Participants	Type of study	Purpose
I	126 social workers, 134 psychologists, 60 child psychiatrists working with children	Explorative	To explore CSA related expertise, attitudes and beliefs among clinicians working with children
II	181 students of psychology, social policy and medicine, 58 students of psychology and 242 psychologists working with children	Explorative	To develop a child sexual abuse attitude and belief scale and to further explore CSA related attitudes and beliefs among psychologists working with children
III	126 social workers, 134 psychologists, 60 child psychiatrists working with children	Experimental	To investigate clinicians' sensitivity to suggestive influences in child interviews as well as clinicians' decision making in CSA cases
IV	83 students with no prior knowledge or experience with CSA	Experimental	To investigate the effects of different kinds of feedback as well as experience on decisions in CSA cases

PARTICIPANTS**Studies I and III**

In both Study I and Study III the participants were 126 child social workers (age $M = 44.42$, $SD = 7.37$), 60 child psychiatrists (age $M = 48.62$, $SD = 9.77$), and 134 child psychologists (age $M = 43.89$, $SD = 8.49$). There were 279 female and 41 male participants. Descriptive statistics concerning the respondents are displayed in Table 2.

Table 2
Means and Standard Deviations of Age and Work Related Variables in the Three Professional Groups

	Social worker (n=126)		Child psychiatrist (n=60)		Psychologist (n=134)		Oneway
	M	SD	M	SD	M	SD	F
Age	44.42 ^a	7.37	48.89 ^b	9.77	43.89 ^a	8.49	1.15**
Years in profession	15.71	7.32	15.83	9.31	15.79	8.38	0.01
Proportion of child cases	61.95 ^a	28.95	86.25 ^b	22.12	72.12 ^c	30.01	15.27***
Nr of CSA cases	3.95 ^a	3.51	16.08 ^b	19.08	5.83 ^a	6.92	33.38***

Means with different superscripts differ significantly from each other (Duncan's multiple range test, $p < .05$)

** $p < .005$

*** $p < .001$

Study II

Study II consisted two parts. In the first part of the study the CSAABS was constructed and in the second part the scale was given to clinicians in the field of CSA.

In part I, in which item selection was conducted, first participants were students of psychology, social policy, and medicine at the Universities of Turku, Helsinki, Tampere, Joensuu, Jyväskylä, Lapland, and Åbo Akademi. One hundred and fifty-four female students (age: $M = 25.5$ years, $SD = 8.3$ years), and twenty-seven male students (age: $M = 23.5$ years, $SD = 3.0$ years) answered the first version of the questionnaire. Because a large number of questionnaires were handed out by persons other than the authors, an exact response rate for the first part of the study could not be calculated. An

estimated 30% of the students that were asked to participate in the pilot study volunteered to do so.

Second, in part I, in which both explorative and confirmatory factor analyses were conducted, participants included the student sample from the first phase, as well as 242 school- and family-counseling psychologists (age: $M = 43.5$ years, $SD = 10.7$ years). There were 214 female and 28 male participants. Altogether 242 out of 583 psychologists participated in the study (40%). The response rate is comparable to other survey studies on sexuality (Fife-Schaw, 2000). For example, Rodriguez (2002) who studied professionals' child abuse related attitudes obtained a response rate of 39%.

The geographic distribution of the respondents reflected the geographic distribution of the population in Finland. That is, the respondents were from all parts of Finland in proportion to the size of the general population in different parts of the country.

Finally, in part I, in which test-retest reliability as well as preliminary predictive validity of the factors was checked, 58 students of psychology at Åbo Akademi University and the University of Turku took part in the first survey and out of these, 26 students answered the questionnaire a second time after three weeks.

The participants in part II of Study II were the same 242 school- and family counseling psychologists (age: $M = 43.5$ years, $SD = 10.7$ years) that had participated in the first part. There were 214 female and 28 male participants. The demographic distribution of the respondents reflected the general population distribution of Finland. Additional descriptive statistics concerning the participants are displayed in Table 3.

Table 3
Means and Standard Deviations of Age and Work Related Variables Among Females, Males, and Total Number of Participating Psychologists as well as Significant Differences Between Sexes, as Measured by t-test

	Total (N = 242)		Female (n = 214)		Male (n = 28)		t-test
	M	SD	M	SD	M	SD	t
Age	43.5	10.7	43.3	10.8	45.3	9.6	0.943
Years in profession	15.94	10.39	15.83	10.45	16.83	10.06	0.477
Proportion of work time with children (%)	53.90 %	21.83 %	54.16 %	21.80 %	52.00 %	22.33 %	0.481
Nr of experienced CSA cases	5.38	7.55	4.98	7.31	8.64	8.66	2.052*
Years at current work place	9.29	9.14	9.18	9.24	10.06	8.45	0.474
Years working with children	13.88	10.43	13.79	10.60	14.58	9.23	0.373
Hours of training in interviewing children (h)	17.55 h	33.22 h	15.74 h	27.36 h	31.00 h	60.53 h	1.292
Hours of training about CSA (h)	24.70 h	43.34 h	22.98 h	39.53 h	37.07 h	74.97 h	1.546

Note. h = hours. * indicates a significant difference between sexes, $p < 0.05$

Study IV

The participants in Study IV were volunteering students recruited through advertisements on the bulletin boards of the universities of Turku and Helsinki, as well as on e-mail lists of all universities and polytechnics in Finland. The precondition for participating was that the students had no previous experience with or knowledge about CSA or CSA investigations or scientific research concerning the area. Altogether 188 students answered the advertisements. Of these 106 took part in the experiment, but since 23 interrupted, the final number of participants was 83. Demographic information concerning the participants is displayed in Table 4.

Some of the participants reported having experience or knowledge about CSA, but when this was evaluated it appeared that their experience or knowledge was so limited (e.g. having read about a CSA case once in a newspaper) that they were allowed take part in the study.

Table 4

Age, Profession and Previous Knowledge of CSA among Participants in Study I

Gender	n	%	Age	n	%	Work	n	%	Knowledge	n	%
Female	75	90.4	19 – 24	47	56,6	Student	59	71,0	No knowledge	67	80,7
Male	8	9.6	25 – 30	25	30,1	Student + working	12	14,5	Some	11	13,3
			31 – 44	11	13,3	Graduated / working	12	14,5	Knowledge	5	6,0
Total	83	100		83	100		83	100		83	100

METHODS

Studies I & III

Questionnaire

Background characteristics. A questionnaire for mental health professionals was constructed in which the respondents were asked to report their age, gender and profession, how many years they had worked in their profession, the proportion of cases involving children of all the cases they worked with, how many cases of CSA they had been involved in and how familiar they were with relevant literature and research in this field and if they had attended courses or lectures concerning CSA. Furthermore, they were asked to evaluate their own expertise concerning CSA on a scale from 1 (no expertise at all) to 7 (broad expertise). Finally, the professionals were

asked to report their assessment of the prevalence of both child sexual abuse in general and incest for boys and girls, respectively.

As a preparatory step in the data analyses, a summary variable for theoretical competence was formed by the number of books and articles read and lectures and courses attended. Open-ended questions about what the participants had read were included to ensure that the literature was indeed relevant and scientific. Cronbach's reliability coefficient was .68 for this variable, which indicates acceptable reliability. Furthermore, discrepancy variables were formed for the estimations of prevalence of CSA and incest by computing the difference between each participant's answer and the prevalence reported in scientific literature (Sariola & Uutela, 1994; 1996).

Attitudes and beliefs concerning sexual abuse. The questionnaire also included ten items developed on the basis of the Juror Bias Scale (Kassin & Wrightsman, 1983) covering a range of beliefs related to CSA. The items of the Juror Bias Scale (Kassin & Wrightsman, 1983) were carefully analyzed and used for formulating items for the present study. This was done by changing the wording of the items so that they would apply specifically to CSA cases. Also, any items specific to the jury system were either discarded or reformulated to be also applicable to an inquisitorial criminal justice system. The resulting items were then read by two experts in the field of CSA investigations who assessed their relevance and readability. The final list of items can be seen in Appendix A.

Response alternatives were -2, -1, 0, +1, +2 with anchors "totally disagree" and "totally agree". The respondents were instructed to indicate their degree of agreement with the ten items. They were also instructed to answer the questions promptly without considering them for long. Items 6 and 9 were phrased so that an "agree" response indicated a pro-accused bias,

whereas for the rest of the items an “agree” answer indicated a pro-victim bias.

In Study III two summary scales based on a principal components analysis in Study I were used. The Pro-Child Beliefs scale had five items with scores ranging from a minimum of -10 to a maximum of 10. The Anti-Criminal Justice System Beliefs scale had four items with scores ranging from a minimum of -8 to a maximum of +8.

Case description. A description of a case of alleged CSA constructed on the basis of a real investigation material was attached to the questionnaire. The names and the facts of the case were altered to protect the individuals involved. In this case a 6-year-old girl, Tiina, of a socio-economically deprived family was observed behaving in a sexualized manner at the daycare and she was referred for further investigation. She was diagnosed as psychiatrically disturbed and also seemed to have problems with concentration. In the two interviews that followed she indicated that her father had sexually abused her by subjecting her to full penetrative sex on at least one occasion.

The materials included a police interview with the father in which he denies any wrongdoing, a police interview with the mother in which she also claims that the charges are wrong, a report to the police by a child psychiatrist in which the psychiatrist summarizes the findings of the investigations concluding that Tiina has been sexually abused, a report by the psychologist on the results of psychological testing and interviews with Tiina in which Tiina is reported to have told that her father has put his penis into her vagina. She is also reported to have described the involvement of the father by saying that “when daddy comes on top and [the child’s word for sexual intercourse] and I am there below then it hurts in my tummy”. The

materials also include a report from a pediatrician who has forensically examined Tiina for physical signs of sexual abuse. The results of this examination are inconclusive. A psychiatric nurse describes in a report Tiina's sexualized behavior on the ward. Finally, a report by child protection services describes the family situation (socioeconomic deprivation) and the background of the case (signs of physical abuse and sexualized behavior from the daycare).

The interviews. The case descriptions also included transcripts of two interviews with Tiina concerning the suspicions of sexual abuse (except in the condition in which no transcripts were presented). To manipulate stereotype induction, emotional tone, as well as threats and rewards, five sentences containing the technique in question were inserted in the questions and remarks of the interviewer in both the first and the second interviews. Altogether ten sentences for each technique were inserted. In order to ensure the authenticity of the manipulation, the sentences were constructed using transcripts from a number of actual CSA cases. Interested readers may contact the corresponding author for more information about the interviews.

Case evaluation. This part contained questions concerning the probability of the alleged abuse, whether the respondents would pursue the charges if they were the prosecutor, whether they would convict if they were the judge as well as questions concerning the likelihood of specific events Tiina mentioned in the interviews.

Design, Study III

The experiment was based on a 2 x 2 x 2 x 2 (leading questions: yes vs no) x (stereotype induction: yes vs no) x (emotional tone: pressure to respond vs no pressure to respond) x (threats and rewards: yes vs no) between-subjects factorial design. It included 16 different versions of the interview transcripts. There was an additional condition in which the material did not contain any interview transcript. Of the participants 214 received interview transcripts and 106 received no interview transcripts. Participants were assigned randomly to the different conditions with a ratio of 2 to 1.

As the number of respondents who had transcripts was 214, the power to detect main effects (with no attempt to analyze interactions as the power would not have been adequate here) was good (107 vs. 107) in each comparison. The additional condition of having participants who did not have any interviews (but instead relied on the statements made by mental health professionals) was not (and could not be) a part of the 2 x 2 x 2 x 2 -design as the variables could not take any values in this instance. Instead, the respondents were randomly allocated to either the transcript condition or into the no-transcript condition and the analyses concerning the no transcript condition were not part of the factorial design. However, as the number of professionals in the area in the whole country was limited, respondents were not assigned to these two conditions 50:50. Instead, more participants were assigned to the transcript condition with a ratio of 2 to 1. The rationale with the no transcript condition was that it presents a situation in which nothing is known about the actual interview techniques utilized.

In two pilot studies it was found that the sentences constituting the manipulation of suggestiveness of the different versions of the interview were equally and clearly noticeable and that evaluations of the case were not

constrained by neither a floor effect nor a ceiling effect which would exclude the possibility of finding significant effects of the experimental manipulation.

Procedure

A questionnaire was sent in cooperation with the appropriate professional organizations to samples of Finnish social workers and psychologists and all the child psychiatrists active in the field. Among social workers and psychologists specifically those who work with children were targeted, by sending the questionnaires to persons working in organizations and units with mainly child clients. A cover letter explaining the purpose of the study as investigating different professional groups' perceptions of CSA cases was attached to each questionnaire. The respondents were offered an opportunity to participate in a lottery draw with a travel abroad (worth €840) as the main prize. The questionnaire was sent to 335 social workers, 193 child psychiatrists and 336 psychologists. The response rate for social workers was 38% (n = 126), 31% for child psychiatrists (n = 60) and 40% (n = 134) for psychologists. Some questionnaires were returned with comments indicating that they had not been answered due to lack of time. The quite low response rate among social workers and psychologists may have been due to the fact that not all of these clinicians are involved with CSA investigations even though they work with children and only those actually working with these issues responded. Thus, the apparent response rate is probably lower than the actual portion of professionals working with children that responded to the questionnaire. There are relatively few child psychiatrists in Finland and the questionnaire was sent to all of them. The proportion of clinicians that do not encounter CSA cases (e.g. those with private practice) might have

been higher than in the other groups and the low response rate indicates that only those working with CSA cases answered.

Study II

Construction of the questionnaire

In part 1 of Study II the purpose was first to test and select efficient and valid items concerning attitudes to CSA and CSA investigations. A questionnaire was constructed based on popular and scientific literature on CSA, for example, Ferrara's Child Sex Abuse Attitude Scale (1999) and items developed on the basis of the Juror Bias Scale (Kassin & Wrightsman, 1983) covering the range of CSA related beliefs used in Studies I and III. A number of items were constructed based on information in news articles and from investigating shows and debates on TV. The last part of the items was constructed based on interviews with a police officer and a social worker, both working intensively with CSA and CSA investigations, and from discussion groups with psychology students at Åbo Akademi University and the University of Turku ($n = 20$). Two experts in the field of CSA investigations then read all items and assessed their relevance and readability.

The university departments of Turku, Helsinki, Tampere, Joensuu, Jyväskylä, Lapland, and Åbo Akademi were asked to distribute an attitude questionnaire about CSA consisting of 112 questions to the students. The questionnaires with a cover letter explaining the purpose of the study were sent by mail to the participating departments after a contact with a willing distributor among the staff members. This person then sent the completed questionnaires back to the authors by mail. Part of the questionnaires was distributed directly to students by the authors, for example during lectures. The preliminary questionnaire items were answered on a 6-point Likert-type

scale (1 = strongly disagree, 6 = strongly agree). Standard demographic questions were used to ascertain gender, age, and level of education.

Cronbach's reliability coefficient for the whole questionnaire was .88, indicating an acceptable level of reliability for the 112 items. Missing values were replaced with the mean value for that item. However, if one item had ten or more missing responses, it was deleted. Because of high inter-correlations in five item pairs, one item in each pair was randomly dismissed. Also, items that had a variable sum correlation of 0 were eliminated (mean value close to 1 or 6, that is. had very low or no variance). This procedure resulted in the final form of the Child Sexual Abuse Attitude and Belief Scale (CSAABS) and included 40 items (see Appendix A). The reliability analysis for the CSAABS revealed Cronbach's α of .92 reflecting a psychometrically acceptable internal consistency.

For the final 40 items, two experts on CSA independently evaluated all the items and categorized them as subjective or factual questions. The experts disagreed on only 3 items, for which they made the categorization together through discussion. The items that reflected attitudes were such that it was impossible to say whether they were right or wrong according to scientific research (e.g. "I often notice things that my colleagues do not notice when dealing with child sexual abuse", "I feel angry when thinking of child sexual abuse" or "In cases of child sexual abuse, the threshold to convict should be kept lower than in many other crimes"), whereas the items that consisted of factual statements were possible to determine as correct or false in the light of empirical research (e.g. "You can tell if a child has been exposed to suggestion", "It is not possible to make a child lie that abuse has happened" or "In 95 % of suspected cases of sexual abuse of children, the suspected perpetrators are guilty").

The 40 items (21 subjective and 19 factual) in the questionnaire were then answered on a 6-point Likert-type scale (1 = strongly disagree, 6 = strongly agree). Participants were asked to answer the questions promptly without considering them for too long. With the permission and support from the Finnish Psychological Association, an e-mail with an introduction letter was sent to all school and family counseling psychologists who were members of the psychologists' union, had given their e-mail address, and were currently working in Finland. These groups were chosen, as they were most likely to target clinicians experienced in working with children and potentially with suspected CSA cases. Recipients were asked to respond anonymously to the final form of the CSAABS, which was located on a password protected Internet site. The CSAABS took approximately 15 minute to complete after which the information was automatically e-mailed to the researchers without revealing the identity of the respondent. Third, an e-mail with an introduction letter was sent to all students of psychology at Åbo Akademi University and the University of Turku. Recipients were asked to respond anonymously to the CSAABS at a password protected Internet site. After they had completed the questionnaire, the information was automatically e-mailed to the researchers without revealing the identity of the respondent. To check test-retest reliability, the same procedure was repeated three weeks later. The first survey also included three vignettes describing cases of sexual abuse in order to get a preliminary assessment of the validity of the scale. For each vignette the participants were asked to indicate whether they believed that the child had been abused or not. It was hypothesized that strongly held attitudes would make the participants more prone to deciding that abuse had occurred. The first vignette described a case of a three-year-old girl who displayed sexualized behavior and signs of fear and anxiety in

the presence of the daycare nanny's husband. The second vignette described a six-year-old boy of divorced parents who lived with his mother and her new husband. The boy had asked his father to touch the boy's penis and explained that this was allowed. The boy had been interviewed five times but had not disclosed abuse. In this case, the participants were also asked whether the child should be interviewed again. The third vignette described a case of a five-year-old girl, who had been anxious and irritated, especially in the presence of her uncle. She had also showed some moderate sexualized behavior. For the third case the participants were told that the suspect, the uncle, had been acquitted in a court of law. They were then asked whether they thought that the court's decision had been correct. Those respondents that thought that the court's decision was false were asked to rate to what extent they experienced emotions of indignation on a scale with the following response options: "no" coded as 0, "yes, to some extent" coded as 1, and "yes, strongly" coded as 2. The belief scale was not analyzed separately in this sample, since it was not reasonable to assume that students would have detailed knowledge about CSA or CSA investigations. Finally, the participants were asked if they had attended a course in forensic psychology that is given at both universities. The course is empirically based and deals with both scientific research and application of scientifically based knowledge in forensic psychology. A part of the course deals with CSA investigations, and it was reasonable to assume that having taken part in such a course could be associated with the responses given to the attitude and belief items. Research in judgment and decision making has shown that empirically based knowledge is essential for good clinical judgment (e.g. Dawes, 1994) and, thus, we thought it reasonable to expect that having taken part in the course would be associated with less extreme attitudes.

Procedure

The 40 items (20 subjective and 20 factual) in the questionnaire were answered on a 6-point Likert-type scale (1 = strongly disagree, 6 = strongly agree) where the respondents were instructed to indicate their degree of agreement. They were asked to answer the questions promptly without considering them for too long. The participants were also instructed to report: (a) their age, (b) gender, (c) if they had children of their own, (d) how many years they had worked in their profession, (e) the proportion of work time involving cases with children, (f) hours of training in how to interview children and about CSA, such as attended courses or lectures, (g) how many cases of CSA they had been involved in, and (h) if they had any personal experience with CSA.

Altogether 242 out of 583 psychologists answered (40%). The response rate is comparable to other survey studies on sexuality (Fife-Schaw, 2000). For example, Rodriguez (2002) studied professionals' child abuse related attitudes and had a response rate of 39%.

Study IV

The questionnaire

The questionnaire consisted of 100 short case descriptions concerning children suspected of having been sexually abused and after each vignette questions concerning whether the abuse had more likely or not happened. The fictive vignettes were constructed using Kendall-Tackett and colleagues' (1993) research concerning the prevalence of different behavioral problems in sexually abused and non-abused children, Sandnabba and colleagues' (2003) and Santtila and colleagues' (2005) studies concerning the normative

sexual behavior of children in Finland, Kauppinen and colleagues' (2000) report on the prevalence of CSA cases reported to the authorities and Sariola's and Uutela's (1994; 1996) articles concerning the prevalence of both CSA and incest in Finland. The vignettes contained information about the children's behavior and other background factors (age, gender, the gender of the suspect and the suspect's relation to the child, who suspected abuse and why) in such a way that in 50 of the described cases the abuse had happened statistically more likely than not and in 50 of the cases it was statistically less likely that the abuse had happened than not. That is, when considering the base rate of abuse and the occurrence of symptoms in both non-abused and abused samples the probability of abuse exceeded .50 for those vignettes in which the correct decision was that the child in question had been abused. For example, in a vignette that described a 10 year old girl who was acting out in a sexual manner and in which the suspected perpetrator was not a member of the family the abuse had more likely occurred than in a vignette in which a father during a custody dispute was suspected of having sexually abused a biological 3 year old son and the son displayed symptoms of depression.

The participants could make four kinds of decisions regarding the cases. They could correctly decide that a child had been sexually abused, which is called a hit. They could falsely decide that a non-abused child had been sexually abuse, which is called a false alarm. Furthermore they could correctly decide that a child had not been abused, that is, a correct rejection, or they could falsely decide that an abused child had not been abused, in other words a miss. Since the vignettes described 50 cases of abuse and 50 false alarms, the maximum number of both hits and correct rejections respectively was 50.

Procedure

Because the questionnaire was so extensive, it was sent to the participants as attachments to e-mails. This ensured that the participants did not check the feedback before they had evaluated the case. The participants got the questionnaire in 10 e-mails each consisting of 10 vignettes. They got a new e-mail with feedback on the evaluated cases and 10 new vignettes when they had evaluated the previous ten vignettes. This way it was also possible to prevent the participants from getting tired, since they could only evaluate 10 cases at a time. After reading the vignettes the participants indicated for each case whether they thought the described child had been sexually abused or not. Since the participants had no previous experience with CSA evaluations it was possible to assess the effect of experience *per se*, that is the number of evaluated cases, as well as the effects of different kinds of feedback on the performance of the participants.

The participants were divided into four groups (see Table 5) in the order in which they contacted the researchers. Before answering the questionnaire, all participants got basic theoretical information about the definition of CSA according to Finnish law, the prevalence of CSA in different groups of children in Finland and about detecting and investigating CSA. This information was taken from the recently published guidelines for investigation of CSA in Finland (Taskinen, 2003). The idea was to give the same basic knowledge about CSA and CSA investigation that it is reasonable to assume that Finnish clinicians have today, whether they are experienced in the field or not. When answering the questionnaire one of the groups got so called cognitive feedback that is, they received information on decision making theory with special emphasis on CSA investigation and seven times during the answering process they were reminded of the information and got

concrete examples of how to solve the cases they read according to Bayesian reasoning. For example, the participants could be shown how to statistically estimate the probability of abuse with certain symptoms of the child by using Bayes's formula. One group received so called outcome feedback, that is for each case they had read they were simply informed whether they had made the right decision or not. One of the groups received both cognitive and outcome feedback and one group did not receive any feedback at all. Thus, the design included 3 independent variables: experience, outcome feedback and cognitive feedback.

Table 5
Design in Study IV

	Outcome feedback	No outcome feedback
Cognitive feedback	GROUP 1 N = 20	GROUP 2 N = 21
No cognitive feedback	GROUP 3 N = 22	GROUP 4 N = 20

STATISTICAL ANALYSES

In **Studies I and III** the data were analyzed with SPSS for Windows, version 10.0. First all missing cases were appropriately substituted. Since a One-Sample Kolmogorov-Smirnov test had revealed that the data were not normally distributed with the exception of age and years in the profession, the results of all the parametric analyses were double checked with appropriate

non-parametric tests. The results of these tests were similar with those of the parametric tests.

In **Study II** all data were analyzed with SPSS for Windows, version 11.0. First, all missing cases were appropriately substituted with the mean value for the item in question before further analyses.

In order to explore the basic dimensionality in part I of Study II, responses to all the 40 items (both attitudes and beliefs) of the CSAABS in both the first student sample and the professional sample were subjected to a preliminary principal component analysis extraction with varimax rotation. Principal component analysis involves a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called *principal components*. As all variability in the 40 items of the CSAABS was of interest, the principal component analysis model was chosen over the principal factors analysis, in which only the variability in an item that it has in common with the other items, is used (Tabachnick & Fidell, 1989, pp. 597-677). The results indicated a similar solution in both the student sample and the professional sample.

The results from these exploratory analyses were used to create a model that was tested with confirmatory factor analyses. The latter analyses were conducted using the software AMOS Graphics 5.0.1. Due to the fairly large sample and the possible limitations of the χ^2 -test (Thompson, 2004), the root-mean-square error of approximation (RMSEA) was used as the main indicator of model fit. The model was regarded as supported if PCLOSE was non-significant and the value of RMSEA was below 0.08 (Browne and Cudeck, 1993). Additionally, the model was considered to be supported if the minimum discrepancy divided by the degrees of freedom ratio (CMIN/DF) was in the range of 2:1 or 3:1 (Carmines, & McIver, 1981), and the goodness-

of-fit index (GFI) was greater than .80 (Arbuckle, & Wothke, 1999). Composite variables were formed on the basis of the results of the factor analyses by summing over the items belonging to each scale and dividing by the number of items.

In part II of Study II the number of missing values had a range of minimum 0 (0%), maximum 20 (8%) with the median at about 1.0 (3.5%). Due to the wide range of values in some of the background variables, these were categorized into four groups for some of the analyses. Responses to all the 40 items (both attitudes and beliefs) of the CSAABS were subjected to a principal component analysis extraction with varimax rotation in order to explore the basic dimensionality. The effects of the different background variables on all 40 items were tested with one-way ANOVAs and *t*-tests. Stepwise linear regression analysis was used to further examine which variables would predict the psychologists' CSA attitudes and beliefs. The beliefs were separately analyzed by computing frequencies and using stepwise regression.

In **Study IV** SPSS for Windows 11.1 was used for the statistical analyses. Kolmogorov-Smirnov-tests revealed that some of the variables were not normally distributed. However, the departures were small and transformations would have made it impossible to compare the variables with each other. The sample was also reasonably large ($N=83$) and, therefore, parametric methods were used for all analyses.

ETHICAL CONSIDERATIONS

In all four studies the participants were volunteers, who remained anonymous to the author. Those wanting to participate in the lotteries arranged for the participants could report their name and address, but this information was not combined with their answers.

RESULTS

Study I

Group differences in training, experience, self-evaluated expertise, and estimated prevalence of CSA.

The child psychiatrists were on average older than the psychologists and social workers and they also had most experience with CSA. Experience with cases involving children in general also varied between the groups with social workers having least experience and the child psychiatrists most. The child psychiatrists also evaluated their expertise regarding CSA to be higher than did the other professionals. Concerning the assessment of prevalence of incest among girls, the psychologists evaluated this prevalence to be lower than did the social workers and the child psychiatrists. All groups of professionals estimated the prevalence of CSA to be slightly higher than it is according to available research evidence (Sariola & Uutela, 1994; 1996).

Number of CSA cases and estimates of abuse base rates.

There was no clear connection between the number of CSA cases the clinicians had been involved in and their estimates of prevalence of abuse. Hence, no support for the existence of an availability bias was found.

The effects of training vs. experience on evaluation of own expertise and prevalence of CSA.

For all participants the strongest predictor for self evaluated expertise was the number of handled CSA cases. Theoretical competence was the second strongest predictor followed by proportion of cases involving children and last general professional experience. For the social workers, only the proportion of cases involving children predicted their self-evaluated expertise. Among the child psychiatrists, the number of CSA cases the professionals had handled was the strongest predictor for self-evaluated expertise. The next strongest predictors in this group were general professional experience as well as professional experience with children. For the psychologists the number of CSA cases was the strongest predictor of self evaluated expertise. The second strongest predictor was theoretical competence.

Differences between groups of clinicians on beliefs concerning CSA

Two separate scales measuring different aspects of CSA related attitudes and beliefs were found. The Pro-Child Beliefs scale (PCB) reflected pro-child attitudes with minimal emphasis on the rights of an accused person. Theoretically, high scorers on this scale are more likely than low scorers to discard problems and shortcomings in the investigation and to estimate the probability of abuse to be higher. The Anti-Criminal Justice System Beliefs

(ACJSB) scale reflected attitudes related to a negative view of the functioning of the criminal justice system (for example, punishments are not harsh enough and guilty persons may be set free as a result of questionable experts, lenient courts, and clever defense lawyers).

Social workers had higher scores on the PCB scale compared to psychologists and child psychiatrists. Child psychiatrists had higher scores on ACJSB scale compared to psychologists whereas social workers did not differ from either group.

The effects of gender

Male participants' scores on the PCB scale were significantly lower than those of female participants, that is, the male participants had less extreme attitudes that could lead them to ignore problems with CSA investigations. This effect remained when professional group was checked for.

Study II

Part I

When analyzing the 40 items that were included in both the student and the psychologist samples, the results from principal components analyses indicated a four-factor solution in both. The ten items loading on the first factor concerned problems with disclosure of CSA (e.g., 'Children must be interviewed repeatedly, because they are often afraid to admit abuse.'). Individuals scoring high on this factor emphasized disclosure by any means necessary and hence it was labelled the *Disclosure subscale*. The nine items loading on the second factor had a clear emphasis on believing in children's reports (e.g. 'Children do not fabricate stories about sexual abuse, because

they do not know anything about such things.’). Individuals scoring high on this factor ascribe to attitudes of unconditional belief in children and dismissal of the possibility of false allegations from children. It was labelled the *Pro-Child subscale*. The content of the eleven items loading on the third factor emphasized intuition. Individuals scoring high on this factor ascribe to attitudes favoring an intuitive approach to CSA investigations (e.g. ‘I trust my intuition.’). The factor was named the *Intuition subscale*. The fourth factor consisted of six items concerning the criminal justice system (e.g. ‘In cases of child sexual abuse courts often hesitate to convict the suspect.’). Individuals scoring high on this factor ascribe to negative attitudes towards the functioning of the criminal justice system in CSA cases. This was labelled the *Anti Criminal Justice System subscale*. The results suggest that the four subscales indeed reflect different types of attitudes and beliefs among the participants and that test-retest reliability for the CSAABS is good.

Predictive validity of the subscales

For the first vignette given to the participants in the third phase of part I, those who believed that the child in this case had been abused received significantly higher scores on each of the subscales than those who did not. For the second vignette, only the mean for the Anti Criminal Justice System subscale was significantly higher for those who believed the child to have been abused. The means for the other subscales were also in the same direction as in the first case. In this case, the participants were also asked whether the child should be interviewed again. It was hypothesized that the participants agreeing with this statement would have higher scores on the Disclosure subscale but not necessarily on the other subscales. In accordance with the hypothesis for the second vignette, the participants deciding that the

child should be interviewed again had significantly higher scores on the Disclosure subscale but not on the three other subscales.

For the third vignette, the participants who identified this child as abused had again higher means on each of the four subscales than those who did not identify the child as abused. However, the differences were only significant for the Pro-Child and Intuition subscales.

The participants were also asked whether they thought that the court's decision to acquit the defendant in the third vignette had been correct. It was predicted that those who thought that the decision of the court was false would score higher on the Anti Criminal Justice System subscale. In accordance with the hypothesis, those who thought that the decision of the court was false scored higher on the Anti Criminal Justice System subscale. However, these participants scored significantly higher on the other three subscales as well. These participants were also asked to indicate their degree of indignation over the court's decision. It was predicted that participants who had experienced such emotions would score higher on the Pro-Child subscale as it would be reasonable to assume that a decision in favor of the suspect and not the child would seem more unfair to participants who hold strong attitudes favoring the child. Stronger feelings were associated with higher scores on the Pro-Child scale as predicted. The association was also positive though weaker between experienced feelings of indignation and scores on the Intuition and ACJS subscales.

The effects of participation in the course in forensic psychology

The means of those who had participated in the forensic psychology course were lower on each of the four subscales. However, the differences

were significant only for the Disclosure, Pro-Child, and Intuition subscales, but not for the Anti Criminal Justice System subscale.

Part II

The male participants had handled more cases of CSA than had the female participants. No other differences in background factors were found.

Different types of CSA related attitudes and beliefs

Four different subscales of CSA related attitudes and beliefs were found. The Disclosure subscale reflected emphasis on disclosure of abuse by any means necessary. The Pro-Child subscale had a clear emphasis on unconditional belief in children and dismissal of the possibility of false allegations. The Intuition subscale reflected attitudes favoring an intuitive approach to CSA investigations and the Anti Criminal Justice System subscale reflected negative attitudes towards the functioning of the criminal justice system in CSA cases. A majority of respondents hold quite moderate attitudes and beliefs, but there were also those who held extreme beliefs.

Variables predicting child sexual abuse attitudes and beliefs

The more hours of training in interviewing children the participants had received, the stronger the attitudes and beliefs on the Disclosure subscale were. Stronger Pro-Child attitudes and beliefs could be predicted from participants' having received more training in interviewing children, being female, and having children of their own. For the Intuition subscale the participants with longer experience with working with children but also those who had more training in interviewing children as well as female participants trusted their intuition more. Finally, for the Anti Criminal Justice System

subscale, lack of faith in the criminal justice system was mainly related to participants being female, but also to having more hours of training in interviewing children than others.

Pre-existing beliefs

The most common misconception among psychologists was that children should be told in the interview situation that they are safe from the perpetrator. Over 70% of the participants agreed with this item. Also over 40% of the participants thought that children should be praised in the interview situation, that professionals do not use suggestive interviewing techniques and that play observation is a good technique for detecting CSA. The participants clearly agreed with incorrect answers or disagreed with correct answers on average on 20% of the questions. The quartile with the highest proportion of incorrect answers got at least 25% of the questions wrong, whereas the quartile with the lowest proportion of incorrect answers got no more than 10% of the questions wrong. Of the participants 4.5% gave no incorrect answers at all, whereas 2% answered 50% or more of the questions incorrectly. Again, the more education in interviewing children the participants had received the more erroneous beliefs they had.

Study III

Clinicians' sensitivity to specific suggestive interview practices

The overall probability of abuse and the likelihood of the different specific events (likelihood of having seen vaginal penetration, likelihood of having heard from father about vaginal penetration, likelihood of having had vaginal penetration, likelihood of father having made sounds during sex)

were judged to be lower when leading questions were used in the interviews compared to when they were not used. None of the other suggestive techniques had any effect on the assessments made by the clinicians. When leading questions were present it was less likely that the respondents thought that the case should be prosecuted or that the suspected father should be convicted. However, the clinicians were not sensitive to the presence of the other suggestive techniques.

When leading questions were present, the likelihood of thinking that the case should be prosecuted decreased to about a third compared to when leading questions were absent. What is more, when leading questions were present, the likelihood of thinking that the suspected father should be prosecuted decreased to approximately a half compared to when leading questions were absent.

Clinicians' awareness of the possibility of suggestive interviewing when no interview transcripts are presented

The number of suggestive techniques did not have a substantial effect on the clinicians' assessments and the absence of transcripts did not lead to substantially lower estimates on the questions regarding the case. Contrary to expectations, cases in which no transcripts had been presented were more similar to cases in which transcripts contained either none to two suggestive techniques than to cases which contained three to four suggestive techniques. Judgments that the case ought to be prosecuted and that the suspected father should be convicted were not significantly related to the presence or absence of the transcripts and the number of suggestive interview techniques present in the transcripts. If anything, the respondents were behaving contrary to the

hypothesis and evaluating the no transcripts condition to be more believable than the 3-4 suggestive techniques condition.

The effects of clinical experience, scientific competence, self-evaluated expertise, and beliefs about CSA cases on the sensitivity of clinicians to specific suggestive techniques

The results showed that the proportion of child cases, experience with CSA cases as well as self-estimated expertise all have a positive influence on the clinicians' sensitivity to suggestive influences. However, this effect was limited to only the probability estimate of abuse having occurred and is observable only when leading questions are included. In terms of the beliefs held by the clinicians, much clearer effects were observed regarding the estimated probability of abuse: those not holding extreme beliefs were sensitive to the suggestion whereas those holding such beliefs were not. Again the effect was limited to when leading questions were included in the number of suggestive techniques variable. Pro-Child Beliefs affected conviction decisions and Anti-Criminal Justice System Beliefs affected prosecution decisions in the same manner. Notably, however, those holding extreme beliefs on the ACJS scale were more likely to think that the suspect should be prosecuted when more of the suggestive techniques, excluding leading questions, had been used in the interviews. Similar trends were observable also in other instances for those holding extreme beliefs. This finding was totally unexpected.

The effects of attitudes and beliefs about CSA cases on the overall evaluation of the case

High scores on both the PCB and the ACJSB scale were associated with high evaluations of the likelihood of abuse having taken place and also with increased likelihood of the participant suggesting that the suspect should be convicted. The only exception was that ACJSB scale was not related to prosecution decisions.

The effects of base rate estimate of intra-familial sexual abuse of girls on the probability estimate

No connection between the participants' estimations of the prevalence of incest among girls and their final probability estimate of abuse in the case they evaluated was found, indicating that the participants did not take base rates into account when making their decision about the case.

Study IV

Background variables

Gender, previous knowledge about CSA, and the field of work or studies did not affect the performance of the participants.

The effects of outcome and cognitive feedback

In accordance with hypothesis 2, the performance of the participants was significantly better when outcome feedback was given than when no such feedback was given. Contrary to the hypothesis, no significant difference in performance was found for cognitive feedback. No significant effect of the combination of feedback compared to either type of feedback or

no feedback at all was found, even though the trend was that both kinds of feedback given together improved the performance more compared to other conditions. Thus, no significant support for hypothesis 3 was found.

The participants receiving only cognitive feedback made significantly less hits than the rest of the participants but had the largest amount of correct rejections. Cognitive feedback did not affect the decisions made if both cognitive and outcome feedback was given together. Those participants that did not receive any feedback at all had the largest number of hits but, on the other hand, made the smallest amount of correct rejections among all the participants. The second largest number of correct rejections was found among participants receiving both kinds of feedback, but the difference in decisions between those and the participants receiving only cognitive feedback was considerable. The effects of cognitive feedback on correct rejections were reduced when outcome feedback was given as well.

The effects of type of feedback and experience on decision making

In the present research experience was defined as the number of cases evaluated by the participants. Contrary to hypothesis 1, a general effect of experience was found: the participants performed significantly better in the second phase.

The hits of participants receiving cognitive feedback did not increase with experience, but when this feedback was not given experience had a positive effect on hits.

Participants receiving cognitive feedback were overall somewhat better in reaching correct rejections than those who did not receive cognitive feedback. In addition, the correct rejections increased with experience for

participants receiving cognitive feedback, but this did not happen if there was no feedback.

The combination of outcome feedback and experience had some positive effects on correct rejections: when about half of the cases had been evaluated, participants receiving outcome feedback had improved in making correct rejections, but the effect did not last till the end of the experiment.

The combination of both kinds of feedback and experience influenced both hits and correct rejections. Participants who received cognitive feedback made less hits than the other participants and this effect increased with experience. When outcome feedback was given together with cognitive feedback this effect disappeared and the number of hits increased.

When it came to correct rejections the effect was reversed. At the end of the experiment participants who received cognitive feedback made more correct rejections than the other groups and again the presence of outcome feedback removed this effect of cognitive feedback. However, the group that did not receive any feedback at all performed at the lowest level.

Over all it seems that outcome feedback together with cognitive feedback diminishes the effect of cognitive feedback. Cognitive feedback seems to reduce the tendency to make positive decisions, which increases with experience. This decreases the number of false alarms.

The performance of all groups of participants on hits and correct rejections respectively are shown in Figures 1 and 2.

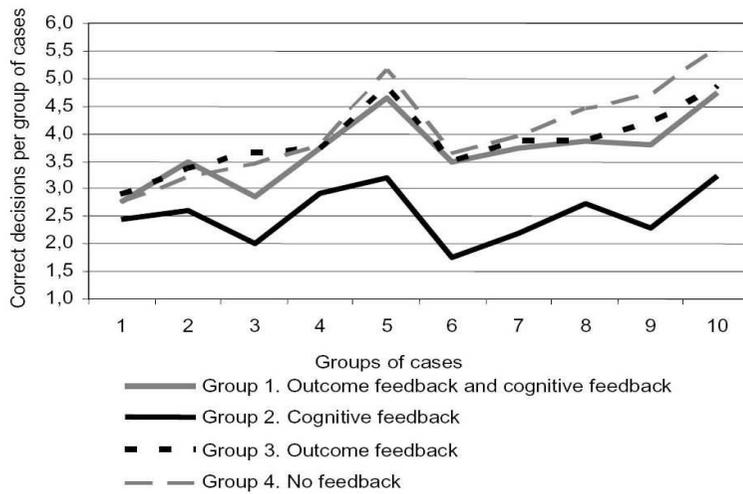


Figure 1. Hits per group of participants

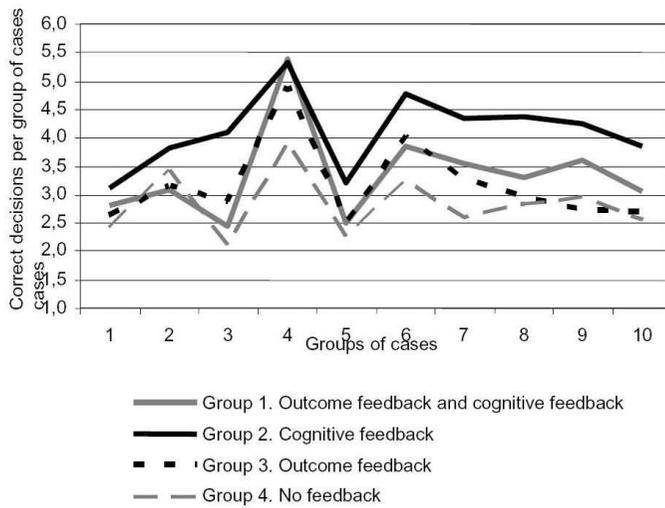


Figure 2. Correct rejections per group of participants

DISCUSSION

Expertise among clinicians conducting CSA investigations

The issue of expertise among clinicians involved in CSA investigations was explored from a number of different angles in the present thesis. The clinicians' own view on what enhances expertise as well as their own opinions on whether they themselves are experts in the field of CSA and CSA investigations were checked. Furthermore, through different questions about CSA and CSA investigations the clinicians' actual empirically based knowledge about this area was investigated. Finally, the clinicians also got to make decisions in a fictive case of alleged CSA and their ability to use base rate estimates as a ground for decisions as well as their sensitivity to suggestive influences in the interviews with the child were explored. The question about expertise was investigated on the basis of definitions of expertise drawn up by researchers in judgment and decision making. Factors influencing the clinicians' ability to reach correct decisions were also explored.

In Study I it was found that clinicians conducting CSA investigations rely more on their professional experience than on scientific, theoretical knowledge when evaluating their own expertise. As a group, clinicians in Finland are not particularly familiar with literature concerning CSA and CSA investigations. This is on the other hand understandable, as all the groups of professionals defined expertise in this field mainly as clinical experience with CSA. However, the clinicians' notion is not in accordance with the definition of expertise reported in scientific literature. For example, Cronbach and

Meehl (1955) as well as Guilford (1954) and Ghiselli and Brown (1955) emphasized that an expert should be able to recognize and avoid judgmental bias. What is more, Einhorn (1974) pointed out that experts should have a common core of knowledge, which leads them to making decisions that are in accordance with other experts in their field and in agreement with their own previous decisions when evaluating the same case a second time. Faust and Ziskin (1988), in turn, define expertise concerning expert witnesses in psychology and psychiatry as the ability to answer forensic questions with reasonable accuracy and the ability to help the members of the court to reach a more accurate conclusion than would otherwise be possible.

Earlier research suggests that the only way to avoid judgmental bias and to reach more accurate conclusions is by using theoretical knowledge and actuarial reasoning as the basis for decision making (Dawes, 1994; Fargason et al., 1997; Tversky & Kahneman, 1974). It appears that clinical experience is not an efficient way to gain such knowledge (Dawes, 1989; 1994), and clinical situations lacking feedback on performance have been identified as so called wicked learning environments, that is an environment that does not educate the decision maker to reach more valid conclusions (Gambrill, 2006; Hogarth, 2001). For this reason, the notion of the professionals in this research is somewhat disquieting. It is also alarming that the lack of theoretical knowledge is shown as unfamiliarity with the phenomenon of CSA: all the clinicians evaluated the prevalence of CSA among boys as higher than it is according to research results. Especially child psychiatrists, who evaluated their expertise as the highest among the groups of professionals, and the social workers, who in real life make many of the influential decisions concerning child care and child protective procedures,

evaluated the prevalence of incest among girls as significantly higher than did the psychologists.

Furthermore, when Study II analyzed to what extent the psychologists working with children held erroneous beliefs about CSA, it was found that some of these beliefs were in fact quite common. For example, over 70% of the participants thought that it should be expressed to the child in the interview situation that the child is safe from the perpetrator, even though this has been considered a suggestive technique that can influence children's reports negatively (Garven, Wood, Malpass, & Shaw, 1998). Furthermore, nearly 50% of the participants thought that children should be praised during the investigative interview, even though this also is considered to be a potentially suggestive technique. Also, more than 40% of the participants thought that play observation is a good method for detecting CSA and believed that professionals do not use suggestive interviewing, even though recent research suggests the opposite (Korkman, Santtila, Westeråker, & Sandnabba, 2008). In addition, when the amount of incorrect answers for each participant was checked, it was found that the participants on average gave incorrect answers for 20% of the questions and that only 4.5% of the participants did not answer any questions incorrectly. A minority of 2% gave incorrect answers to more than 50% of the questions. It is quite alarming that some misconceptions related to CSA investigations are this common among clinicians dealing with alleged CSA and that there are clinicians who possess a large amount of erroneous beliefs. This can result in serious mistakes being made in CSA investigations, both when it comes to methods used and decision making processes.

Finally, in Study III, when the participants had to evaluate a case of alleged CSA it was found that the hypothesis that clinicians would be

sensitive to suggestive influences overall received only limited support in the results. Specifically, only the presence of leading questions had a significant effect on the clinicians' responses regarding the case. As expected, such questions led to lower estimates of the probability of abuse and of the likelihood of several specific events touched upon in the transcripts. Also, they decreased the likelihood of the participants suggesting that the case should be prosecuted and that the father under suspicion should be convicted. These results show that the clinicians are either well aware of the empirical evidence (e.g. Bruck & Ceci, 1995; Garven & al., 1998) pertaining to the negative effects of leading questions on the reliability of children's testimony and that this affects their decision making, or that no special knowledge of empirical research is required in order to understand that leading questions may have such an effect. However, there was no evidence that the clinicians would have realized the potentially harmful effects of the other three suggestive influences on children's testimony. This is surprising as there is ample evidence that a number of suggestive interviewing techniques apart from leading questions may have such effects (Bruck & Ceci, 1995; Finniliä et al., 2003; Garven et al., 1998). This is an alarming finding, because if the clinicians do not recognize such influences as harmful, it would not be possible for them to take steps to avoid them when interviewing children themselves. Furthermore, clinicians in advisory roles could falsely state that suggestive interviews have been conducted appropriately. The finding also indicates that merely being a child professional is no guarantee for knowing how children should be interviewed and that assertion by such professionals to the effect that no suggestive methods were used in a particular interview should be evaluated with skepticism. These findings are in line with a number of studies concerning interviewing of alleged victims of sexual abuse

showing a number of shortcomings in the actual clinical practice of interviewers (Cederborg et al., 2000; Korkman, 2006; Lamb, Sternberg, Esplin, Hershkowitz & Orbach, 1997; Sternberg, Lamb, Davies, and Westcott, 2001)

The results pertaining to the hypothesis that the evaluations of the material without transcripts would be similar to the evaluations of the material including several suggestive influences were also surprising. In the majority of the analyses no significant effects were observed at all. For two estimates of the likelihood of specific events having taken place, significant effects were noticeable. However, closer inspection of the results showed that these contradicted the hypothesis. When no transcripts were included in the material, the estimates given were similar to estimates given when the material contained either none to two suggestive techniques. On the contrary, for one of the two effects these estimates were significantly higher than estimates given when the material contained three to four suggestive techniques. The conclusion is that when suggestive influences were not blatantly clear from available transcripts, the clinicians did not consider the possibility of such influences to be of any concern. In fact, they seemed to suppose that the interviews in such cases had been conducted appropriately. The clinicians probably trusted their colleagues to have acted appropriately and not to have used suggestive techniques in their interviews. Previous research has shown that memory for exactly how an interview happens and whether it includes suggestive influences is fallible (Bruck, Ceci, & Francoeur, 1999; Warren & Woodall, 1999). Taken together with the finding that the clinicians were not sensitive to three of the four suggestive influences, great care should be taken when cases with no transcripts of the investigative interviews are evaluated, for example, in a court of law.

In addition, in Study III, contrary to expectations, the clinicians' base rate estimate of incest abuse of girls was not related to the probability estimate in the case presented to them. This result shows that the basis for the clinician's judgment when evaluating the material was not the base line estimate of the phenomenon in question. In other words, Bayesian reasoning was not used as a tool for decision making. This result is not surprising since it has been shown that this way of making decisions is not usually the spontaneous strategy people use (Kahneman & Tversky, 2002; Tversky & Kahneman, 1974) and that clinicians conducting CSA investigations do not base their decisions on Bayesian reasoning (Wood, 1996).

In recent years some researchers have claimed that heuristics in human thinking can in fact be of use in judgment and decision making since in real life it is not possible to successfully make use of actuarial decision making, as situations oftentimes are too complicated and all necessary data is often unavailable (Gigerenzer & Engel, 2006; Gigerenzer & Todd, 1999). In this vein, Hogarth (2001) proposes that intuitive thinking can be developed and even consciously educated in so called kind learning environments through feedback on performance and that intuition thus is learned through similar processes as expertise and can be successfully used as a basis for decision making. In light of this view it could be argued that clinicians using their intuition could perform quite well as investigators of CSA, even better than an investigator who tries to reach a decision based on actuarial data that might be incomplete or too complicated to comprehend, even when using available formulas or rules. It follows that clinicians who rely on their experience with CSA cases and on their intuition when investigating CSA could in fact be skilled experts in the field. However, it should be noted that Hogarth's definition of intuition is based on the notion that successful

intuitive decision makers possess an unconscious body of knowledge that is in fact in line with scientific data and helps in reaching valid decisions. This kind of intuition has been developed in kind environments as opposed to wicked ones, that is, the criteria for learning from experience (Dawes, 1994) have been met. The results of this thesis showed that even experienced clinicians hold beliefs about how CSA should be investigated that are not in line with empirical data and are potentially harmful (e.g. that children should be repeatedly or suggestively interviewed) and make decisions that could have detrimental effects for families involved in CSA investigations (e.g. are more prone to decide that a case should be prosecuted when more suggestive techniques are used in the interview with the child). In other words, it can be concluded that the clinicians in this research had not obtained through experience the necessary knowledge for investigating CSA successfully and that intuitive thinking or heuristics did not serve as a useful alternative for conscious knowledge.

The results of this thesis have important implications for future training of clinicians investigating CSA: it is important not only to provide information about the prevalence and context of CSA and incest abuse, but also about the decision making process *per se* and actuarial judgment, so that clinicians understand how to apply research knowledge during the investigation (Edwards & von Winterfeldt, 1986).

CSA related attitudes and beliefs among Finnish clinicians

Since it has been documented in previous literature on decision making in forensic settings that strong attitudes and beliefs may influence the outcome of legal processes (e.g. McPherson, 1996), the pre-existing attitudes

and beliefs of the clinicians investigating CSA in Finland were explored in the present thesis. Based on previous research it is logical to assume that strongly held attitudes can overwhelm empirically based knowledge, and, thus, also affect expert judgment. In Study I a short instrument for detecting beliefs and attitudes related to CSA and CSA investigations was included and in Study III this was taken one step further by developing a 40 item Child Sexual Abuse Attitude and Belief Scale (CSAABS). In Studies I-III both the effects of strongly held attitudes and beliefs, and factors influencing such attitudes and beliefs, were explored.

In Study I two distinct sets of attitudes and beliefs were found, which were expected to be related to the way in which a person evaluates evidence pertaining to an ambiguous alleged CSA incident. The first set of items, which was labeled the Pro-Child Beliefs Scale, had a clear emphasis on protecting children and a minimal emphasis on the rights of an accused person. High scorers were expected to disregard problems and shortcomings in the evidence and estimate the probability of abuse to be higher and to be more prone to convict an accused person than low scorers. The other set of items, which was called the Anti-Criminal Justice System Beliefs Scale, reflected a negative view of the functioning of the criminal justice system: the punishments are not harsh enough and, furthermore, guilty persons may be set free as a result of questionable experts, lenient courts, and good defense lawyers. As a whole, it was found that both child psychiatrists and social workers scored more highly on the belief scales and thus held more extreme attitudes and beliefs than psychologists. Also, male participants received lower scores on the belief scales, that is, held less strong attitudes and beliefs than female participants.

In Study II somewhat similar results to those in Study I were found, concerning the attitudes and beliefs of the participants. As a 40 item attitude and belief scale was used in Study II, which was four times longer than the one used in Study I, it was possible to get a clearer picture of the attitudes and beliefs related to CSA among Finnish professionals. Four sets of attitudes and beliefs were found, two of which were similar to the ones found in Study I: individuals scoring high on the Disclosure subscale emphasized the importance of disclosure at any expense. Individuals with high scores on the Pro-Child subscale ascribed to believing in children and dismissing the possibility of false allegations from them. The Intuition subscale assessed whether individuals favored an intuitive approach to CSA investigations. Individuals who scored high on the Anti Criminal Justice System subscale ascribed to negative attitudes towards the functioning of the criminal justice system in CSA cases. Background variables explained only a minority of the differences in all of the CSA attitudes. However, some significant differences were found, which will be discussed below.

Finally, Study III showed that high scores on the two belief scales found in Study I were related to the evaluation of the case in the hypothesized manner, indicating that such a bias has similar effects in the context of CSA as has previously been shown for rape trials (Kassin & Wrightsman, 1983). The results show that holding strong beliefs related to CSA does indeed have an effect on the evaluation of evidence: such beliefs are probably related to the disregarding of problems and shortcomings in the evidence. Also, it may be that such beliefs are related to seeking to compensate for perceived problems in the criminal justice system. Decreased sensitivity to suggestive influences was expected from those scoring high on the two belief scales, the Pro-Child Beliefs Scale and the Anti-Criminal Justice System Beliefs Scale.

The results supported these expectations. However, an unexpected effect was found suggesting that those holding extreme beliefs on the two scales were actually more likely to think that the case should be prosecuted when there were many of the suggestive techniques other than leading questions present. This paradoxical finding might be caused by the clinicians being themselves influenced by the suggestive techniques used by the interviewer. It could be that the use of such techniques was interpreted by the clinicians as an indication of the interviewer being more certain of the accuracy of the allegation and that this affected the clinicians. The other explanation is that these techniques affected their representations of the accused person and the child's responses directly by suggestion.

Factors influencing expertise and attitudes among clinicians investigating CSA

Based on Study I it would appear that the profession of the clinicians influences their knowledge about CSA. Child psychiatrists and social workers evaluated the prevalence of incest among girls as significantly higher than did the psychologists. An explanation for this pattern could be that it is due to an availability bias effect: child psychiatrists in Finland have more real life experience with CSA and, thus an illusion of a higher prevalence has been created. This could, according to previous research, be avoided through scientific knowledge (Dawes, 1994; Fargason et al., 1997). In the case of the social workers, this effect could be due to the fact that social workers in Finland traditionally are the group of professionals whose responsibility it is to make sure that the CSA case is investigated. Therefore, their role is much more active than that of the other clinicians and they might be more involved

than the others in the cases they encounter. Hence, the possibility of an availability bias cannot be excluded.

Furthermore, in Study I it was also found that both the social workers and the child psychiatrists held stronger attitudes and beliefs than the psychologists. The social workers had the highest scores on the Pro-Child Beliefs Scale whereas the child psychiatrists had the highest scores on the Anti-Criminal Justice System Beliefs Scale. There are several plausible explanations for this result. First, it is possible that the fact that both child psychiatrists and social workers in Finland up until very recently have played a much more active part in the CSA investigations than the psychologists may have caused them to have more extreme opinions about such an emotionally charged phenomenon. Second, differences in training could be a further explanation. The beliefs on the PCB scale consisted of assertions that the psychologists and child psychiatrists might have agreed with simply due to their basic education in child psychology (e.g. children tell that they have been victimized although it is not true). An explanation for the negative view of the criminal justice system among the child psychiatrists could be that they have dealt with cases where they have become convinced of someone being guilty of CSA but not being convicted for reasons they do not understand or learn about. It is worth noting that in Finland, most CSA trials are conducted *in camera* which means that crucial evidence relating to a case may not come to the attention of the mental health professionals in charge of, for example, interviewing the child. Experiences like this could have led them to hypothesize possible reasons for these incidents in the functioning of the criminal justice system. Since in Finland it is the child psychiatrists' responsibility to testify in court in CSA cases, they are in closer contact with the legal system. Therefore, the effect could be especially strong among these

professionals. Being the person who is responsible for the investigation and reporting to the legal system might also lead child psychiatrists to having stronger feelings towards the justice system. This could in turn be reflected as an endorsement of the items. The psychologists, on the other hand, have only recently become more directly involved with the legal system in CSA investigations in Finland. The social workers do not have the same role as the child psychiatrist and, due to their training, might have a better understanding of how the justice system works.

Gender also seems to affect CSA related attitudes and beliefs among clinicians. In Study I the male participants held less extreme beliefs than the female participants. This could also be seen on the PCB scale where the female participants had significantly higher scores. Also in Study II female psychologists held significantly stronger pro-child attitudes and beliefs compared to male psychologists. Previous research also suggests that women might have more child-believing attitudes (Ford Peters, 2001; Gabora et al., 1993). One explanation to this could be that women in general are more prone to have stronger opinions when it comes to issues concerning children. These attitudes might reflect women's nurturing and care taking role in our society (Hubbartt & Singg, 2001). In addition, gender was the most important factor predicting anti criminal justice system attitudes and beliefs in Study II, and was a major factor predicting the Intuition subscale. This might be due to the fact that male participants were significantly more experienced in handling CSA cases. However, due to the highly unequal number of male versus female participants in Study II, generalized conclusions cannot be made based on these results.

According to Study II, the psychologists who had children of their own seemed to possess stronger pro-child attitudes and beliefs than the ones who

did not have children. These results do not support previous findings by Davey and Hill (1995), which might be explained by the differences in their sample. Davey and Hill (1995) also studied other professionals, such as medical professionals, social workers and police officers, and focused on the differences between CSA attitudes of different professionals.

Studies II and III explored how experience influenced expertise, decisions as well as attitudes and beliefs, since experience is frequently referred to as an important base for expertise and good decision making. In Study I it had also been found that clinicians themselves hold experience as the most important factor for enhancing expertise. In light of existing recommendations and clinicians' opinions, the results were surprising.

In Study II it was shown that no form of experience predicted the amount of incorrect answers on the CSAABS. The results suggest that experience does not correct misconceptions held by investigators of CSA. Furthermore, it was found that extensive experience in the area of clinical work with children results in having a more intuitive approach to CSA. For example, the experienced psychologists tended to agree more often than less experienced psychologists that they are able to see when a child is telling the truth. As Horner et al. (1993) concluded, clinicians with extensive experience in handling CSA cases might selectively use pieces of investigative evidence as a basis for rationalizing their intuitive perceptions. Also, research on judgment and decision making shows that when making decisions in clinical situations professionals might pay more attention to information that confirms their pre-existing beliefs and that this effect can even get stronger with experience through availability and hindsight biases (e.g. Dawes, 1994). Both the danger of using inappropriate suggestive methods when interviewing a child and the risk of making decisions on only a selection of

the available evidence might grow due to extreme intuitive attitudes, which might lead to false decisions about CSA.

The results of Study III did not strongly support the hypothesis that clinical experience and experience with CSA cases in particular, as well as scientific competence and self-evaluated expertise, would be positively related to sensitivity to suggestive influences. The proportion of child cases, experience with CSA cases as well as self-estimated expertise all had a small positive influence on the clinicians' sensitivity, but this was limited to sensitivity to leading questions. No enhanced sensitivity was observed for the other suggestive techniques. Also, the fact that the effect was only observed for one of the dependent variables suggests that it is not particularly strong.

Overall, these results are in line with earlier research on clinical experience (Dawes, 1994; Gambrell, 2006). However, it is alarming that a factor that has been seen as an invaluable tool for investigating CSA, probably strongly increasing the credibility of mental health professionals as witnesses in legal proceedings concerning sex crimes against children, appears to have quite limited positive effects on the decision making process of the clinicians involved in these investigations.

In the literature on judgment and decision making as well as on CSA investigations, training has been pointed out as the most important means for enhancing expertise and correct decisions in clinical work in general, as well as in CSA investigations. In Study II special attention was paid to the effects of training on attitudes and beliefs. Training was found to be a major factor explaining differences on the separate attitudes and beliefs scales. Hours of training in interviewing children seemed to be associated with higher scores on all subscales, even though the number of years working with children was a stronger predictor for the Intuition subscale and gender was the strongest

predictor for the Anti Criminal Justice System. Furthermore, it was found that training in interviewing children predicted the amount of erroneous beliefs that the participants held: the more training in interviewing children the participants had received, the more erroneous beliefs they held. These results were surprising, as training in theory could reduce the possibility of clinicians relying on their intuition when making decisions about CSA (e.g. Dawes, 1994) by, for example, helping them to be more aware of their own CSA attitudes, and help them to correct their false beliefs about CSA. However, in Study II it was not known in detail what kind of training the participants had received or by whom it had been provided. The results suggest that all training might not be effective for correcting erroneous beliefs, and thus improving decision making in CSA cases. On the contrary it can increase both extreme attitudes and erroneous beliefs about CSA and thus induce bias. Based on Study II it cannot be determined which components of training produced these effects, since it was not known what kind of training in interviewing the respondents had received. However, the fact that both high scores on the attitude and belief subscale and the amount of erroneous beliefs were predicted by training in child interviewing, the present results can be interpreted as showing a potential problem with the training of CSA investigators.

The effects of experience and feedback on non-experienced evaluators' decisions in CSA investigations

The issue addressed in Study IV was whether experience, cognitive feedback and outcome feedback or combinations of these would improve decision making among non-experienced evaluators of sexual abuse. In

previous research the effects of cognitive feedback have been somewhat contradictory – some studies show that it improves decision making whereas others do not show this effect (Balzer et al., 1989; Sengupta & Tarek, 1993). Some researchers have also emphasized outcome feedback as a condition for learning from experience (e.g. Dawes, 1989; 1994). The results of Study IV indicate that cognitive feedback and outcome feedback affected the decisions differently. Cognitive feedback seemed to make decision makers so cautious that it significantly increased the number of correct rejections, but also the number of misses. On the other hand, outcome feedback increased the number of hits but also the number of false alarms. However, this latter result should be interpreted cautiously, since the difference in decision making between those who received outcome feedback and those who did not receive feedback at all was so small. It could be argued that outcome feedback does not remove the tendency to make positive decisions, which occurs when no feedback is given and that therefore decision makers who receive outcome feedback make fairly similar decisions as those who do not receive feedback at all. This result is in accordance with previous research, which suggests that outcome feedback alone does not improve decision making in complicated tasks (Balzer, 1989; Brehmer, 1980), especially not if the evaluator possesses pre-existing beliefs and does not use a hypothesis testing approach (Garb, 1989). So called cognitive illusions may also prevent learning from outcome feedback, since positive feedback might be better remembered than negative feedback (Edwards & von Winterfeldt, 1986)

One of the hypotheses of Study IV was that the decision making of the participants would improve the most if they had as much information as possible, that is, received both kinds of feedback. However, the results showed that cognitive feedback did not affect the decisions when given

together with outcome feedback. One explanation could be that participants receiving outcome feedback tried to figure out the frequency of “abused”- versus “non-abused”-cases in the given material, even though the cases were given in a random order. For example, a participant could have thought that when he or she had got two cases with non-abused children the next one had to be a case with an abused child. This explanation, however, does not seem very plausible, since the performance of participants receiving outcome feedback differed so little from that of those participants who did not receive feedback at all, and, therefore, could not have used this strategy. A more likely explanation could be that when faced with both cognitive feedback and outcome feedback the participants relied only on outcome feedback since this was easier to comprehend. This explanation is supported by earlier research which suggests that mechanisms affecting decision making processes under uncertainty are so complex that people do not understand them spontaneously but rely on more easily comprehended rules for decision making (e.g. Edwards & von Winterfeldt, 1986).

Concerning the effects of experience per se, Study IV suggests that experience has some positive effects on the number of hits. The increased number of positive results could be explained by the so called availability bias (Dawes, 1994; Edwards & von Winterfeldt, 1986; Tversky & Kahneman, 1974). In other words, the more cases the participants were confronted with, the more they believed in abuse having occurred. The participants receiving cognitive feedback seemingly avoided this effect, since the number of positive decisions did not increase with experience for this group. In fact, for these participants the number of negative decisions increased with the number of cases read. As a result of this, the number of positive decisions remained small for this group throughout the experiment.

A similar effect of experience in combination with outcome feedback could be seen in the beginning of the experiment, but this effect did not last as the experience grew. Thus, even though cognitive feedback might decrease the availability bias, an opposite bias toward making too many negative decisions is created, possibly due to extreme cautiousness. On the other hand, the interaction of cognitive as well as outcome feedback and experience showed that outcome feedback together with cognitive feedback diminishes and even prevents the effects of cognitive feedback. Concerning the correct rejections the difference between participants receiving only cognitive feedback and the rest of the participants was clear especially in the second half of the experiment. When it comes to the hits, the difference clearly grew with experience. The results seem to reflect the limited capacity of human decision making: if large amounts of information and many clues for decision making are given, only a small part, which is easy to comprehend, influences the decisions made.

To some extent the results of Study IV contradict earlier research, which has indicated that outcome feedback is necessary for improving clinical decision making (Dawes, 1989). On the other hand the results are in accordance with research on cognitive feedback in complicated and uncertain situations (Balzer et al, 1989). The results showed that cognitive feedback in itself does not improve decision making, but can have detrimental effects in the form of a bias towards negative decisions. The implications for CSA investigations are of course serious: if too much emphasis is put on cognitive feedback, that is information about human decision making in general, this could lead to an increased number of false negative decisions, misses. In other words, a significant number of sexually abused children would be unidentified. On the other hand, if no feedback is given at all, which is often

the case in real CSA investigations, the results of this study indicate that a great number of innocent people could be accused of sexual abuse that has not occurred.

Methodological limitations

In Studies I-III low response rates can be seen as a potential problem. However, it should be noted that in previous research concerning such emotionally charged and sensitive issues as sexuality and sexual abuse of children, it has been found that response rates are generally lower than in other areas of research and that, therefore, lower response rates are acceptable (Fife-Schaw, 2000; Rodriguez, 2002). Furthermore, in Finland there are very few clinicians who work only with CSA investigations and thus all clinicians working with children may at some point work with these issues. It is therefore not possible to determine only on the basis of profession or workplace which child experts actually conduct CSA investigations or in some way get involved in these cases. In Studies I and III all child psychiatrists as well as a representative sample of all social workers and psychologists working with children were targeted and in Study II all psychologists working with children in Finland received the questionnaire. It is reasonable to assume that only those actually working with CSA participated, and, therefore, response rates among clinicians working with these issues may have been much higher than it seems. In Study IV, it was not possible to determine the response rate, but as a statistically representative sample was obtained and as the participants did not differ in terms of essential background factors, the actual response rate was not relevant for the study.

In Studies I, II and III rather limited measures of the participants' previous knowledge and training were obtained. Since training is emphasized as an essential tool for improving expertise it is important to further investigate what kind of training the clinicians conducting CSA investigations have received and to investigate in more detail what theoretical literature they have read. In Study II an important finding was that training in interviewing children may in fact cause more extreme attitudes and more erroneous beliefs. However, it was not possible to establish what kind of training the participants had received, and therefore more thorough investigations in this area are needed.

In Study III an important question concerning the practical applicability of the findings is the extent to which the results of the study are generalizable. Great care was taken to make the materials as realistic as possible by using a real suspected sexual abuse case as the basis. Also, the manipulation of the suggestiveness of the interviews was created by taking examples from other real suspected sexual abuse cases. The main objection against the results being generalizable is that the clinicians were aware of taking part in an experiment. The fact that their decisions in this case were not as consequential as real-life decisions they make for courts could have made them less careful in the way they evaluated the materials. It could be argued, however, that this problem is not as grave as it might seem at first. The fact that the clinicians were participating in an experiment where they knew the results would be published should have lead to great care being taken to make as careful a decision as possible. It should also be observed that the findings of the present research are applicable only to ambiguous cases in which the child's witness statement usually is the most central evidence in a criminal trial.

The shortcomings of Study IV lie in the fact that an experimental design was used, which in many aspects differed from real life situations. Even though the participants got more information about the described children than only reports of symptoms, the amount of information in real CSA investigations is nearly always larger. In addition, the interview of the child, which was not included in this study, is often the most important part of the CSA investigation. On the other hand, it must be remembered that a clear picture of the effects of feedback and experience on their own could not be obtained in a real life situation, where the factors influencing decisions are so complex that they cannot be isolated and measured. Also, it is thought that the same mechanisms direct human decision making in both simple and complicated tasks (e.g. Dawes, 1994). In addition, some clinicians may not rely on children's interviews when making decisions in CSA cases and all children, especially very young ones, cannot be interviewed in a reliable way and thus, the findings of the present study could be more applicable to CSA investigations than it initially may seem. Finally, it should be noted that the strength of this study also lies in the experimental design: in real life situations and using real clinicians, it would be impossible to isolate the variables explored in this study. The participants would be different in terms of experience and knowledge about CSA and objective outcome feedback could not be given regarding real cases.

CONCLUSIONS AND IMPLICATIONS

The main issues addressed in this thesis were a) whether clinicians investigating CSA rely more on scientific knowledge or clinical experience

when making decisions, b) whether clinicians hold attitudes and erroneous beliefs or lack essential knowledge concerning CSA and CSA investigations, which in turn influence the decision making process, c) what factors influence such attitudes and beliefs and d) whether experience and different kinds of feedback improve decisions in CSA investigations. The results were both surprising and alarming. Clinicians involved in CSA investigations do not seem to understand what factors enhance expertise and rely more on experience than on scientific knowledge. They hold many erroneous beliefs concerning CSA and many professionals have strong subjective attitudes to CSA. When making decisions clinicians do not rely on actuarial judgment and do not take base rate estimates into consideration. Furthermore, even experienced clinicians are not able to detect suggestive interviewing techniques and take this into account when deciding on whether a child has been abused or not. Surprisingly, experience and training do not correct these errors, but can rather make the situation even worse. Finally, when investigators get experience in the absence of feedback, which is often the situation in real life cases of alleged CSA, they seem to become biased towards making a positive decision about the suspected abuse.

The obvious conclusion that can be drawn from the present thesis is that this lack of expertise among clinicians both when it comes to CSA as a phenomenon and to decision making processes in CSA cases as well as knowledge about valid methods for investigating CSA, constitutes a great risk when clinicians assist the police in such cases. Earlier research has established that clinicians indeed use suggestive interviewing in CSA investigations (Korkman, 2006). To some extent this thesis offers explanations for these previous findings: it is quite understandable that clinicians do not use good interviewing techniques if they do not know what

kind of questions are in fact suggestive, and, on the other hand, if they possess attitudes according to which children should indeed be interviewed repeatedly with direct questions in cases of suspected CSA. This thesis also draws a somewhat more complicated picture of the problems concerning CSA investigations than research into forensic interviews with children has revealed. Even if clinicians were to use good interviewing practice, the conclusions of the case might still be highly biased if the clinicians use their uneducated intuition, do not understand the basics of actuarial judgment or emphasize disclosure of abuse at any cost. Especially the finding related to the absence of the interview transcripts in Study III supports this hypothesis. A possible explanation for this result is that clinicians simply do not base their decision in CSA investigations on the evidence of the interview with the child, even if this would be the only evidence at hand in an ambiguous case. The findings concerning the belief scales also show that it is possible to make crucial decisions even if the interview is highly suggestive and that crucial decisions can be made even though there is no interview at all. Could it be that the clinicians base their decisions on completely different reasoning than considering the evidence of the interview? For example, it is possible that the findings of the present study can be explained as a result of representative thinking, that is, the child's sexualized behavior was described in the material and it is possible that the participants saw this as representative of the category "abused child" despite the presence or absence of an interview. In studies of juror competence, researchers have found that when presented with complex trial material it is possible that jury members start using heuristics and subjective thinking instead of objectively considering the evidence when trying to reach a decision in a case (Honest, Levi, & Charman, 1998). Sex crimes against children are highly emotionally charged but often also only a

part of a complex case of a family in crisis. It is possible that the problems in the investigations of CSA originate even more from other biases in the decision making process than from the suggestive interviews.

It has been hypothesized that the quality of CSA investigations in Finland might in fact already be better than recent research suggests (Korkman, 2006), since new and better guidelines, as well as an empirically validated interview protocol, have been offered to CSA investigators in 2003 (Taskinen, 2003), that is, after the data for the most recent studies have been collected. Based on the present thesis there is reason to adopt a skeptical attitude towards this assumption. If the reason for the poor quality of CSA investigations lies in subjective attitudes and erroneous beliefs, it is not reasonable to assume that formal guidelines will correct the problem.

Another conclusion of the present thesis is that the emphasis on clinical experience that has been very strong in guidelines for CSA investigations does not seem to have a basis in empirical research. It is not self evident that experience *per se* enhances correct decisions in CSA investigations; on the contrary, experience without feedback can induce a bias. Also, the present thesis showed that the problem with feedback is in fact more complicated than was previously thought, since different kinds of feedback can potentially induce bias in different directions.

The results of the present thesis have implications for future training of CSA investigators. Based on previous research it has been recommended that investigators of CSA should get extensive training in interviewing children and continuous feedback on their performance (Korkman, 2006; Sternberg, Lamb, Orbach, Esplin, & Mitchell, 2001). Based on the present thesis it can be concluded that such training should focus not only on interviewing techniques but also on decision making processes and expertise, pre-existing

beliefs and subjective attitudes. In the present thesis it was found that those with extensive training in interviewing children held even stronger attitudes and more erroneous beliefs than those without such training. Therefore, it is of utmost importance to further investigate what sort of training might result in such effects before starting training programs with CSA investigators. Furthermore, the differential effects of different kinds of feedback should be taken into consideration when training clinicians in investigating CSA. Also, since extensive training seems to predict stronger attitudes and false beliefs, it would be important to investigate whether in fact some clinicians possess these attitudes and beliefs before they receive training and as a result of this embrace and interpret the training selectively. If this is the case, it could even be argued that to a greater extent than has been done previously, it should be determined beforehand whether some clinicians are suitable for investigating CSA in the first place. It might well be that some clinicians should not be given this task, since their orientation is better suited for, for example, therapeutic work. An attitude and beliefs scale such as the one developed in the present thesis could be used for selecting clinicians both for training programs and to work at units investigating CSA. All in all, since the results of the present thesis suggest that formal education does not necessarily correct false beliefs, it is reasonable to conclude that future training programs should focus on the changing of attitudes (i.e. focus both on cognitive and emotional components of attitudes and beliefs) rather than on solely providing information. Finally, courts should be more aware of the fact that simply being an expert on children, child psychiatry or developmental psychology does not automatically mean that an expert witness is an expert on CSA or CSA investigations or even suitable for conducting such investigations and testify in court.

SUGGESTIONS FOR FUTURE RESEARCH

More research is needed to investigate the effects of different kinds of training on beliefs and attitudes of CSA investigators. The nature of the training currently given to CSA investigators in Finland should also be examined. In the present thesis it was not known what kind of training the clinicians had received. It was also not known whether the clinicians had adopted extreme beliefs and false beliefs solely from the training or whether they had in fact possessed these attitudes and beliefs earlier, which due to this bias had become even stronger as a result of training. More research is needed to establish what sort of training is constructive for CSA investigators. Future research may also focus more thoroughly on how CSA attitudes and beliefs affect the clinical decision making process of CSA investigations, as well as how CSA attitudes and beliefs develop and change with experience and training, and, more specifically, how pre-existing attitudes may affect how information and scientific knowledge is received. In future studies, the CSAABS may possibly be used a valuable research tool. Furthermore, the reasoning and decision making of CSA investigators should be studied more in detail in the future than was possible here due to the limited information given by the questionnaires used in the present research. For example, it is not yet quite clear to what extent investigators of CSA rely on background data, for example behavioral symptoms and base rate estimates, as well as on the forensic interview with the child in question and how they integrate these information sources.

In the future, research on the effects of experience and feedback in CSA cases could focus on the effects of individual cognitive feedback, that is, individual feedback to each participant on how well they used decision

making strategies and on their reasoning in each individual case. It would be interesting to see if this would affect decisions in the same way as in the present thesis. Research on how attitudes and beliefs influence the use of feedback is also an area worth exploring further in the future.

There has been some discussion in the literature concerning the potential usefulness of heuristics and intuitive thinking in legal settings (Gigerenzer & Engel, 2006). The results of the present thesis suggest that intuitive thinking is not a good alternative to actuarial decision making in cases of alleged CSA. However, since CSA investigations most certainly represent situations in which it is difficult to obtain and calculate the effects of all necessary data, it would be of great help if there were other useful strategies for making valid decisions in these investigations. Nevertheless, considering the potentially detrimental effects of flawed decision making in cases of alleged child abuse, considerably more research is needed to determine whether there are any aspects of investigations of alleged CSA in which heuristics could be of use and furthermore, whether and how intuitive thinking can be developed so that decisions made by experts in CSA investigations would be more valid.

REFERENCES

American Professional Society on the Abuse of Children (APSAC). (1990). *Guidelines for psychosocial evaluations of suspected sexual abuse in young children*. Chicago, IL: American Professional Society on the Abuse of Children

- Balzer, W.K., Doherty, M.E., & O'Connor, R. Jr. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin*, *106*, 410-433.
- Baron, J. (1994). *Thinking and deciding*. 2nd ed. Cambridge University Press, NY, USA.
- Bays, J.& Chadwick, D. (1993). Medical diagnosis of the sexually abused child. *Child Abuse & Neglect*, *17*, 91-110.
- Berliner, L. & Conte, J. (1993) Sexual abuse evaluations: Conceptual and empirical obstacles. *Child Abuse & Neglect*, *17*, 111-125.
- Brehmer, B. (1980). In one word: Not from experience. *Acta Psychologica*, *45*, 223-241.
- Brigham, J. C., & Spier, S. A. (1992). Opinions held by professionals who work with child witnesses. In H. Dent & R. Flin (Ed.), *Children as witnesses* (pp. 93-111). John Wiley et Sons. Chichester.
- Bruck, M. & Ceci, S. (1995). Amicus brief for the case of *State of New Jersey vs. Michaels* presented by Committee of Concerned Social Scientists. *Psychology, Public Policy, and Law*, *1*, 272-322.
- Bruck, M., Ceci, S., & Francoeur, E. (1999). The accuracy of mothers' memories of conversations with their preschool children. *Journal of Experimental Psychology: Applied*, *5*, 1-18.
- Bruck, M., Ceci, S. J., Francoeur, E., & Renick, A. (1995). Anatomically detailed dolls do not facilitate preschoolers' reports of a pediatric examination involving genital touching. *Journal of Experimental Psychology: Applied*, *1*, 95-109.
- Bruck, M., Ceci, S., & Hembrooke, H. (1998). Reliability and credibility of young children's reports. From research to policy and practice. *American Psychologist*, *53*, 136-151.

- Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*, *113*, 403-439.
- Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. American Psychological Association. Washington DC, USA.
- Cederborg, A-C., Orbach, Y., Sternberg, K.J., & Lamb, M.E. (2000). Investigative interviews of child witnesses in Sweden. *Child Abuse & Neglect*, *24*, 1355-1361.
- Cheung, M., & McNeil Boutte-Queen, N. (2000). Emotional responses to child sexual abuse: A comparison between police and social workers in Hong Kong. *Child Abuse & Neglect*, *24*, 1613-1621.
- Collings, S. (1997). Development, reliability, and validity of the child sexual abuse myth scale. *Journal of Interpersonal Violence*, *12*, 665-675.
- Crohnbach L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281 – 302.
- Cronch, L., Viljoen, J., & Hansen, D.(2006). Forensic interviewing in child sexual abuse cases: Current techniques and future directions. *Aggression and Violent Behavior*, Vol 11(3), 195-207.
- Dammeyer, M.D. (1998). The assessment of child sexual abuse allegations: using research to guide clinical decision making. *Behavioral Sciences and the Law*, *16*, 21 – 34.
- Davey, R. I., & Hill, J. (1995). A study of the variability of training and beliefs among professionals who interview children to investigate suspected sexual abuse. *Child Abuse & Neglect*, *19*, 933-942.

- Davies, G. M., Westcott, H. L., & Horan, N. (2000). The impact of questioning style on the content of investigative interviews with suspected child sexual abuse victims. *Psychology, Crime and Law*, 6, 81–97.
- Dawes, R (1989). Experience and validity of clinical judgment: the illusory correlation. *Behavioral Science and The Law*, 7, 457 – 467.
- Dawes, R (1994). *House of Cards – Psychology and Psychotherapy Built on Myth*. Simon & Schuster Inc. NY, USA.
- Dawes, R., Faust, D., & Meehl, P. (1989). Clinical versus actuarial judgment. *Science*, 243, 1668 – 1674.
- DeLoache, J., & Marzolf, D. (1995) The use of dolls to interview young children: Issues of symbolic representation. *Journal of Experimental Child Psychology*, Vol 60(1), 155-173.
- Drach, K. Wientzen, J., & Ricci, L. (2001) The diagnostic utility of sexual behavior problems in diagnosing sexual abuse in a forensic child abuse evaluation clinic. *Child Abuse & Neglect*, Vol 25(4), 489-503.
- Edwards, W. & von Winterfeldt, D. (1986). On cognitive illusions and their implications. *Southern California Law Review*, 59, 401 – 451.
- Einhorn, H. J. (1974). Expert judgment: some necessary conditions and an example. *Journal of Applied Psychology*, 59, 562 – 571
- Einhorn H. J. & Hogarth R. (1978). Confidence in judgment: persistence of the illusion of validity. *Psychological Review*, 85, 395 – 416.
- Ellonen, N., Kääriäinen, J., Salmi, V., & Sariola, H. (2008). Lasten ja nuorten väkivaltakokemukset: Tutkimus peruskoulun 6. ja 9. luokan oppilaiden kokemasta väkivallasta [Violence against Children and Adolescents in Finland]. Poliisiammattikorkeakoulun raportteja 71/2008.

- Oikeuspoliittisen tutkimuslaitoksen tutkimustiedonantoja 87 [Police College of Finland, Reports No. 71. National Research Institute of Legal Policy, Communications no. 87]. English summary can be found at <http://www.optula.om.fi/44610.htm>
- Everson, M., & Boat, B. (1997). Anatomical dolls in child sexual abuse assessments: A call for forensically relevant research. *Applied Cognitive Psychology*, Vol 11(Spec Issue), 1997. pp. S55-S74.
- Fargason, C.A., Peralta-Carcelen, M.C., Fountain, K.E., Amaya, M.I., & Centor, R. (1997) Impact of uncertainty on non-medical professionals' estimates of sexual abuse probability. *Journal of Sexual Abuse*, 6, 49 – 63.
- Faust, D. & Ziskin, J. (1988). The expert witness in psychology and psychiatry. *Science*, 241.
- Ferrara, F. F. (1999). Validation of the child sexual abuse attitude scale through confirmatory factor analysis. *Structural Equation Modelling*, 6, 99-112.
- Festinger, L. (1957). *A theory of cognitive dissonance*. New York: Row, Peterson & Company.
- Fife-Schaw, C. (2000). Surveys and sampling issues. In G. M. Breakwell, S. Hammond & C. Fife-Schaw (Ed.), *Research methods in psychology* (pp. 88-104). London: Sage.
- Finkelhor, D., Hotaling, G., Lewis, I., & Smith, C. (1990). Sexual abuse in a national survey of adult men and women: Prevalence, characteristics, and risk factors. *Child Abuse & Neglect*, Vol 14(1), 19-28.

- Finnilä, K., Mahlberg, N., Santtila, P., Sandnabba, K., & Niemi, P (2003).
Validity of a test of children's suggestibility for predicting responses
to two interview situations differing in their degree of suggestiveness.
Journal of Experimental Child Psychology, 85, 32-49
- Finnish Psychological Association (2002). Information about the profession
for psychologists. *Psykologi*, 6, 19 – 27
- Ford Peters, D. (2001). Examining child sexual abuse evaluations: The types
of information affecting expert judgment. *Child Abuse & Neglect*, 25,
149-178.
- Friedrich, W.N. (1998). Behavioral manifestations of child sexual abuse.
Child Abuse & Neglect, 22 (6). 523-531
- Gabora, N. J., Spanos, N. P., & Joab, A. (1993). The effects of complainant
age and expert psychological testimony in a simulated child sexual
abuse trial. *Law and Human Behavior*, 17, 103-119.
- Gambrill, E. (2006). *Critical thinking in clinical practice: improving the
quality of judgments and decisions*. 2nd Edition. Wiley-Interscience,
John Wiley & Sons Inc., Hoboken, New Jersey, USA.
- Garb, H. N. (1989). Clinical judgement, clinical training, and professional
experience. *Psychological Bulletin*, 105, 387-396.
- Garven, S., Wood, J. M., Malpass, R. S., & Shaw, J. S. (1998). More than
suggestion: The effect of interviewing techniques from the McMartin
preschool case. *Journal of Applied Psychology*, 83, 347-359.
- Gergen, K. & Gergen, M. (1986). *Social psychology*. 2nd Edition. Springer-
Verlag. West Hanover, Massachusetts, USA.
- Ghiselli, E. E. & Brown, C. W. (1955). *Personnel and industrial psychology*,
McGraw-Hill, NY.

- Gigerenzer, G. & Engel, C. (2006). *Heuristics and the law*. Dahlem Workshop reports, MIT Press & Dahlem University Press. USA.
- Gigerenzer, G. & Todd, P. (1999). *Simple heuristics that make us smart*. Oxford University Press. New York, USA.
- Goldberg, L.R., Faust, D, Kleinmutz, B., & Dawes, R. (1991). Clinical versus statistical prediction. In Cicchetti et al (Eds.). *Thinking clearly about psychology: essays in honor of Paul E Meehl*, pp 173 – 264, University of Minnesota Press, Minneapolis, USA.
- Guilford, J. P. (1954). *Psychometric methods*. McGraw-Hill, NY.
- Hartman G. L., Karlson, H., & Hibbard, R. A. (1994). Attorney attitudes regarding behaviors associated with child sexual abuse. *Child Abuse & Neglect*, 18, 657-662.
- Helweg Larsen, K. & Bøving Larsen, H. (2002). Unges trivsel år 2002. En undersøgelse med fokus på seksuelle overgreb [The well-being of youth in 2002. An investigation focused on sexual abuse]. Statens Institut for Folkesundhed (SIF), København.
- Herman, S. (2005). Improving decision making in forensic child sexual abuse evaluations. *Law and Human Behavior*, 29, 87-120.
- Hetherington, J., & Beardsall L. (1998). Decisions and attitudes concerning child sexual abuse: Does the gender of the perpetrator make a difference to child protection professionals? *Child Abuse & Neglect*, 22, 1265-1283.
- Hewitt, S. (1999). *Assessing allegations of sexual abuse in preschool children – understanding small voices*. Interpersonal Violence: The Practice Series. SAGE publications. Thousand Oaks, CA, USA.
- Hewitt, S. & Arrowood, A. (1994) Systematic touch exploration as a screening procedure for child abuse: A pilot study. *Journal of Child Sexual Abuse*, Vol 3(2), 31-43.

- Hogarth, R.M. (2001). *Educating intuition*. University of Chicago Press, Chicago. USA.
- Honess, T.M., Levi, M., & Charman, E. A. (1998). Juror competence in processing complex information: implications from a simulation of the Maxwell trial. *Criminal Law Review*, 761 – 839
- Horner, T. M., Guyer, M. J., & Kalter, N. M. (1993a). The biases of child sexual abuse experts: Believing is seeing. *Bulletin of the American Academy of Psychiatry and Law*, 21, 281-292.
- Horner, T. M., Guyer, M. J., & Kalter, N. M. (1993b). Clinical expertise and the assessment of child sexual abuse. *Journal of the American Academy of Child and Adolescent Psychiatry*, 32, 925-931.
- Horowitz, S., Lamb, M., Esplin, P., Boychuk, T., Reiter-Lavery, L. & Krispin, O. (1995). Establishing ground truth in studies of child sexual abuse. *Expert Evidence*, 4, 42 – 51.
- Hubbart, S. A., & Singg, S. (2001). Attitudes toward perpetration of sexual abuse against children: A study of legal and social service oriented students. *North American Journal of Psychology*, 3, 243-252.
- Kahneman, D. & Tversky, A. eds. (2002). *Choices, Values and Frames*. New York, Cambridge University Press.
- Kassin, S. M. & Wrightsman L. S. (1983). The construction and validation of a Juror Bias Scale. *Journal of Research in Personality*, 17, 423-442

- Kauppinen, S., Sariola, H., & Taskinen, S., (2000). *Lasten seksuaalinen hyväksikäyttö sosiaali- ja terveydenhuollon viranomaisten tietoon tulleet hyväksikäyttöepäilyt 1.5.1998 – 30.4.1999*. [Child sexual abuse and suspected child sexual abuse that was reported to the authorities May 1, 1998 to April 30, 1999] STAKES Tilastoraportti 200/5 [National Research and Development Centre for Welfare and Health, Statistical report 200/5]
- Kendall-Tackett, K. A., Meyer Williams, L., & Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. *Psychological Bulletin*, 113, 164-180.
- Koocher, G.P., White, C.S., Sivan, A.B., Goodman, G.S., Friedrich, W.N., & Reynolds, C.R. (1995). Psychological science and the use of anatomically detailed dolls in child sexual-abuse assessments. *Psychological Bulletin*, 118, 199-222.
- Korkman, J., Santtila, P., Drzewiecki, T., Sandnabba, K. (2008). Failing to keep it simple: Language use in child sexual abuse interviews with 3-8-year-old children. *Psychology, Crime & Law*, Vol 14(1), pp. 41-60.
- Korkman, J., Santtila, P., & Sandnabba, N.K. (2006). Dynamics of verbal interaction between interviewer and child in interviews with alleged victims of child sexual abuse. *Scandinavian Journal of Psychology*, 47, 109-119.
- Korkman, J., Santtila, P., Westeråker, M., Sandnabba, K. (2008) Interviewing techniques and follow-up questions in child sexual abuse interviews. *European Journal of Developmental Psychology*, Vol 5(1), 108-128.
- Kuehnle, K. (1998). Child sexual abuse evaluations: The scientist-practitioner model. *Behavioral Sciences and the Law*, 16, 5-20.

- Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1998). Conducting investigative interviews of alleged sexual abuse victims. *Child Abuse & Neglect*, 22, 813-823.
- Lamb, M.E. (1994). The investigation of child sexual abuse: An interdisciplinary consensus statement. *Child Abuse & Neglect*, 18, 1021-1028.
- Leichtman, M. D., & Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology*, 31, 568-578.
- Lepore, S. J., & SESCO, B. (1994). Distorting children's reports and interpretations of events through suggestion. *Applied Psychology*, 79, 108-120.
- Lipian, M., Mills, M., & Brantman, A. (2004) Assessing the verity of children's allegations of abuse: A psychiatric overview. *International Journal of Law and Psychiatry*, Vol 27(3), 249-263.
- London, K., Bruck, M., Ceci, S.J., & Shuman, D.W. (2005). Disclosure of child sexual abuse. What does research about the ways that children tell? *Psychology, Public Policy, and Law*, 11, 194-226.
- Maynard, C., & Wiederman, M. (1997). Undergraduate students' perceptions of child sexual abuse: effects of age, sex, and gender-role attitudes. *Child Abuse & Neglect*, 21, 833-844.
- McPherson, S. B. (1996). Psychosocial investigation in death penalty mitigation: Procedures, pitfalls and impact. In G. Davies, S. Lloyd-Bostock, M. McMurrin, and C. Wilson (Eds.), *Psychology, Law, and Criminal Justice* (pp. 286-295). Berlin: Walter de Gruyter.

- Melton, G. B. (1987). Bringing psychology to the legal system. *American Psychologist*, 42, 488 – 495.
- Murphy, K. R. (1987). Detecting infrequent deception. *Journal of Applied Psychology*, 72, 611 – 614.
- Oikeuspoliittinen tutkimuslaitos [the National Research Institute of Legal Policy]. (2007). *Rikollisuustilanne 2007. Rikollisuus ja seuraamusjärjestelmä tilastojen valossa*. [The crime situation in 2007. Criminality and the consequence system in the light of statistics]. Can be found on the internet on <http://www.optula.om.fi/44467.htm>.
- Poole, D., Dickinson, J. (2005). The future of the protocol movement: Commentary on Hershkowitz, Horowitz, and Lamb (2005). *Child Abuse & Neglect*, 29, 1197-1202.
- Poole, D. & Lamb, M. (1998): *Investigative interviews of children – a guide for helping professionals*. Washington DC, American Psychological Association.
- Realmuto, G., & Wescoe, S. (1992). Agreement among professionals about a child's sexual abuse status: Interviews with sexually anatomically correct dolls as indicators of abuse. *Child Abuse & Neglect*, 16, 719-725.
- Rodriguez, C. M. (2002). Professionals' attitudes and accuracy on child abuse reporting decisions in New Zealand. *Journal of Interpersonal Violence*, 17, 320-342.
- Santtila, P., Korkman, J., & Sandnabba, N. K. (2004). Effects of interviewer phase, repeated interviewing, presence of a support person, and anatomically detailed dolls on child sexual abuse interviews. *Psychology, Crime and Law*, 10, 21-35.

- Sariola, H. (1990). *Lasten väkivalta- ja seksuaalikokemukset* [Experiences of violence and sexuality among children]. Lastensuojelun Keskusliiton julkaisu 85. Helsinki.
- Sariola, H. (2005). Onko lasten seksuaalinen hyväksikäyttö lisääntynyt? [Has the sexual abuse of children increased?] *Duodecim*, 20.
- Sariola, & Uutela (1994). The prevalence of child sexual abuse in Finland. *Child Abuse & Neglect*, 18, 827-835.
- Sariola, & Uutela (1996). The prevalence and context of incest abuse in Finland. *Child Abuse & Neglect*, 20, 843-850.
- Sengupta, K., & Tarek, K. (1993). Alternative conceptions of feedback in dynamic decision environments: An experimental investigation. *Management Science*, 411
- Sisäasiainministeriö [Ministry of the Interior], 2006. *Lapsen kohtaaminen poliisitoiminnassa ja esitutkinnassa*, Sisäasiainministeriön ohje SM-2006-02026/Ri-2. [Encountering Children in Police Work. Directive of the Ministry of the Interior SM-2006-02026/Ri-2.]
- Skinner, L. (1996). Assumptions and beliefs about the role of AD dolls in child sexual abuse validation interviews: are they supported empirically? *Behavioral Sciences & the Law*, 14, 167-185.
- Sternberg, K.J., Lamb, M.E., Davies, G.M., & Wescott H.L. (2001.). The memorandum of good practice: theory versus application. *Child Abuse & Neglect*, 25, 669-681
- Sternberg, K.J., Lamb, M.E., Orbach, Y., Esplin, P.W., & Mitchell, S. (2001). Use of a structured investigative protocol enhances young children's responses to free recall prompts in the course of forensic interviews. *Journal of Applied Psychology*, 86, 997-1005.
- Suomen laki [Finnish Law], 2008. Talentum. Helsinki

- Svedin, C. G. & Back, C. (2003). Varför berättar de inte? Om att utnyttjas i barnpornografi. Rädda barnen, Stockholm, Sverige [Why do they not tell? On being exploited in child pornography. Save the Children, Stockholm, Sweden]
- Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics*. 2nd ed. New York: HarperCollins Publishers, Inc.
- Taskinen, S. ed. (2003). *Lapsen seksuaalisen hyväksikäytön ja pahoinpitelyn selvittäminen*. STAKES oppaita 55. [Investigation of child sexual abuse and child physical abuse, National Research and Development Centre for Welfare and Health, guide 55]
- Thierry, K., Lamb, M., Orbach, Y., & Pipe, M. (2005) Developmental differences in the function and use of anatomical dolls during interviews with alleged sexual abuse victims. *Journal of Consulting and Clinical Psychology*, 73, 1125-1134.
- Thompson, W. C., Clarke-Stewart, K. A., & Lepore, S. J. (1997). What did the janitor do? Suggestive Interviewing and the Accuracy of Children's Accounts. *Law and Human Behavior*, 21, 405-424.
- Trocme, N. & Bala, N. (2005): False allegations of abuse and neglect when parents separate. *Child Abuse & Neglect*, 29, 1333-1345.
- Trute, B., Adkins, E., & MacDonald, G. (1992). Professional attitudes regarding the sexual abuse of children: Comparing police, child welfare and community mental health. *Child Abuse & Neglect*, 16, 359-368.
- Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124 – 1131.

- Wagner, W. G., Aucoin, R., & Johnson, J. T. (1993). Psychologists' attitudes concerning child sexual abuse: The impact of sex of perpetrator, sex of victim, age of victim, and victim response. *Journal of Child Sexual Abuse*, 2, 61-74.
- Warren, A., & Woodall, C. E. (1999). The reliability of hearsay testimony: How well do interviewers recall their interviews with children? *Psychology, Public Policy, and the Law*, 5, 355-371.
- Wellman, M. M. (1993). Child sexual abuse and gender differences: Attitudes and prevalence. *Child Abuse & Neglect*, 17, 539-547.
- Wood, J.M. (1996). Weighing evidence in sexual abuse evaluations: An introduction to Bayes' Theorem. *Child Maltreatment*, 1, 25 – 36.
- Wood, J. & Wright, L., (1995) Evaluation of children's sexual behaviors and incorporation of base rates in judgements of sexual abuse. *Child Abuse & Neglect*, 19, 1263-1273.
- Övreeide (2002). *Yrkesetik i psykologiskt arbete*. Studentlitteratur. Lund, Sweden