

ORIGINAL ARTICLE

False confessions in the lab: Do plausibility and consequences matter?

ROBERT HORSELENBERG^{1,2}, HARALD MERCKELBACH², TOM SMEETS², DIRK FRANSSSENS², GJALT-JORN YGRAM PETERS², & GWENNY ZELES²

¹*Department of Psychology, Open University Heerlen, The Netherlands and* ²*Department of Experimental Psychology, Maastricht University, The Netherlands*

Abstract

The present paper describes three studies that examined false confessions in the laboratory. Studies 1 ($N=56$) and 2 ($N=9$) relied on the by now classic computer crash paradigm introduced by Kassin and Kiechel (*Psychological Science*, 7, 125–128, 1996). Study 3 ($N=12$) employed a novel paradigm in which undergraduate participants were falsely accused of exam fraud. Our data indicate that false confessions do occur, even when conditions become more ecologically valid. Furthermore, we explored whether individual differences in compliance, suggestibility, fantasy proneness, dissociation, and cognitive failures are related to false confessions. Of these, only fantasy proneness was associated with false confessions.

Keywords: *False confessions, individual differences, fantasy proneness*

Introduction

Throughout the history of legal practice, triers of fact have attributed much evidentiary weight to confessions (Kassin & McNall, 1991; Kassin & Sukel, 1997; Kassin & Wrightsman, 1985). In general, confessions have greater impact on judicial decisions than eyewitness testimonies (Kassin & Neumann, 1997). Germane to this issue are those individuals who have been wrongly convicted on the basis of their false confessions (for a review, see Gudjonsson, 2003). Estimates of the base rate of false confessions vary. Whereas Cassell (1996) argued that they have a relatively low base rate, Gudjonsson and Sigurdsson (1994; see also Sigurdsson & Gudjonsson, 1996) found that 12% in their sample of Icelandic prison inmates claimed to have falsely confessed. Although this latter finding might be an overestimation, Bedau and Radelet (1987) noted that false confessions constitute a real problem in the sense that their consequences are far-reaching. These authors analysed 350 miscarriages of justice and in at least 10% of these cases, false confessions played a crucial role. Recently, Scheck, Neufeld, and Dwyer (2000) reported

Correspondence: Robert Horselenberg, Department of Psychology, Open University, PO Box 2960, 6401 DL, Heerlen, The Netherlands. E-mail: Robert.Horselenberg@OU.NL

a false confession incidence of 22% in their Innocence Project, which supports the conclusions of Bedau and Radelet (1987). The large-scale surveys of Gudjonsson and colleagues (Gudjonsson, Sigurdsson, Bragasson, Einarsson, & Valdimarsdottir, 2004; Gudjonsson, Sigurdsson, & Einarsson, 2004) among college and university students found self-reported base rates of 2.8% (false confessions to police officers) and 8.8% (false confessions to parents and teachers).

In a recent extensive review, Gudjonsson (2003) proposed a modified taxonomy for false confessions. This taxonomy intends to take into account critiques on the classifications introduced by Kassin and Wrightsman (i.e. voluntary, coerced-compliant, and coerced-internalized false confessions; Kassin & Wrightsman, 1985) and by Ofshe and Leo (i.e. voluntary, coerced-compliant, stress-compliant, non-coerced-persuaded, and coerced-persuaded false confessions; Ofshe & Leo, 1997). In short, Gudjonsson suggests that for a better understanding of the various false confession types, the term “coerced” should be redefined as “pressured” and that the source of pressure should be categorized as either internal or external. Accordingly, the following classification of false confessions can be made. To begin with, there is the voluntary false confession, where pressure stems from an internal source. This source refers to the individual having internal motives (e.g. to protect a significant other or a morbid desire for notoriety) to confess falsely. A second type of false confession is the pressured-compliant, where the source of pressure may be either caused by custodial (e.g. during a police interrogation) or non-custodial circumstances. Non-custodial pressure originates from people other than police officers (e.g. peers, cell mates, customs and immigration officers, undercover police officers). It is important to note that with this type of confession, the person does not believe that he or she committed the crime confessed. A third type of false confession is the pressured-internalized type. Here, suspects actually come to believe that they are guilty of a crime they did not commit. Again, the source of pressure can be either custodial or non-custodial.

Kassin and Kiechel (1996) examined whether analogues of the various false confession types may occur under laboratory conditions. These authors instructed their undergraduate participants to copy letters that were presented on a computer screen. Participants were told not to touch the “Alt” key because otherwise the computer would crash. During the task, the computer did crash and participants were falsely accused of having touched the forbidden key. Next, participants were asked to sign a written confession and they were approached by a confederate of the experimenter who presented as a naive subject. This confederate asked the participant about what happened. Kassin and Kiechel found that many participants (depending on the precise conditions, between 35 and 100%) were willing to sign a written confession. A smaller portion of them (depending on the conditions, between 0 and 65%) apparently had internalized their false confessions because they related to the confederate how they had touched the forbidden “Alt” key.

Several studies (Candel, Merckelbach, Luyen, & Reyskens, in press; Forrest, Wadkins, & Miller, 2002; Forrest & Wilson, 2003; Horselenberg, Merckelbach, & Josephs, 2003; Klaver, Gordon Rose, & Lee, 2003; Lee, Klaver, & Gordon Rose, 2003; Redlich & Goodman, 2003) have replicated the basic findings of Kassin and Kiechel (1996). For example, Candel and colleagues (in press) used the Kassin and Kiechel paradigm to examine false confessions in young children and found that 36% of the children falsely confessed having touched a forbidden computer key. As well, 89% of the falsely confessing children had internalized their confession (see also Redlich & Goodman, 2003). Forrest et al.’s (2002) study focused on whether stress may contribute to laboratory induced false confessions. These authors reported that stress did not affect false confession and

internalization rates in women, but it did so in men. When stress was absent, women tended to confess and internalize more than did men. The results of Lee et al. (2003) parallel those of Forrest et al. (2002). That is, Lee and colleagues (2003) reported a heightened rate of false confessions for women (80%) compared with men (20%). Yet, this gender difference was only evident for signing confessions, not for internalization rates. Forrest and Wilson (2003) examined the effects of rumination about past events (e.g. reminding someone of his or her criminal history) on false confessions. Although they came up with high overall rates of false confessions and internalizations (90% and 46.3%, respectively), no effect of rumination could be demonstrated.

Horselenberg et al. (2003) studied whether negative consequences (i.e. monetary penalties) would reduce false confession rates. These authors found that, despite monetary penalties, the large majority of their participants (82%) were willing to sign a false confession (see also Redlich & Goodman, 2003). They also noted that in about half of their participants, false confessions were accompanied by internalization and confabulation. In a large sample ($N=204$) of undergraduates, Klaver et al. (2003) studied the effects of interrogation strategies (i.e. maximization or minimization) on false confession rates. Half of their participants were interrogated emphasizing the accidental nature of the act and blaming the sensitivity of the keyboard (i.e. minimization). During interrogations of the other half, the significant loss of data and the supervisor's irritation was emphasized (i.e. maximization). Within both conditions, some participants were accused of having pressed the more plausible "Alt" key, while others were accused of having touched the rather implausible "Esc" key. More confessions were signed in the minimization condition than in the maximization condition (56% and 30%, respectively). Furthermore, when plausibility was low, confession rates dropped dramatically. More specifically, when participants were accused of having pressed the "Alt" key, 59% signed a confession. However, when they were accused of having hit the "Esc" key, 13% signed a confession.

In sum, then, experimental studies suggest that there are many factors that may contribute to a false confession. Case studies have shown that these factors can best be grouped into two classes, namely situational and personal factors. Situational factors include, for example, the interrogation process itself (Leo & Ofshe, 1998), the techniques used by the police (Gudjonsson, 2003), over-zealousness of police officers (Bedau & Putnam, 1996), or presentation of false evidence (e.g. statements supposedly made by an accomplice or a staged eyewitness identification; Kassin, 1997). Personal factors that may increase susceptibility to false confessions are naiveté, stress, anxiety, fatigue, low intelligence (e.g. mental disorders), young age, interpersonal trust, compliance, suggestibility, and alcohol or drug use (Gudjonsson, 2003; Kassin, 1997).

In case reports about false confessions, compliance and suggestibility have received much attention (e.g. Gudjonsson, 2003). However, experimental studies using Kassin and Kiechel's paradigm yielded little or no evidence for the idea that compliance or suggestibility modulates participants' susceptibility to confess falsely (e.g. Candel et al., in press; Horselenberg et al., 2003; Lee et al., 2003). Perhaps, these null findings have to do with the fact that extremely compliant or suggestible individuals are seldom found among undergraduates. A related point is that sample sizes in this research domain are often limited. Interestingly, Redlich and Goodman (2003) found some tentative evidence that especially those participants who yielded to misleading questions were prone to falsely confess. In more general terms, it is fair to conclude that most laboratory work on false confessions has looked at situational factors (interrogation techniques, effects of false evidence, etc.), while relatively few studies systematically looked at relevant personality traits.

Besides compliance and suggestibility, there are other individual difference measures that might be relevant to false confessions. Candidates that have some *prima facie* plausibility are fantasy proneness, dissociation, and self-reported cognitive failures. The point here is that at least some types of false confessions (e.g. the pressured-internalized type; Gudjonsson, 2003; Ost, Costall, & Bull, 2001) can best be considered as pseudomemories. There is an extensive literature showing that fantasy proneness, dissociation, and/or cognitive failures predict an individual's tendency to develop pseudomemories (e.g. Candel, Merckelbach, & Kuijpers, 2003; Hyman & Billings, 1998; Merckelbach, Muris, Rassin, & Horselenberg, 2000; Ost, Vrij, Costall, & Bull, 2002; Winograd, Peluso, & Glover, 1998; but see also Horselenberg et al., 2000; Spanos, Burgess, Burgess, Samuels, & Blois, 1999). It should be noted that these traits overlap to some extent (Kihlstrom, Glisky, & Angiulo, 1994; Merckelbach, Horselenberg, & Muris, 2001). There are good reasons to believe that, through suggestion or imagery, these traits act so as to endow internal events with features that are usually typical for "real" memories (i.e. memories of external events; Merckelbach et al., 2001). Thus, it may well be the case that, because they misinterpret their detailed imagery or fantasies about their involvement in a crime as real memories, people high on fantasy proneness, dissociation, and/or cognitive failures are susceptible to false confessions. Alternatively, individuals high on these traits may interpret their lack of robust memories about the crime as an indication that they must have repressed these memories. This line of reasoning comes close to what Gudjonsson and MacKeith (1982) have termed memory distrust.

The present paper describes three studies which attempted to elicit false confessions in a laboratory situation. Using Kassin and Kiechel's paradigm, study 1 manipulated the plausibility of the accusation (pressing the "Windows" key versus the "F12" key). We tried to replicate Klaver et al.'s (2003) finding that low plausibility goes hand in hand with a relatively low false confession rate. Following the same paradigm, study 2 focused on the severity of consequences. What happens when participants are accused of having caused a computer crash and they know that confessing means paying for the damage? Horselenberg et al.'s (2003) results seem to indicate that even under these circumstances, many participants falsely confess (see also Redlich & Goodman, 2003). In order to create an ecologically valid situation, participants in study 3 were accused of cheating on an exam. We expected that, since cheating is less frequent than wrongfully touching a computer key and as it involved severe consequences, this accusation would drastically reduce the number of false confessors in comparison to other studies. The studies to be described below are not intended to unravel the precise causal antecedents of the various false confession types. Rather they seek to give a solid empirical proof of false confessions under laboratory conditions. In all three studies, we explored whether compliance, suggestibility, fantasy proneness, dissociation, and cognitive failures predispose individuals to falsely confess.

We are, of course, aware of the ethical aspects of our studies. Laboratory studies on false confession require careful planning and so we sought in an early stage the advice of our ethical committee. After extensive discussions, all three studies were approved by the standing ethical committee of the Faculty of Psychology at Maastricht University. Before the studies began, participants signed an informed consent in which it was explicitly stated that they had the right to withdraw from the study whenever they wanted to. Furthermore, after the studies, all participants were debriefed thoroughly and this was done in two ways: they were told about the ideas and procedures behind each study, but they also received a letter in which detailed background information was given. Most importantly, we decided beforehand that we would discontinue a study whenever participants became too distressed. In that case we also had a scenario that specified how participants had to be

reassured and which other measures had to be taken. For study 3, we decided that we would give participants a follow-up debriefing after approximately 5 months. In addition, by pre-testing participants we were able to exclude highly anxious individuals from participation in study 3. Nevertheless, after we had run a limited number of participants in studies 2 and 3, we decided to call these studies to a halt. The manipulations were experienced by some participants and experimenters as too upsetting.

STUDY 1

Method

Participants

Participants were 72 undergraduate psychology students. After the study, all participants were given an exit interview in which they were asked whether they had known all along that the computer task was fake. It was explicitly stated that acknowledging this would not result in any consequences. Of all participants, 16 were discarded from further analysis as they saw through the experimental set-up. Thus, the remaining sample consisted of 56 participants (46 women). Their mean age was 20.6 years (range 18–33 years).

Materials

Participants completed the Gudjonsson Compliance Scale (GCS; Gudjonsson, 1989), the Gudjonsson Suggestibility Scale (GSS1; Gudjonsson, 1984), the Creative Experiences Questionnaire (CEQ; Merckelbach et al., 2001), the Dissociative Experiences Scale-C (DES-C; Wright & Loftus, 1999), and the Cognitive Failures Questionnaire (CFQ; Broadbent, Cooper, Fitzgerald, & Parkes, 1982).

The GCS ($\alpha = 0.45$) is a self-report questionnaire consisting of 20 true/false statements that have to do with how individuals deal with conflicts and confrontation (e.g. “I give in easily to people when I am pressured.”). Responses are summed up to obtain a total score, with higher scores reflecting a stronger tendency to comply.

The GSS (version 1) comprises a short story that is read out loud to participants. This is followed by 20 specific questions about the story. Fifteen of the questions convey misleading information about the story. The extent to which an individual gives in to leading questions constitutes the Yield score (range 0–15). The other five questions measure memory performance. Following the 20 questions, participants are told in an authoritative manner that they made many mistakes and that all questions have to be answered for a second time. The number of times a participant changes his/her answers constitutes the Shift score (range 0–20). Yield and Shift scores are summed to obtain a total suggestibility score (range 0–35), with a higher score indicating higher interrogative suggestibility. A number of other parameters can be derived from the GSS, but for practical reasons the present study only focused on total suggestibility scores.

The CEQ ($\alpha = 0.75$) is a 25-item yes/no index of fantasy proneness. The CEQ items were derived from extensive case descriptions of fantasy proneness provided by Wilson and Barber (1982). Typical CEQ items are: “In general, I spend at least half of the day fantasizing or daydreaming” and “My fantasies are so vivid that they are like a good movie”. “Yes” answers are summed to obtain a total CEQ score, with higher scores indicating higher fantasy proneness levels.

The DES-C ($\alpha=0.95$) consists of 28 items that pertain to dissociative phenomena (e.g. feelings of derealization, depersonalization, disturbances in memory). Respondents indicate on 11-point scales (anchors: 0 = *much less than others*; 11 = *much more than others*) how often they have these experiences compared with others. Scores are averaged to obtain a total DES score. Higher total DES scores indicate stronger dissociative tendencies.

The CFQ ($\alpha=0.77$) is a 25-item self-report instrument that measures the frequency of everyday lapses and minor blunders (e.g. “Do you forget appointments?”). Participants indicate on five-point scales how often they experienced each cognitive failure in the past month (anchors: 0 = *never*; 4 = *very often*). Scores are summed to obtain a total CFQ score, with higher scores indicating a higher frequency of self-reported cognitive failures.

For the computer task, a Dell PC was used running a program in MS-DOS. Stimuli appeared on a 17-inch screen in lower case in yellow colour against a grey background. Responses were recorded online to correct for participants actually hitting the “Windows” or “F12” key. None of the participants actually hit these keys.

Procedure

Participants were randomly assigned to one of two groups. All participants were told that they were contributing to an ergonomic study evaluating the effectiveness of new keyboard configurations. They were also told that they were randomly assigned to one of three keyboard conditions. To increase the credibility of the cover story, participants completed several questionnaires about their typing abilities, typing performance, and problems they previously encountered with various keyboard arrangements. They also completed the CEQ, DES-C, and CFQ. All participants were then informed that they were assigned to a control group and, therefore, had to carry out a task on a standard keyboard. Subsequently, they were instructed to type as quickly and accurately as possible various letters appearing on the screen. Depending on their condition, participants were instructed that they should refrain from touching either the “Windows” key or the “F12” key and that touching this key would result in a fatal computer crash and complete loss of data.

After the computer had crashed, the experimenter falsely accused participants of having touched the forbidden key and insisted that she saw it happening with her own eyes (i.e. false incriminating evidence). Next, the experimenter invited participants to sign a handwritten confession stating – and this was stressed explicitly by the experimenter – that the data were lost because the participant had hit the forbidden key. If participants refused to sign the confession form, the experimenter once again encouraged them to sign it. The experimenter then left the room and advised the participants to think for a while about what had happened. After about 5 minutes, a confederate entered the room and explicitly stated that he was not associated with the ongoing experiment. The confederate asked participants what had happened. Participants’ answers were later written down by the confederate, so as to derive a measure of internalization. Finally, the experimenter re-entered and fully debriefed participants and asked them not to discuss the experiment with their colleagues. Approximately 2 years later, participants were once again invited to the laboratory to complete the GCS and GSS. The reason for this considerable time interval was twofold. First of all, with this period, we tried to avoid carry-over effects from the false confession experiment to the GCS and GSS. In contrast to the CEQ, DES-C, and CFQ, it is easy for participants to see that the GCS and GSS are highly relevant to the false confession experiment. Secondly, gathering GCS and GSS data 2 years later mimics real-world conditions. After all, during a judicial procedure, it takes months before an expert

witness comes into play (Gudjonsson, 2003). The disadvantage of this set-up was that only 34 participants could be reached and were willing to complete the GCS and the GSS.

Results and discussion

Results showed that confession rates were highest (77%) in the high plausibility condition (i.e. the “Windows” key) and were lowest (58%) in the low plausibility condition (i.e. “F12” key). This difference, however, was not significant: Fisher’s exact $p=0.16$, two-tailed. Interestingly, internalization only occurred in the “Windows” key condition with 26% of the confessors showing internalization during their conversation with the confederate: Fisher’s exact $p=0.01$, two-tailed.

Table I shows mean scores on the GCS, GSS-shifts, GSS-total, CEQ, DES-C, and CFQ. These scores are well in line with those reported by previous studies using undergraduate samples (e.g. Gudjonsson, 1997; Merckelbach et al., 2000). To explore whether individual difference variables were related to absence or presence of false confessions (i.e. signing the confession form), group means within each condition (i.e. “F12” versus “Windows” key) were compared, using independent samples t -tests. No significant differences between confessors and non-confessors were found (all t s < 1.2 ; all p s > 0.10). Yet, when the data were collapsed across conditions, a significant difference between confessors and non-confessors with regard to their CEQ scores was found [$t(54) = 2.32, p = 0.02$]. That is, false confessors scored higher on fantasy proneness than non-confessors (means being 7.8 and 5.8, respectively). No other significant group differences in individual difference measures emerged. Given the fact that there were only six participants who internalized, we did not compare internalizers versus others with respect to mean individual difference scores.

In line with a number of previously mentioned studies, our results show that it is relatively easy to make participants confess, irrespective of the plausibility of the key they allegedly hit. Apparently, the sheer presence of a pseudo-eyewitness who claims to have seen that the participant touched the forbidden key is enough for many to confess. However, plausibility does affect internalization rates. Only when the act was plausible, participants tended to internalize the false confession (i.e. came to believe their false confession). Whereas the internalization rate was lower than that found in our earlier study (Horselenberg et al., 2003; 26% versus 42%, respectively), the figure of 26% accords well with the overall internalization rate (28%) reported by Kassin and Kiechel (1996). Importantly, some

Table I. Mean scores of participants in study 1 in each condition, who did or did not falsely confess on measures of compliance (GCS), shift score (GSS-shift) suggestibility (GSS-total), fantasy proneness (CEQ), dissociative experiences (DES-C), and cognitive failures (CFQ).

	“F12” key ($N=26$)		“Windows” key ($N=30$)	
	Non-confessors ($n=11$)	False confessors ($n=15$)	Non-confessors ($n=7$)	False confessors ($n=23$)
GCS*	7.6 (SD = 7.5, $n=8$)	8.8 (SD = 3.9, $n=9$)	7.5 (SD = 0.7, $n=2$)	9.2 (SD = 1.8, $n=15$)
GSS-shift*	2.6 (SD = 1.4, $n=8$)	3.2 (SD = 4.1, $n=9$)	1 (SD = 1.4, $n=2$)	2.3 (SD = 2, $n=15$)
GSS-total*	7.6 (SD = 7.5, $n=8$)	6 (SD = 5.2, $n=9$)	3.5 (SD = 4.9, $n=2$)	5.8 (SD = 3.1, $n=15$)
CEQ	6 (SD = 2.8)	8.1 (SD = 4.9)	5.4 (SD = 1.6)	7.7 (SD = 3.9)
DES-C	42.4 (SD = 10.9)	40.3 (SD = 11.9)	35.1 (SD = 15.8)	40.6 (SD = 13.9)
CFQ	56 (SD = 8.6)	53.7 (SD = 10.6)	59.7 (SD = 8.6)	54.3 (SD = 7.4)

*Due to missing data, the number of participants differs from total N .

tentative evidence was found suggesting a relationship between fantasy proneness and false confessions. No other significant connections were found.

Although experimental studies have shown that it is relatively easy to elicit false confessions, critical points have been raised with regard to the traditional Kassin and Kiechel paradigm. For example, Forrest et al. (2002) argued that a computer crash does not involve any consequences apart from embarrassment. In contrast, falsely confessing to a crime usually leads to penalties of all sorts. With this in mind, we accused participants in study 2 of having caused the computer to crash and then they were told to pay for the damage.

STUDY 2

Method

Participants

Participants were 13 undergraduate psychology students. Of these, four were excluded after exit interviews had made it clear that they knew from the start the cover story was fake (e.g. because they had been warned). Thus, the remaining sample consisted of nine participants (six women) with a mean age of 19.6 years (range 18–21 years).

Materials

The same materials as in study 1 were used. Cronbach α s for the GCS, CEQ, DES-C, and CFQ were 0.46, 0.59, 0.93, and 0.84, respectively. As in study 1, GCS and GSS data were gathered approximately 2 years after the computer crash manipulation. We were able to locate eight participants.

Procedure

The procedure of study 1 was closely followed. Participants were told not to touch the “plausible” “Windows” key. After the computer had crashed, they were told that they had to sign a written confession stating that they were being held accountable for the financial costs of the crash they had caused. Participants were told that costs would amount to about €250.

Results and discussion

Results can be summarized as follows. One participant was willing to sign a confession (11%). In her conversation with the confederate, this participant did not show signs of internalization. Given the small sample size, formal statistical analysis was impossible. For this reason, only descriptive data about the individual difference measures is given (see Table II). Compared with mean scores of non-confessors, the false confessor had a higher GSS-shift and GSS-total score and, to a lesser extent, higher DES-C and CFQ scores. However, the participant who falsely confessed did not have an elevated fantasy proneness score.

Whereas earlier studies relying on Kassin and Kiechel’s paradigm found false confession rates up to 100%, the false confession rate in the current study (11%) was considerably lower. In any case, the figure of 11% is much lower than the 77% obtained in the high plausibility condition of study 1. Except for the consequences of confessing, study 2 and the

Table II. Mean scores of participants in study 2 on measures of compliance (GCS), shift score (GSS-shift), suggestibility (GSS-total), fantasy proneness (CEQ), dissociative experiences (DES-C), and cognitive failures (CFQ).

	Non-confessors (<i>n</i> = 8)	False confessor (<i>n</i> = 1)
GCS*	11.3 (SD = 2.8, range: 8–16; <i>n</i> = 7)	11
GSS-shift*	1.7 (SD = 2.0, range: 0–2; <i>n</i> = 7)	6
GSS-total*	5.4 (SD = 4.7, range: 1–15; <i>n</i> = 7)	11
CEQ	6.8 (SD = 3.2, range: 3–11)	5
DES-C	33.7 (SD = 13.5, range: 11–52)	39.3
CFQ	60.9 (SD = 10.9, range: 44–76)	64

*Due to missing data, the number of participants differs from total *N*.

high plausibility condition of study 1 were highly comparable in terms of participant characteristics (e.g. age, educational level), procedure and so on. Therefore, we compared false confession rates, using the Fisher exact probability test. This yielded a significant difference ($p = 0.001$; two-tailed), indicating that articulating the consequences of false confessions reduces the rate of this type of confession. In fact, the rate of 11% comes close to what has been found in field studies (Gudjonsson & Sigurdsson, 1994; Sigurdsson & Gudjonsson, 1996). Klaver et al.'s (2003) laboratory study also reports such low rates when a low plausibility event and minimization techniques were combined. Together, these data suggest that how people perceive the consequences of confessing may be a crucial determinant of whether or not they are going to confess (see also Kassin & McNall, 1991).

Of course, one could argue that the low false confession rate found in the current study cast doubts on the assumption that normal, healthy people can be easily brought to falsely confess. Note, however, that the undergraduates in our study were intelligent and critical participants. The simple fact that one of them falsely confessed, while accepting the consequences, demonstrates in our opinion that even in such populations, false confessions may arise. Still one could maintain that the typical Kassin and Kiechel experiment suffers from a lack of ecological validity (Inbau, Reid, Buckley & Jayne, 2001). Being involved in an ergonomic study, in which one is falsely accused of having touched a key, is not what normally might happen to undergraduates. With this in mind, study 3 was conducted. Here, participants were accused of cheating on an exam for which they would be excluded from future exams.

STUDY 3

Method

Participants

Twenty-two university undergraduate students served as participants. Six of them were excluded after exit interviews had made it clear that they had been familiar with false confession research. All in all, 16 participants (seven women) with a mean age of 19.4 years (range: 19–25 years) participated in what they thought was a study on intelligence and certain personality variables. Participants were first screened on the Dutch version of the State Trait Anxiety Inventory-trait (STAI-t; van der Ploeg, Defares & Spielberger, 1980). It was anticipated that participants with a STAI-t score higher than 46 (the third quartile) might experience the manipulation as too stressful. Consequently, four participants

were not invited to take part in the next phase of the study, leaving 12 participants in the final sample. Undergraduates received one course credit for participating in this study.

Materials

The same materials as in Studies 1 and 2 were used, except that now the DES (Bernstein, & Putnam, 1986) instead of the DES-C was used. Cronbach α s for the GCS, CEQ, DES, and CFQ were 0.41, 0.51, 0.91, and 0.91, respectively.

Procedure

Participants were approached by an experimenter, carrying a stack of papers including the test materials and a sketch version of an upcoming exam. The experimenter showed the participants to the experimental room. Upon entering, he put the papers with the exam on top of the desk in clear sight of the participants. He then asked participants to sit down and briefly explained the goal of the study (e.g. "You are required to fill in two questionnaires and then complete the Raven's Intelligence Test"). After the participants had signed a consent form, the experimenter left the room, "forgetting" the stack of papers with the exam still on top of them. Participants began filling out the questionnaires. When they started with the intelligence test, the experimenter returned with the excuse that he had left some papers on the desk.

While leaving again, he noticed that one of the corners of the exam booklet was creased (i.e. false incriminating evidence). He then turned around and asked the participant whether s/he had looked at the exam, accusing the participant of having cheated. Initially, all participants denied the allegation. The experimenter left the room, saying that he was going to discuss further proceedings with the examination board, leaving the participant to think about what s/he had done. Approximately 5 minutes later, the experimenter returned saying that the examination board was in a meeting and right now had no time to discuss the issue with the participant. The experimenter then wrote down a statement (i.e. "I looked at the exam and therefore committed exam fraud") and again accused the participant, arguing that the creased corner was clear evidence and that the examination board had informed him that refusing to confess would result in the participant not being permitted to take any upcoming exams until further notice. Participants were told that signing the confession would prevent the exclusion from the exams.

To explore whether participants would exhibit internalization, a confederate acting as if he were the next participant entered the room saying that he had been waiting for 15 minutes. The experimenter acknowledged this, excused himself and left the room to get the course credit form. The confederate asked the participant why things had taken so long. If the participant said that they had actually looked at the exam, this was scored as internalization. All sessions were videotaped using two hidden cameras to verify that participants did not peek at the exam (but see below). Finally, all participants were extensively debriefed and explicitly instructed not to talk with other students about the experiment as it would be of vital importance that future participants were naive as to the goal of the study.

In order to obtain GCS and GSS data, participants were asked 5 months later if they were willing to participate in another experiment. To ensure that participants were in no way biased, the "second experiment" was introduced as a separate and unrelated study and

with another experimenter obtaining the data. Again, participants were thoroughly debriefed after this second session.

Results and discussion

Of the 12 participants, one participant falsely confessed (8%), while another participant actually did browse the exam, but did not read it. This participant confessed, albeit that she denied the browsing part and reported later that in her experience, she had falsely confessed (e.g. false denier). There were no indications that these participants had internalized their confessions. Given the small number of false confessors, no statistical analysis could be performed on individual difference measures. Therefore, the scores of these two individuals were compared with the data obtained from non-confessing participants by examining their individual scores (see Table III).

Relative to non-confessors, the participant who falsely confessed and the false denier appeared to have elevated GCS scores. However, the false denier also had a low score on the GSS. The false confessor exhibited a different pattern. Compared with non-confessors, he had raised scores on the GSS and CEQ.

For the judicial context, falsely confessing to exam fraud is more ecologically valid than is falsely confessing that one has touched a forbidden key. This has to do with the element of, what lawyers term, *mens rea*, i.e. the presence of explicit intent to do what one is not supposed to do (Kassin & Kiechel, 1996). Inspecting an exam that is not supposed to be inspected is purposeful behaviour. Touching a forbidden key might be an unintended side-effect of routine behaviour. The current study shows that even in a small sample of non-anxious and intelligent undergraduates, some people can be manipulated so as to falsely confess to evidently wrongful behaviour. Interestingly, there was also a false denier. Thus, the current paradigm is not only useful for studying false confessions and true denials, but also to explore false denials. Gudjonsson (2003) has recently pointed out that false denials form an interesting phenomenon. In large Icelandic college and university student samples, Gudjonsson *et al.* (Gudjonsson, Sigurdsson, Bragasson *et al.*, 2004; Gudjonsson, Sigurdsson, & Einarsson, 2004) showed that about 14% of those who had been interrogated by the police reported that they had falsely denied.

General discussion

Our studies provide additional evidence for the claim that one can induce false confessions in healthy, intelligent adults by using subtle tactics. That is, clever undergraduates falsely confessed in the laboratory, even if implausible acts, serious consequences, or *mens rea* were

Table III. Mean scores of participants in study 3 on measures of compliance (GCS) shift score (GSS-shift), suggestibility (GSS-total), fantasy proneness (CEQ), dissociative experiences (DES), and cognitive failures (CFQ) .

	Non-confessors (<i>n</i> = 10)	False confessor (<i>n</i> = 1)	False denier (<i>n</i> = 1)
GCS	8 (SD = 1.8, range: 5–11)	11	14
GSS-shift	5.1 (SD = 2.1, range: 2–9)	3	1
GSS-total	10.3 (SD = 3.7, range: 5–18)	10	2
CEQ	5.4 (SD = 2.4, range: 2–9)	11	3
DES	15.8 (SD = 10.5, range: 3.6–42.7)	18.8	8.2
CFQ	64.1 (SD = 12.1, range: 41–75)	49	67

involved. When serious consequences were involved, confession rates, as predicted, were lower and showed a remarkable resemblance to figures obtained in field studies (e.g. Gudjonsson & Sigurdsson, 1994; Sigurdsson & Gudjonsson, 1996).

There are good reasons to believe that in the real world, false confessions are strongly linked to interrogation tactics used by the police. Thus, Kassin and Wrightsman (1985; p. 77) have argued that pressured-compliant confessions are likely to occur with “powerful and highly salient techniques of social control”. Alternatively, internalizations are expected when “more subtle, less coercive” techniques are used. In our studies, interrogations lasted only a few minutes, with participants being accused approximately twice. These accusations were accompanied by false incriminating evidence (e.g. eyewitness or creased corner). All in all, the tactics used in the present studies were relatively mild compared to, for example, the techniques recommended by some interrogations manuals (e.g. Kassin, 1997). Nevertheless, our participants experienced the whole procedure as stressful. This may have contributed to the pressured-compliant confessions.

Clearly, our findings can also be related to the idea that people’s choice to falsely confess or not depends on a cost–benefit analysis they make (Kassin & McNall, 1991). So, in the face of the interrogation tactics with which they were confronted (e.g. false evidence, etc.), participants will have made an analysis of costs and benefits of confessing. Indeed, during informal interviews after the experiment, most confessors stated spontaneously that their confession was based on a weighing of pros and cons. For example, in our third study, the false confessor was convinced that, eventually, the examination board would conclude that the accusations were unreasonable. So, he saw his confession as an easy way to escape from a painful situation.

In study 1, low plausibility was not related to a significantly lowered false confession rate. However, it was only with high plausible accusations that internalization of false confessions occurred. This is reminiscent of Pezdek, Finger, and Hodge’s (1997) work, which showed that full-blown false memories are easier to elicit when the target event is highly plausible (e.g. lost in a shopping mall) than when the target is a low-probability event (e.g. having had a rectal enema).

A subsidiary aim of the present studies was to explore the relationship between false confessions and individual difference measures of compliance, suggestibility, fantasy proneness, dissociative experiences, and cognitive failures. Note that the mean scores on these measures come close to those reported in earlier studies (Gudjonsson, 1997; Horselenberg, Merckelbach, van Breukelen & Wessel, 2004; Platt, Lacey, Iobst & Finkelman, 1998). Of these measures, only fantasy proneness appeared to be consistently related to false confessions, although in study 2, we also found some tentative evidence for heightened suggestibility levels (i.e. GSS) in the person who false confessed (see also Redlich & Goodman, 2003).

At first sight, it seems odd that fantasy proneness rather than compliance (as indexed by the GCS) is related to false confessions. After all, the large majority of false confessions that we obtained were of the pressured-compliant type (Gudjonsson, 2003). That is, in only a few cases, we found indications that false confessors really came to believe their confessions (i.e. internalized their confession). On the basis of false memory literature, one would expect fantasy proneness to be related to such pressured-internalized confessions, as they can be conceptualized as full-blown false memories (see also Ost et al., 2001). One way to look at the connection between fantasy proneness and false confessions is to assume that the various types of false confessions (i.e. pressured-compliant and pressured-internalized) do not represent distinct categories, but rather points on a continuum. It may well be the case

that an excellent ability to fantasize about a forbidden act contributes via reality monitoring problems to false confessions. Yet, these reality monitoring problems may not be robust and severe enough to be manifested in internalization. At low intensities, they might only give rise to doubts about one's memory and a pressured-compliant confession. With further interrogation and as fantasy-related reality monitoring confusion about the forbidden act grows, such a pressured-compliant confession may develop into a pressured-internalized confession. This point clearly warrants further study.

On theoretical grounds, one could argue that compliance is related to pressured-compliant false confessions, whereas suggestibility is related to pressured-internalized false confessions (Gudjonsson, 2003). However, research addressing this issue has yielded mixed results. While it is true that several case (see for an overview, Gudjonsson, 2003) and field studies (e.g. Gudjonsson & Sigurdsson, 1994; Sigurdsson & Gudjonsson, 1996) found evidence for a connection between compliance, suggestibility, and false confessions, recent work by Gudjonsson *et al.* (Gudjonsson, Sigurdsson, Bragasson *et al.*, 2004; Gudjonsson, Sigurdsson, & Einarsson, 2004) failed to find any relationship between these variables. Such a null finding was also reported by Horselenberg *et al.* (2003). These authors speculated that individual differences in compliance and suggestibility may only come in to play when severe consequences of confessing are involved. Although our studies 2 and 3 met this criterion, they did not yield impressive cues that false confessors are extremely compliant and suggestible individuals. One could, of course, counter that sample sizes in these studies were very small. But, then, in real court cases, the numbers are always small and the question arises to what extent compliance and suggestibility scores of a suspect convey diagnostic information about the probability of a false confession. This issue remains unanswered and warrants further study.

Some limitations of the present studies must be acknowledged. A major concern, besides the ethical issues, is the small sample size in studies 2 and 3. We decided to stop these studies because some participants were distressed by the procedures. The small sample sizes make it difficult to be very precise about the prevalence rates of false confessions in the laboratory and they make it impossible to speculate about the causal antecedents of false confessions. Future studies should focus on the causality issue. Another shortcoming of our studies is their reliance on undergraduate participants. Thus, it would be highly informative to conduct false confession experiments with, for example, forensic patients, i.e. people who probably have elevated compliance and suggestibility scores.

In sum, then, do false confessions occur when plausibility is low and consequences are serious? The answer is: Yes, they do. There are intelligent and healthy undergraduates who falsely confess under these circumstances.

References

- Bedau, H. A., & Radelet, M. L. (1987). Miscarriages of justice in potentially capital cases. *Stanford Law Review*, *40*, 21–179.
- Bedau, H. A., & Putnam, C. (1996). False confessions and other follies. In D. S. Connery (Ed.), *Convicting the innocent* (pp. 69–83). Cambridge, MA: Brookline Books.
- Bernstein, E. M., & Putnam, F. W. (1986). Development, reliability and validity of a dissociation scale. *Journal of Nervous and Mental Disease*, *174*, 727–735.
- Broadbent, D. E., Cooper, P. F., Fitzgerald, P., & Parkes, L. R. (1982). The Cognitive Failure Questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology*, *21*, 1–16.
- Candel, I., Merckelbach, H., & Kuijpers, M. (2003). Dissociative experiences are related to commissions in emotional memory. *Behaviour Research and Therapy*, *41*, 719–725.
- Candel, I., Merckelbach, H., Loyen, S., & Reyskens, H. (in press). 'I hit the Shift-key and then the computer crashed': Children and false confessions. *Personality and Individual Differences*.

- Cassell, P. G. (1996). All benefits, no costs: The grand illusion of Miranda's defenders. *Northwestern University Law Review*, 90, 1084–1124.
- Forrest, K. D., & Wilson, S. L. (2003). Thinking about what you have done in the past: Does rumination elicit false confessions? Paper presented at the International and Interdisciplinary Psychology & Law conference, Edinburgh, 7–12 July 2003.
- Forrest, K. D., Wadkins, T. A., & Miller, R. L. (2002). The role of preexisting stress on false confessions: An empirical study. *Journal of Credibility Assessment and Witness Psychology*, 3, 23–45.
- Gudjonsson, G. H. (1984). A new scale of interrogative suggestibility. *Personality and Individual Differences*, 5, 303–314.
- Gudjonsson, G. H. (1989). Compliance in an interrogation situation: A new scale. *Personality and Individual Differences*, 10, 535–540.
- Gudjonsson, G. H. (1997). *The Gudjonsson suggestibility scales*. Hove: Psychology Press.
- Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions: a handbook*. Chichester: John Wiley & Sons.
- Gudjonsson, G. H., & MacKeith, J. A. C. (1982). False confessions: psychological effects of interrogation. A discussion paper. In A. Trankell (Ed.), *Reconstructing the past: the role of psychologists in criminal trials* (pp. 253–269). Deventer, the Netherlands: Kluwer.
- Gudjonsson, G. H., & Sigurdsson, J. F. (1994). How frequently do false confessions occur? An empirical study among prison inmates. *Psychology, Crime & Law*, 1, 21–26.
- Gudjonsson, G. H., Sigurdsson, J. F., Bragasson, O. O., Einarsson, E., & Valdimarsdottir, E. (2004). Confessions and denials and the relationship with personality. *Legal and Criminological Psychology*, 9, 121–133.
- Gudjonsson, G. H., Sigurdsson, J. F., & Einarsson, E. (2004). The role of personality in relation to confessions and denials. *Psychology, Crime & Law*, 10, 125–135.
- Horselenberg, R., Merckelbach, H., van Breukelen, G., & Wessel, I. (2004). Individual differences in the accuracy of autobiographical memory. *Clinical Psychology and Psychotherapy*, 11, 168–176.
- Horselenberg, R., Merckelbach, H., & Josephs, S. (2003). Individual differences and false confessions: A conceptual replication of Kassir and Kiechel (1996). *Psychology, Crime & Law*, 9, 1–8.
- Horselenberg, R., Merckelbach, H., Muris, P., Rassin, E., Sijsenaar, M., & Spaan, V. (2000). Imagining fictitious childhood events: The role of individual differences in imagination inflation. *Clinical Psychology and Psychotherapy*, 7, 128–137.
- Hyman, I. E., & Billings, F. J. (1998). Individual differences and the creation of false memories. *Memory*, 6, 1–20.
- Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2001). *Criminal interrogation and confessions* (4th edn). Gaithersburg, MD: Aspen.
- Kassin, S. M. (1997). The psychology of confession evidence. *American Psychologist*, 52, 221–233.
- Kassin, S. M., & Kiechel, K. L. (1996). The social psychology of false confessions: Compliance, internalization, and confabulation. *Psychological Science*, 7, 125–128.
- Kassin, S. M., & McNall, K. (1991). Police interrogations and confessions. *Law and Human Behavior*, 15, 233–251.
- Kassin, S. M., & Neumann, K. (1997). On the power of confession evidence: an experimental test of the fundamental difference hypothesis. *Law and Human Behavior*, 21, 469–484.
- Kassin, S. M., & Sukel, H. (1997). Coerced confessions and the jury: An experimental test of the 'harmless error' rule. *Law and Human Behavior*, 21, 27–46.
- Kassin, S. M., & Wrightsman, L. S. (1985). Confession evidence. In S. Kassin & L. Wrightsman (Eds.), *The psychology of evidence and trial procedure* (pp. 67–94). Beverly Hills, CA: Sage Books.
- Kihlstrom, J. F., Glisky, M. L., & Angiulo, M. J. (1994). Dissociative tendencies and dissociative disorders. *Journal of Abnormal Psychology*, 103, 117–124.
- Klaver, J., Gordon Rose, V., & Lee, Z. (2003). Differential effects of minimization and maximization interrogation techniques and the role of plausibility in an experimental false confession paradigm. Paper presented at the International and Interdisciplinary Psychology & Law conference, Edinburgh, 7–12 July 2003.
- Lee, Z., Klaver, J., & Gordon Rose, V. (2003). The role of gender, race, and personality variables in eliciting false confessions. Poster presented at the International and Interdisciplinary Psychology & Law conference, Edinburgh, 7–12 July 2003.
- Leo, R. A., & Ofshe, R. J. (1998). The consequences of false confessions: Deprivations of liberty and miscarriages of justice in the age of psychological interrogation. *Journal of Criminal Law and Criminology*, 88, 429–497.
- Merckelbach, H., Horselenberg, R., & Muris, P. (2001). The Creative Experiences Questionnaire (CEQ): A brief self-report measure of fantasy proneness. *Personality and Individual Differences*, 31, 987–995.
- Merckelbach, H., Muris, P., Rassin, E., & Horselenberg, R. (2000). Dissociative experiences and interrogative suggestibility in college students. *Personality and Individual Differences*, 29, 1133–1140.

- Ofsche, R. J., & Leo, R. A. (1997). The social psychology of police interrogations. The theory and classification of true and false confessors. *Studies in Law, Politics and Society*, 16, 189–251.
- Ost, J., Costall, A., & Bull, R. (2001). False confessions and false memories: A model for understanding retractors' experiences. *Journal of Forensic Psychiatry*, 12, 549–579.
- Ost, J., Vrij, A., Costall, A., & Bull, R. (2002). Crashing memories and reality monitoring: Distinguishing between perceptions, imaginations and 'false memories'. *Applied Cognitive Psychology*, 16, 125–134.
- Pezdek, K., Finger, K., & Hodge, D. (1997). Planting false childhood memories: The role of event plausibility. *Psychological Science*, 8, 437–441.
- Platt, R. D., Lacey, S. C., Iobst, A. D., & Finkelman, D. (1998). Absorption, dissociation, and fantasy-proneness as predictors of memory distortion in autobiographical and laboratory-generated memories. *Applied Cognitive Psychology*, 12, S77–S89.
- Redlich, A. D., & Goodman, G. S. (2003). Taking responsibility for an act not committed: The influence of age and suggestibility. *Law and Human Behavior*, 27, 141–156.
- Scheck, B., Neufeld, P., & Dwyer, J. (2000). *Actual innocence: When justice goes wrong and how to make it right*. New York: Signet.
- 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.
- Spanos, N. P., Burgess, C. A., Burgess, M. F., Samuels, C., & Blois, W. O. (1999). Creating false memories of infancy with hypnotic and non-hypnotic procedures. *Applied Cognitive Psychology*, 13, 201–218.
- van der Ploeg, H. M., Defares, P. B., & Spielberger, C. D. (1980). *Handleiding bij de Zelf-Beoordelings Vragenlijst: Een Nederlandstalige bewerking van de Spielberger State-Trait Anxiety Inventory*. Swets & Zeitlinger BV, Lisse. [Guide for the Self-Assessment Questionnaire: A Dutch version of the Spielberger State-Trait Anxiety Inventory].
- Wilson, S. C., & Barber, T. X. (1982). Fantasy-prone personality: implications for understanding imagery, hypnosis, and parapsychological phenomena. *PSI-Research*, 1, 94–116.
- Winograd, E., Peluso, J. P., & Glover, T. A. (1998). Individual differences in susceptibility to memory illusions. *Applied Cognitive Psychology*, 12, S5–S27.
- Wright, D. B., & Loftus, E. F. (1999). Measuring dissociation: Comparison of alternative forms of the dissociative experiences scale. *American Journal of Psychology*, 112, 497–519.