

See discussions, stats, and author profiles for this publication at:  
<https://www.researchgate.net/publication/233290219>

# Offenders' lies and truths: An evaluation of the Supreme Court of Sweden's criteria for credibility assessment

Article in *Psychology Crime and Law* · September 2012

DOI: 10.1080/1068316X.2010.548815

CITATIONS

8

READS

86

2 authors:



[Rebecca M. Willén](#)

Ronin Institute

15 PUBLICATIONS 103 CITATIONS

[SEE PROFILE](#)



[Leif A Strömwall](#)

University of Gothenburg

91 PUBLICATIONS 2,444 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Victim blaming in rape cases: Empirical studies of understudied situations

[View project](#)



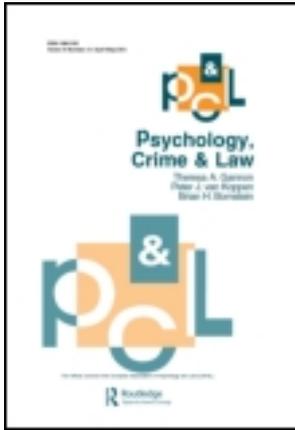
IGDORE [View project](#)

This article was downloaded by: [University of Gothenburg]

On: 04 November 2012, At: 02:41

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Psychology, Crime & Law

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/gpcl20>

### Offenders' lies and truths: an evaluation of the Supreme Court of Sweden's criteria for credibility assessment

Rebecca M. Willén<sup>a</sup> & Leif A. Strömwall<sup>a</sup>

<sup>a</sup> University of Gothenburg, Department of Psychology, Gothenburg, Sweden

Version of record first published: 27 Jun 2011.

To cite this article: Rebecca M. Willén & Leif A. Strömwall (2012): Offenders' lies and truths: an evaluation of the Supreme Court of Sweden's criteria for credibility assessment, *Psychology, Crime & Law*, 18:8, 745-758

To link to this article: <http://dx.doi.org/10.1080/1068316X.2010.548815>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## Offenders' lies and truths: an evaluation of the Supreme Court of Sweden's criteria for credibility assessment

Rebecca M. Willén and Leif A. Strömwall\*

*University of Gothenburg, Department of Psychology, Gothenburg, Sweden*

*(Received 25 June 2010; final version received 6 December 2010)*

The criteria used by Swedish courts for assessing credibility of plaintiffs' accounts were for the first time scientifically evaluated. Furthermore, unlike much previous deception detection research, we used offenders as participants instead of college students. False and truthful confessions by 30 offenders were analysed, and few significant effects were obtained. Truthful confessions were rated as having a higher degree of *clarity* than false confessions. Women's truthful confessions were rated as more credible than their false confessions. The offenders who were most experienced in being interviewed by the police gave a stronger impression of talking about something *self-experienced* in their false than in their truthful confessions; hence, it seems that offenders with more police interview experience have developed a kind of expertise in telling a convincing lie about crime. Overall, the criteria for credibility assessment used by Swedish courts had very limited usefulness in discriminating truthful and false confessions. A critique of the current status of evaluating statements in Swedish courts is provided.

**Keywords:** statement analysis; credibility assessment; true and false confessions; Supreme Court of Sweden; offenders

### Introduction

Evidentiary assessment may be extraordinarily difficult in certain types of criminal cases – when the only evidence available are the statements from the plaintiffs and/or defendants. Such cases are often sexual assaults and domestic violence. It is of importance that a defendant can be convicted on the basis of statements only; otherwise it would often be impossible to make a conviction at all (Gregow, 1996). The statement is furthermore often the only evidence in asylum procedures (Noll, 2005). Since the statement is fundamental in many cases there is a pressing need for a method to analyse the statement in order to assess its credibility. Judicial decision makers in courts around the world have developed several criteria to be used for such assessments.

The procedure for assessing the credibility of statements in Swedish courts has developed gradually between 1980 and 1996 through rulings in the Supreme Court of Sweden (SCoS) (Strömwall, 2010). The origin of these criteria is not clear but Schelin (2006) has argued that the SCoS probably has been influenced by another technique for statement analysis called *formal structure analysis*. The formal structure analysis was developed by Trankell mainly during the 1950s and 1960s and the trained

---

\*Corresponding author. Email: leif.stromwall@psy.gu.se

evaluators using the technique were quite often employed as expert witnesses in Swedish courts during the 1980s up until 1992. In a ruling in 1992 (NJA, 1992) the SCoS more or less banned experts using this particular method from appearing as expert witnesses in Swedish courts. This was the end of years of controversy and criticism regarding the method used by the expert witnesses. The method was for instance criticized for only reflecting common sense, and the courts were considered able to practise statement analysis without outside help (Strömwall, 2010). Diesen (2008) has argued that the courts are practising very complex psychological techniques for assessing credibility without any proper psychology training.

### *Statement analysis*

Techniques for assessing the veracity of statements were developed in parallel in Germany and Sweden during the 1960s and 1970s. The techniques were to be used primarily on the plaintiffs' statement in criminal cases concerning child sexual abuse. The German technique was further developed and became the method called Statement Validity Assessment (SVA). The core part of SVA, Criteria-Based Content Analysis (CBCA), are 19 criteria that should occur more frequently in truthful than in deceptive statements (Köhnken, 2004). SVA is today used as evidence in child sexual abuse trials in Germany, and is allowed as evidence in some other countries (Vrij, 2005).

Another technique for statement analysis is Reality Monitoring (RM). RM was originally developed by Johnson and Raye (1981) in order to separate one's own real and false memories. In the beginning of the 1990s Alonso-Quecuty (1992) suggested that RM might be useful also to detect deception in others. RM has similarities to CBCA since it consists of a list of criteria, but RM is still under development and not used in practice. Both techniques reach overall accuracy rates of above 70% when scientifically evaluated (see Vrij, 2008, for the most recent review of both methods).

The Swedish technique, formal structure analysis, had a wider scope of use than the German counterpart SVA; it was used not only in cases concerning child sexual abuse but also in for example arson cases (Trankell, 1965). The core of Trankell's method was a list of reality criteria which were supposed to occur more often in truthful statements than in false. In addition, there was a list of control criteria serving as complements to the reality criteria. Several criteria concerned different types of details, for instance superfluous details; details that were not necessary to mention in order for the interviewer to understand what had happened (Trankell, 1967). The Swedish technique was never scientifically tested (although it might have been based upon the scientific knowledge of its time) before its use was for all intents and purposes stopped in 1992 (Strömwall, 2010). As mentioned, it is quite possible that the criteria presently used by Swedish courts have been influenced by Trankell's method (Schelin, 2006), since there are several similarities.

### *The criteria used by Swedish courts*

Swedish courts mainly use statement analysis when assessing the credibility of statements in cases concerning child sexual abuse and domestic violence. The Swedish district courts referred to at least one such criterion in 60% of 268 rulings in sexual abuse cases between 1994 and 2003 (Schelin, 2006). Schelin (2006) went

through rulings made by the SCoS and identified the criteria the court did express as basis for their rulings. The set of criteria is not a fixed set; they are derived from several rulings by the SCoS (Strömwall, 2010). When or how a criterion is used is decided from case to case by the District court (Sutorius & Kaldal, 2003). Schelin identified five reality criteria and six control criteria. Furthermore, specific circumstances shall be considered when the plaintiff is a child. The reality criteria are supposed to occur in truthful statements while the control criteria are supplementary to the reality criteria.

*Reality criteria.* (1) The account should be *long*. This criterion is referred to in several rulings by the SCoS, but the criterion is never further described. (2) The account should be *coherent*. Schelin (2006) suggests that this criterion should be understood as a logic statement which is 'held together in a natural way'. The application of this criterion in the District courts shows that it sometimes is given the meaning of a narrative told in chronological order. (3) The account should be *clear*. Further definitions are not given by the SCoS. (4) The account should be *detailed*. The type of given details are important although not described further than for example reports of conversations and 'trenchant' details. (5) The account should be *consistent* between different statements. Changes in the account between different occasions may decrease the credibility (Schelin, 2006).

*Control criteria.* (1) The account should not contain *difficult-to-explain points*. Implausible points, strong exaggerations, contradictions or fabrications in a statement or other reasons to doubt the information decreases the credibility. (2) The plaintiff must fulfil her or his *burden of explanation* if necessary. The plaintiff must be able to give a good explanation of any weaknesses in the account. The credibility of a statement is strengthened if such explanations are given; hence, weaknesses in a statement may not be detrimental for the credibility as long as they can be well explained. (3) It is important *how the statement was given*. It strengthens the credibility if the account was given in a serious, thoughtful or spontaneous manner or with credible gestures and facial expressions. Non verbal cues are considered important. (4) The account should give a *credible impression* and (5) an *impression of self-experience*. These two criteria are based on common sense and intuition. They are overall judgements of the statement and therefore arguably the two most influential criteria. (6) The credibility of the whole statement is affected negatively if some information turns out to not be true; hence, the account is then considered to be *contaminated*. This criterion also works in the opposite direction: the credibility increases when specific information in the statement can be verified (Schelin, 2006).

The most commonly found criteria in the District courts in sexual abuse cases between 1994 and 2003 were *impression of credibility* and *self-experience* and quantity of *details*. It was also common to refer to *coherence*, *clarity*, *consistency* and *difficult-to-explain points*. The *length* of a statement was never found (Schelin, 2006).

Previous deception research shows support for some of the criteria: Truthful accounts tend to be *longer* (Sporer & Schwandt, 2006) and more *detailed* (Vrij, 2005) than deceptive accounts. *Clarity* has received support when it is further defined as clear, sharp and vivid in contrast to vague (Vrij, 2008); truthful statement then tend to be more clear than false (Masip, Sporer, Garrido, & Herrero, 2005). However, deception research does not support that truthful statements would be *told in chronological order* or be more *coherent* (Vrij, 2005) than false accounts. In addition,

Granhag, Strömwall, and Jonsson (2003) have found that false statements can be consistent to at least the same degree as true statements. *Nonverbal cues* have generally been found to not be reliable cues in veracity assessments (DePaulo et al., 2003; Vrij, 2008). No research has yet been carried out regarding *difficult-to-explain points*.

Impressions of *credibility* and *self-experience* are overall judgements closely related to the direct question on whether the plaintiff is lying or not (Schelin, 2006). The meta-analysis by Bond and DePaulo (2006) showed that people's ability to detect deception is not much better than chance level even when they are supposed to be experts in lie catching (e.g. police detectives).

### ***Credibility criteria used in asylum decisions***

Apart from court decisions in criminal cases such as child sexual abuse, assessments of credibility are crucial in other legal decisions. Perhaps the most important area of application is the decision to grant asylum or not to an applicant where the consequences for the individual in worst cases may be torture or death. The basis for judgments in these cases is primarily the asylum seekers' statement (Noll, 2005). Most of the rejections made tend to be due to a 'lack of credibility'. The UN High Commissioner for Refugees (UNHCR, 1992) tried to make the credibility assessments more precise in order to protect asylum seekers from arbitrary judgements (Kagan, 2003). For instance, it is written in the UNHCR handbook (1992), p. 33: 'The applicant's statements must be coherent and plausible, and must not run counter to generally known facts'. According to Kagan (2003) a coherent statement should be told in chronological order. Kagan has created a list of criteria used for determining credibility of a statement: the statement should be *detailed and specific* (contrary to vague), *consistent* (not involve contradictions) and *plausible* (not implausible). In addition the applicant should *reveal all key facts early* (not delay the revelation of key facts). Kagan (2003) argues that a list of criteria is an important step towards objective credibility assessments in refugee status determination. In the UK, the Home Office has set up Asylum Policy Instructions (Home Office, API, 2010), in which the credibility assessment is made of a combination of internal consistency, external consistency and a judgement of plausibility; see Sweeney (2009) for a critique of these guidelines. Despite its importance, credibility-based decisions in refugee and asylum cases are frequently based on personal judgement that is inconsistent from one adjudicator to the next, unreviewable on appeal, and potentially influenced by cultural misunderstandings. Some of the people who need protection most are especially likely to have trouble convincing decision-makers that they and their reasons for asylum should be believed (Kagan, 2003; Millibank, 2009).

### ***Type of lies and truths***

Most deception detection and statement analysis research have employed truth-tellers performing a non-criminal act and compared their behaviours and statements with deceivers who deny having made any transgression. However, in real-life cases people's lies are not only denials, but could for example take the form of a false confession, especially when trying to protect someone else or voluntarily taking the

blame in organized crime circles. An Icelandic study showed that 12% of almost all prisoners being incarcerated in Iceland (during a period of one year) reported to have made a false confession at least once (Gudjonsson, 2003). Sigurdsson and Gudjonsson (1996) found that almost 1% of the prisoners claimed that they currently were incarcerated for a crime they had falsely confessed to. Ofshe and Leo (1997) points out that the frequency of false confessions may be underestimated; the known cases may be the tip of an iceberg. Common reasons for giving a false confession have been reported to be perceived pressure from the police, protection of someone else and to avoid detention (Sigurdsson & Gudjonsson, 1996). Researchers and practitioners have, according to Gudjonsson (2003), neglected those suspects who make false confessions in order to protect someone else.

The possibilities of detecting false confessions through the use of statement analysis have not yet been studied. We believe this is the first such study.

### ***The present study***

Only two published studies (Bond & Lee, 2005; Lee, Klaver, & Hart, 2008) have employed offenders (i.e. persons with criminal experience) as participants in statement analytic studies. Since offenders' statements are evaluated in criminal courts, it is of importance to examine if the suggested credibility assessment techniques are applicable to offenders' statements and not confined to the plaintiffs' or the witnesses' statements.

The present study intended to test the SCoS's criteria for assessing the credibility of statements. The criteria that were tested in this study were: *length*, *coherence*, *clarity*, *details*, *difficult-to-explain points*, *credible impression* and *impression of self-experience*. In line with the assumption that the criteria could discriminate between truthful and false statements, it was hypothesized that truthful accounts would be *longer*, more *coherent*, *clear* and *detailed* than deceptive statements (Hypothesis 1a). In addition it was expected that deceptive accounts would contain more *difficult-to-explain points* (Hypothesis 1b). True statements were hypothesized to give a stronger impression of *credibility* and *self-experience* than false accounts (Hypothesis 1c).

People in general do not perform much better than chance level when detecting deception (Bond & DePaulo, 2006). Offenders seem to actually perform better although being lie biased (Bond, Malloy, Arias, Nunn, & Thompson, 2005; Bond, Thompson, & Malloy, 2005; Hartwig, Granhag, Strömwall, & Andersson, 2004). A reason for their skills may be that they receive valuable feedback through their criminal lifestyle. This assumption is referred to as *the feedback hypothesis* (Hartwig et al., 2004; Vrij & Semin, 1996). In line with the feedback hypothesis it was expected that the offenders with more police interview experience would tell more convincing lies than offenders with less such experience (Hypothesis 2).

## **Method**

### ***Participants***

Thirty offenders serving time in Swedish prisons participated: nine women ( $M = 39$  years,  $SD = 11$ ) and 21 men ( $M = 32$  years,  $SD = 8$ )<sup>1</sup>. All prisons were classified as low security prisons. The men's experiences of crime were typically violence

crimes, drug offences and theft. The typical crime experience of the women involved drug offences, theft and fraud. The participants' experience of police interviews ranged from 4 to over 200, with an average (median) of 23 police interviews. Time served in prison ranged from one month to 15 years ( $Md = 22$  months). Participation was voluntary. All participants kept their daily wage from the correctional treatment facility and received a small additional compensation to the amount of about 6€.

### **Questionnaire**

Participants answered a short questionnaire before being interviewed. Three questions regarded their police interview/interrogation experience as suspect, witness and plaintiff and were answered by specifying the number of interviews experienced or choosing the alternative 'never'. An open question regarding crime experience allowed the participants to write down what type of crime(s) s/he had some time committed. There were also questions about time served in prison and the participants' age and gender.

### **Procedure**

The participants were verbally informed about the study and its procedure during a meeting, and they also received written information. It was emphasized that the truthful statement must be all true (i.e. not contain distortions of any kind) and that the false statement must be all false (i.e. not contain any true parts)<sup>2</sup>. The written information included the most important points about the procedure: (i) in an interview to tell the truth about a crime s/he has committed and been convicted for, (ii) in another interview telling a lie (falsely confess) to a crime which s/he has never committed, (iii) in both interviews doing her/his best to convince the interviewer that s/he is telling the truth. Participants were also informed both verbally and in writing about anonymity and ethics.

Each participant was interviewed twice and each interview lasted about five–10 minutes. Half of the participants began with their truth, and the other half began with their lie. Before each interview the participant received instructions from a research assistant telling the participant if s/he was going to lie or tell the truth in the upcoming interview. The interviews were carried out in an adjacent room by the experiment leader. Participants arrived in pairs to adjacent rooms inside the prison. One participant was interviewed while the other participant received instructions from the assistant.

The assistant gave out crime scenarios from a pre-prepared list. Participants did not know the topic of their lie until they were handed out the scenario five–10 minutes before telling it. The selection of scenarios was carried out with regard to: (i) what type of crime the participant chose to tell the truth about; our aim was to not give out more severe crime scenarios than the participant oneself chose to talk about, (ii) what time interval the truth ranged; hence, we regarded whether the crime was one isolated incident or lengthy (e.g. several months).

Each interview started out with the interviewer pointing out that she did not know whether the participant was going to lie or tell the truth in the immediately upcoming interview. The participant was reminded to give the statement in a

convincing manner independently of the objective veracity status. After a free recall part the interviewer used at least two and never more than five (depending on how much had been said in the free recall) open questions/prompts in order to elicit a complete statement: 'Can you tell me more about how the crime was carried out?', 'Do you remember anything else?', and 'What happened before/after the immediate crime incident?' (for example). All interviews were audio taped.

### **Data preparation**

The interviews were transcribed and de-identified. When judicial consequences of the crime (e.g. detention or trial) had been mentioned it was deleted and hence not considered in the analyses. A three-point rating scale (0–2) was used for the ratings of the six reality criteria and the *difficult-to-explain points*. For the rating of *credible impression* and *impression of self-experience* a six-point rating scale (1–6) was used.

All criteria were defined as described above. *Credible impression* and *impression of self-experience* were further distinguished; *credible impression* was fulfilled when the event in the statement was considered to be all authentic (however, not necessary self-experienced), while a fulfilled *impression of self-experience* required that it actually had to be the speakers' own genuine experience.

All 60 statements were rated by two independent coders. None of the coders had received any specific judicial training in how to use the criteria, but had experience in conducting both CBCA and RM ratings. The criteria and ratings were discussed and disagreements solved during a conference between the two coders. Despite this, the initial inter-rater reliability was low (Pearson correlations ranged from  $r = 0.24$  to  $0.74$ ). Since the inter-rater reliabilities were generally low, we decided to average both coders' scores for each statement and each criterion. This had the advantage of creating a five-point scale (0–0.5–1–1.5–2), which in turn made possible the use of parametric analyses.

## **Results**

### **Order effects**

To examine possible order effects, a MANOVA was carried out. There was no multivariate effect,  $F(7,22) = 0.86$ ,  $p > 0.05$ ,  $\eta^2 = 0.22$ , or any univariate effects. Hence, it did not matter if the lie or the truth was told first.

### **The discriminating power of the SCoS's method**

A MANOVA was carried out to test Hypotheses 1a, 1b and 1c. The SCoS's criteria were the dependent variables and Authenticity the independent variable. There was no multivariate effect,  $F(7,23) = 1.43$ ,  $p = 0.24$ ,  $\eta_p^2 = 0.30$ . However, there was a significant univariate effect of *clarity*,  $F(1,29) = 4.37$ ,  $p = 0.04$ ,  $\eta_p^2 = 0.13$ . The truthful statements were significantly more *clear* than the deceptive statements. No other univariate analyses showed any significant effects (see Table 1 for descriptive and inferential statistics). Hypothesis 1a partly received support while Hypothesis 1b and 1c received no support.

Table 1. The credibility criteria as a function of Authenticity.

Criterion	Truth		Lie		$F(1,29)$
	$M$	SD	$M$	SD	
Length	0.97	0.69	1.00	0.54	0.07
Coherence	1.07	0.58	1.12	0.58	0.15
Clarity	1.02	0.56	0.80	0.43	4.37*
Details	1.08	0.63	1.02	0.44	0.36
Difficult-to-explain points <sup>a</sup>	1.68	0.33	1.50	0.57	1.98
Credible impression	4.13	1.06	3.80	0.93	1.54
Self-experience	4.15	1.12	3.87	0.89	1.12

Notes. Mean scores ranged from 0 to 2, except for credible impression and self-experience where the mean scores ranged from 1 to 6. High scores indicate truth.

<sup>a</sup>Lie criterion which has been reversed. High scores indicate truth.

\* $p < 0.05$ .

### Effects of interview experience

Participants were divided into two groups through a median split with regard to the number of police interviews they had experienced (less/more experience). A MANOVA was carried out with Authenticity and Interview experience as independent variables and the credibility criteria as dependent variables. There were no multivariate main effects of Interview experience,  $F(7,22) = 1.38$ ,  $p = 0.26$ ,  $\eta_p^2 = 0.31$ , or Authenticity,  $F(7,22) = 1.38$ ,  $p = 0.26$ ,  $\eta_p^2 = 0.30$ . There was no multivariate interaction effect,  $F(7,22) = 0.96$ ,  $p = 0.48$ ,  $\eta_p^2 = 0.24$ . However, a

Table 2. The credibility criteria as a function of Authenticity and Interview experience.

Criterion	Experience	Truth		Lie		$F(1,28)^b$
		$M$	SD	$M$	SD	
Length	Less	0.97	0.79	0.90	0.51	0.62
	More	0.97	0.61	1.10	0.57	
Coherence	Less	1.20	0.68	1.13	0.55	0.83
	More	0.93	0.46	1.10	0.63	
Clarity	Less	1.13	0.69	0.80	0.41	1.28
	More	0.90	0.39	0.80	0.46	
Details	Less	1.20	0.68	0.93	0.37	3.48
	More	0.97	0.58	1.10	0.51	
Difficult-to-explain points <sup>a</sup>	Less	1.80	0.32	1.57	0.56	0.14
	More	1.57	0.32	1.43	0.59	
Credible impression	Less	4.50	0.98	3.77	0.90	2.32
	More	3.77	1.03	3.83	0.99	
Self-experience	Less	4.37	1.13	3.53	0.88	4.85*
	More	3.93	1.12	4.20	0.80	

Notes. Mean scores ranged from 0 to 2, except for credible impression and self-experience where the mean scores ranged from 1 to 6. High scores indicate truth.

<sup>a</sup>Lie criterion which has been reversed. High scores indicate truth.

<sup>b</sup> $F$ -values for the interaction Interview experience  $\times$  Authenticity.

\* $p < 0.05$ .

significant univariate interaction effect was found for *impression of self-experience*,  $F(1,28) = 4.85$ ,  $p < 0.04$ ,  $\eta_p^2 = 0.15$ . For the more experienced offenders the statements were rated higher on *impression of self-experience* in the deceptive statements, while the less experienced offenders were rated higher on this criterion in their truthful statements (see Table 2 for descriptive statistics). Hence, the more experienced participants seemed to have developed a kind of expertise in telling a convincing lie. This can be viewed as support for Hypothesis 2.

In addition, there was a non-significant tendency,  $F(1,28) = 3.48$ ,  $p = 0.07$ ,  $\eta_p^2 = 0.11$ , indicating that deceptive accounts by the more experienced offenders contained more *details* than their truthful accounts, while the truthful statements by the less experienced offenders contained more details than their false statements.

### *Effects of gender*

A MANOVA with Gender and Authenticity as independent variables and the SCoS's criteria as dependent variables was carried out. It showed a multivariate interaction effect of Authenticity  $\times$  Gender,  $F(7,22) = 2.65$ ,  $p = 0.04$ ,  $\eta_p^2 = 0.46$ . The women's lies received in general lower scores than their truths. On the univariate level *credible impression* was significant,  $F(1,28) = 5.42$ ,  $p = 0.03$ ,  $\eta_p^2 = 0.16$ . The women's false statements ( $M = 3.06$ ,  $SD = 0.85$ ) were judged to be less *credible* than their truthful statements ( $M = 4.28$ ,  $SD = 1.33$ ).

There was no multivariate main effect of Authenticity (although very close),  $F(7,22) = 2.37$ ,  $p = 0.06$ ,  $\eta_p^2 = 0.43$ . Univariate analyses did however reveal a significant effect of Authenticity on *credible impression*,  $F(1,28) = 4.64$ ,  $p = 0.04$ ,  $\eta_p^2 = 0.14$ . The offenders' truthful statements ( $M = 4.13$ ,  $SD = 1.06$ ) were rated higher on *credible impression* than their deceptive statements ( $M = 3.80$ ,  $SD = 0.93$ ).

There was no multivariate main effect of Gender,  $F(7,22) = 1.13$ ,  $p = 0.38$ ,  $\eta_p^2 = 0.26$ . However, a significant univariate effect of *coherence* was obtained,  $F(1,28) = 5.61$ ,  $p = 0.02$ ,  $\eta_p^2 = 0.17$ ; women's statements ( $M = 0.81$ ,  $SD = 0.62$ ) were less coherent than the men's statements ( $M = 1.21$ ,  $SD = 0.33$ ). In general, women's statements tended to receive lower scores on the credibility criteria than men's statements.

### **Discussion**

The purpose of this study was to test the criteria for veracity assessments used by the Swedish courts. These criteria have never before been tested scientifically. It was hypothesized that the criteria would distinguish truthful statements from deceptive, but none of these hypotheses (Hypotheses 1a, 1b and 1c) received full support. The only discriminative criterion was *clarity*: truthful accounts tended to be more *clear* than the false. This is in line with the research on RM (Masip et al., 2005). It is important though to notice that the *clarity* criterion tested in this study is not exactly the same as *clarity* in RM. The latter criterion is more precisely defined (Schelin, 2006).

In line with Hypothesis 2 police interview experience did have an impact on the scores of the criteria. Gender was not expected to influence the scores but did in fact turn out to be of some importance. Hypotheses 1a, 1b and 1c partly received support

in that women offenders' truthful and false statements could be distinguished if all criteria were used as one tool; that is, the multivariate effect was significant. However, that is not how the SCoS's criteria are used; it is decided from case to case if and which criterion shall be applied (Schelin, 2006). The single best discriminative criterion in women's statements in this study was *credible impression*. The results in this study did not point to remarkable differences with regard to interview experience, but the only significant difference was found in the criterion *impression of self-experience*; a criterion very close to *credible impression*. According to Schelin (2006) these two criteria are the most influential on the credibility assessment as performed by the courts.

Women's statements in general tended to give a less strong impression of credibility, although a significant effect was found only for one criterion: *coherence*. The reason for this may be that women offenders tend to have a background history of for instance more violence and more emotionally charged experiences (e.g. Yourstone, Lindholm, & Kristiansson, 2008), which may be retold in a less coherent fashion than less emotional experiences. Only further research can shed light on the gender differences issue in statement analysis.

The offenders who were most experienced regarding police interviews told deceptive statements which gave a stronger *impression of self-experience* than their truthful statements. Hence, they seemed to have developed a kind of expertise in telling convincing lies. This is giving support to the feedback hypothesis (Hartwig et al., 2004; Vrij & Semin, 1996).

A defendant often denies having committed the crime and hence the statement tends to be short independently on whether s/he is guilty or not. It is therefore much more difficult to carry out a statement analysis on the defendant's account than on a plaintiff's account (Gregow, 1996). This is sometimes made (Schelin, 2006), but there does not seem to be any specific guidelines for how to analyse statements made by defendants/suspects.

The issue of whether statement analysis in its current form can or should be carried out also on defendants' or suspects' statements is ripe for more research. Very few studies on deception (the polygraph studies excluded) have involved participants with criminal experience. The authors of this paper know of only two studies where well-known techniques for statement analysis have been tested on offenders: one study on CBCA (Lee et al. 2008) and one on RM (Bond & Lee, 2005). Hence, research on the usefulness of statement analysis on suspect's accounts is so far a neglected topic. The current study contributed by using actual offenders' statements and by analysing these with actually used credibility criteria.

*Quantity of details* is in CBCA a criterion with extensive support: truthful statements tend to be more rich in detail than false (Vrij, 2005). There were no significant results in this study on the *details* criterion, albeit a tendency towards the opposite pattern for the offenders with most experience of police interviews: those offenders' lies tended to be more detailed than their truths. Bond and Lee (2005) tested RM on offenders' accounts and found more *spatial details* in their lies than in their truths, which was an unexpected result but finds some support in this study. Strömwall and Granhag (2003) investigated what police detectives, prosecutors and judges believed to be discriminative cues when assessing veracity. Interestingly, the beliefs of police detectives significantly differed from the other groups: lies were thought to be more detailed than truths. The results of this study do not fully

support the police detectives' beliefs; however, the question of the applicability of the detail criterion for experienced offenders' truths and lies does arise.

### **Limitations**

The inter-rater reliabilities were generally low. Reasons may be that the definitions of the SCoS's criteria are not precise enough (or in some cases completely lacking) or perhaps that it takes certain judicial training or experience to understand and use them correctly. Judges in Swedish courts are however not trained in any systematic way about the method; yet they are expected to understand and use the criteria. In addition, the criteria are used by judicial decision-makers in many countries, for instance in asylum procedures (Kagan, 2003). In line with our finding Kagan argued that credibility assessments based on these criteria seem to be frequently influenced by subjectivity. It is important to further investigate the inter-rater reliability of the criteria since they are widely practised by lawyers and other judicial decision-makers around the world. As Kagan (2003) and Vrij (2008) among others has pointed out, it is extremely important to carry out credibility assessments that are not arbitrary – objectivity is needed in these decisions which often have far-reaching consequences for the individual. Two judges need to reach the same conclusion from the same statements – one wonders how they are able to do so when the criteria used are not clearly defined. Especially criteria such as *impression of self-experience* and *credible impression* seem to be very hard to pin down.

Another limitation is that ground truth was not established in this study. It is not fully clear whether the participants lied as they were instructed (although the instructions were very thorough). The truths told in this study are not problematic in this respect, since they were crimes the offenders had committed and being sentenced for. As far as the lies go, a ground 'truth' establishing was not possible.

Furthermore, the study used a small sample as gaining access to this type of participants and have them volunteer is undeniably more difficult than using the usual sample of university students. The low sample size resulted in it not being meaningful to separate the lies and truths told in different categories to find out if differences in credibility criteria would emerge for statements about violence crimes as compared to property crimes, for example.

In addition, experimental conditions differ from real life conditions in several aspects. For instance, the participants in the study knew that they would suffer no consequences from their truthful and false confessions and neither did they have anything to gain from convincing the interviewer about their guilt. The influence of high stakes on offenders' uncoerced confessions, and the analysis of these should be addressed in future research.

### **Conclusions and future directions**

The Swedish courts are at present applying a complex psychological technique for statement analysis (Diesen, 2008). The criteria used by the Swedish courts lack scientific support. Despite this they are applied and thus affecting many people. Similar criteria are also used in international criminal courts and in asylum procedures (May & Wierda, 2002; Noll, 2005). Hence, it is not a problem restricted to Swedish courts.

The most experienced offenders in this study seem to have developed an expertise in telling convincing lies. If offenders do receive feedback through their criminal lifestyle as suggested by for example Vrij and Semin (1996), then they possess knowledge which may be useful for supplementary training and continuing education for law enforcement personnel. It is of course a drawback for all of society if offenders are the only group who receive essential feedback concerning deception and its detection.

It must be possible to convict a defendant on the basis of statements only since perpetrators in some types of crime otherwise rarely would be possible to convict (Gregow, 1996). However, the techniques used for statement analysis in court must be anchored in science, which is not the case at present in Sweden as well as in many other countries. It is imperative to carry out more research on the criteria and decision rules used by judicial decision makers around the world.

More research is also needed on how statement analysis can be carried out on accounts given by suspects and defendants, especially research with criminal experienced participants. This study is one of the first contribution to the exploration of how statement analysis can be applied on offenders' accounts.

### Notes

1. Women and men did not differ in number of interviews,  $t(28) = 1.31$ ,  $p = 0.20$ , or in number of months in prison,  $t(28) = 1.82$ ,  $p = 0.08$ .
2. In a different study carried out simultaneously, some of the offenders reported using the strategy of staying 'close to the truth' to convince the interviewers when deceiving (Strömwall & Willén, 2010). At least some of the false statements could therefore be expected to contain truthful parts, despite the instruction to create an all false story.

### Acknowledgements

Thanks to the offenders and prisons who participated. Thanks to Naria Bing for help with data collection. Parts of this research were presented at the 19th European Conference on Psychology and Law in Sorrento, Italy in September 2009. This research was financially supported by a grant from the Swedish Research Council to the second author.

### References

- Alonso-Quecuty, M.L. (1992). Deception detection and reality monitoring: A new answer to an old question? In F. Lösel, D. Bender, & T. Bliesener (Eds.), *Psychology and law: International perspectives* (pp. 328–332). Berlin: Walter de Gruyter.
- Bond, C.F., & DePaulo, B.M. (2006). Accuracy of deception judgements. *Personality and Social Psychology Review*, *10*, 214–234.
- Bond, G.D., & Lee, A.Y. (2005). Language of lies in prison: Linguistic classification of prisoners' truthful and deceptive natural language. *Applied Cognitive Psychology*, *19*, 313–329.
- Bond, G.D., Malloy, D.M., Arias, E.A., Nunn, S.N., & Thompson, L.A. (2005). Lie-biased decision making in prison. *Communication Reports*, *18*, 9–19.
- Bond, G.D., Thompson, L.A., & Malloy, D.M. (2005). Vulnerability of older adults to deception in prison and nonprison contexts. *Psychology and Aging*, *20*, 60–70.
- DePaulo, B.M., Lindsay, J.J., Malone, B.E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*, 74–118.

- Diesen, C. (2008). Juridiska perspektiv på bevisvärdering [Legal perspectives on evidentiary assessment]. In P.A. Granhag & S.Å. Christianson (Eds.), *Handbok i rättspsykologi* (pp. 395–405). Stockholm: Liber.
- Granhag, P.A., Strömwall, L.A., & Jonsson, A.-C. (2003). Partners in crime: How liars in collusion betray themselves. *Journal of Applied Social Psychology, 33*, 848–868.
- Gregow, T. (1996). Några synpunkter på frågan om bevisprövning och bevisvärdering i mål om sexuella övergrepp mot barn [Some standpoints on the issues of evidence testing and evidentiary assessment in cases concerning sexual abuse]. *Svensk Juristtidning [Swedish Law Review]*, 7, 509–523.
- Gudjonsson, G.H. (2003). *The psychology of interrogations and confessions – A handbook*. Chichester: John Wiley & Sons.
- Hartwig, M., Granhag, P.A., Strömwall, L.A., & Andersson, L.O. (2004). Suspicious minds: Criminals' ability to detect deception. *Psychology, Crime & Law, 10*, 83–95.
- Home Office, Asylum policy instruction. *Assessing credibility in asylum and human rights claims*. Downloaded June 23, 2010 from [www.bia.homeoffice.gov.uk/sitecontent/documents/policyandlaw/asylumpolicyinstructions/apis/credibility.pdf](http://www.bia.homeoffice.gov.uk/sitecontent/documents/policyandlaw/asylumpolicyinstructions/apis/credibility.pdf)
- Johnson, M.K., & Raye, C.L. (1981). Reality Monitoring. *Psychological Review, 88*, 67–85.
- Kagan, M. (2003). Is truth in the eye of the beholder? Objective credibility assessment in refugee status determination. *Georgetown Immigration Law Journal, 17*, 367–381.
- Köhnken, G. (2004). Statement Validity Analysis and the 'detection of the truth'. In P.A. Granhag & L.A. Strömwall (Eds.), *Detection of deception in forensic contexts* (pp. 41–63). Cambridge: Cambridge University Press.
- Lee, Z., Klaver, J.R., & Hart, S.D. (2008). Psychopathy and verbal indicators of deception in offenders. *Psychology, Crime & Law, 14*, 73–84.
- Masip, J., Sporer, S.L., Garrido, E., & Herrero, C. (2005). The detection of deception with the Reality Monitoring approach: A review of the empirical evidence. *Psychology, Crime & Law, 11*, 99–122.
- May, R., & Wierda, M. (2002). *International criminal evidence*. Ardsley: Transnational Publishers.
- Millibank, J. (2009). The ring of truth: A case study of credibility assessment in particular social group refugee determinations. *International Journal of Refugee Law, 21*, 1–33.
- NJA. (1992). *New Legal Archive*, 446–470.
- Noll, G. (2005). Introduction: Re-mapping evidentiary assessment in asylum procedures. In G. Noll (Ed.), *Proof, evidentiary assessment and credibility in asylum procedures* (pp. 1–10). Leiden: Martinus Nijhoff Publishers.
- Ofshe, R.J., & Leo, R.A. (1997). The social psychology of police interrogation: The theory and classification of true and false confessions. *Studies in Law, Politics and Society, 16*, 189–251.
- Sigurdsson, J.F., & Gudjonsson, G.H. (1996). The psychological characteristics of 'false confessors'. A study among Icelandic prison inmates and juvenile offenders. *Personality and Individual Differences, 20*, 321–329.
- Schelin, L. (2006). *Bevisvärdering av utsagor i brottmål [Evidentiary assessment of statements in criminal cases]*. Doctoral thesis, University of Stockholm: Faculty of Law, Stockholm.
- Sporer, S.L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta analytic synthesis. *Applied Cognitive Psychology, 20*, 421–446.
- Strömwall, L.A. (2010). Assessing reliability by analyzing the verbal content: The case of Sweden. In P.A. Granhag (Ed.), *Forensic psychology in context: Nordic and International Approaches* (pp. 264–280). Cullompton, UK: Willan Publishing.
- Strömwall, L.A., & Granhag, P.A. (2003). How to detect deception? Arresting the beliefs of police officers, prosecutors and judges. *Psychology, Crime & Law, 9*, 19–36.
- Strömwall, L.A., & Willén, R.M. (2010). *Inside criminal minds: Offenders' strategies when lying*. Manuscript submitted for publication.
- Sutorius, H., & Kaldal, A. (2003). *Bevisprövning vid sexualbrott [Evidentiary assessment in sexual crimes]*. Stockholm: Norstedts Juridik.
- Sweeney, J.A. (2009). Credibility, proof and refugee law. *International Journal of Refugee Law, 21*, 700–726.

- Trankell, A. (1965). *Vittnespsykologin som hjälpmedel i svensk rättegång [The witness psychology as aid in Swedish trials]*. Stockholm: Liber.
- Trankell, A. (1967). *Vittnespsykologins arbetsmetoder [The methods of witness psychology]* (2nd edn). Stockholm: Liber.
- UNHCR, (1992). *Handbook on procedures and criteria for determining refugee status under the 1951 convention and the 1967 protocol relating to the status of refugees*. Geneva: UNHCR.
- Vrij, A. (2005). Criteria-based content analysis. A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law*, 11, 3–41.
- Vrij, A. (2008). *Detecting lies and deceit. Pitfalls and opportunities* (2nd ed). Chichester: Wiley.
- Vrij, A., & Semin, G.R. (1996). Lie experts beliefs about nonverbal indicators of deception. *Journal of Nonverbal Behavior*, 20, 65–80.
- Yourstone, J., Lindholm, T., & Kristiansson, M. (2008). Women who kill: A comparison of the psychosocial background of female and male perpetrators. *International Journal of Law and Psychiatry*, 31, 374–383.