Behavioral Sciences and the Law
Behav. Sci. Law **31**: 239–255 (2013)
Published online 26 April 2013 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/bsl.2063

"Yes, I Have Sometimes Stolen Bikes": Blindness for Norm-Violating Behaviors and Implications for Suspect Interrogations

Melanie Sauerland, Dr.*[†,‡], Jennifer Maria Schell, MSc[†], Jorg Collaris, BA[‡], Nils Karl Reimer, BA Student[†], Marian Schneider, BA[†] and Harald Merckelbach, Prof. Dr.[‡]

Across two experiments, we studied a phenomenon akin to choice blindness in the context of participants' accounts of their own history of norm-violating behaviors. In Experiment 1, *N* = 67 participants filled in an 18-item questionnaire about their history of norm-violating behaviors (QHNVB). Subsequently, they were questioned about four of their answers, two of which had covertly been manipulated by the experimenter. Of the 134 manipulations, 20 (14.9%) remained undetected concurrently and 13 were accepted in retrospect (9.7%). In Experiment 2 (*N* = 37), we inserted a one-week interval between questionnaire and interview. Twenty-seven (36.5%) of the 74 manipulations remained undetected concurrently and three were accepted in retrospect (8.1%). Data obtained in a four-week follow-up indicated that our manipulations may have long-term effects on participants' perception of their own history of norm-violating behaviors. Implications for the occurrence of false confessions during the course of an interrogation are discussed. Copyright © 2013 John Wiley & Sons, Ltd.

INTRODUCTION

Our lives are full of decisions and most of the time we feel that we are aware of our preferences, choices, and intentions. Recent experimental research on a phenomenon called *choice blindness* (Hall, Johansson, Tärning, Sikström, & Deutgen, 2010; Johansson, Hall, & Sikström, 2008; Johansson, Hall, Sikström, & Olsson, 2005) challenges this view. For example, in one typical choice blindness experiment (Johansson et al., 2005), participants were presented with two stimuli (e.g., female faces) and had to decide which one they found more attractive. After making a decision, participants were handed the chosen face and indicated why they had made this decision. In some trials, however, the participant's choice was manipulated with the help of a magical card trick. As a consequence, participants ended up with the very face they did not choose. Surprisingly, only 13% of the manipulated trials were detected at the

*Correspondence to: Melanie Sauerland, Section Forensic Psychology, Faculty of Psychology and Neuroscience, Maastricht University, P.O. Box 616, 6200 MD Maastricht, the Netherlands. E-mail: melanie.sauerland@maastrichtuniversity.nl
[†]University College Maastricht, The Netherlands
[‡]Faculty of Psychology and Neuroscience, Department of Clinical Psychological Science, Maastricht University, The Netherlands

time of the manipulation (concurrent detection), while 87% of the participants were *choice blind* (i.e., accepted the manipulated item as their own choice).

Interestingly, some participants based their motivation statement on facial features that were absent from the (manipulated) face they held in their hand (e.g., "she had a nice smile" while the person on the photo was not smiling). Likewise, some participants gave reasons that could only refer to the non-chosen face (e.g., refer to a smile when only the non-chosen photo displayed a smile). What is more, the majority of participants who failed to detect the manipulations (84%) expressed the belief that they would be able to do so during a post-test interview. This metacognition phenomenon was dubbed *choice blindness blindness* (Johansson et al., 2005).

Merckelbach, Jelicic, and Pieters (2011b, Experiment 2) used a paradigm akin to the choice blindness paradigm to test participants' blindness for their own psychological symptoms. Specifically, participants were asked to fill in the Symptom Checklist-90 (SCL-90; Derogatis, Lipman, & Covi, 1973), which measures a wide range of psychological symptoms on a five-point Likert scale. During a short break, the experimenter discreetly increased two of the answers given by two full scale points. Next participants were asked to explain their answers to the two manipulated target items as well as to the eight control items. The results showed that 57% of the participants were blind to both manipulations (75% for target item 1, 68% for target item 2). Contrary to baseline testing, blind participants showed a tendency to score higher on the target items, compared with control items, when filling in a short 30-item version of the SCL-90 at a later time. This suggests long-term effects of the conducted manipulations.

Merckelbach, Jelicic, and Pieters (2011a) replicated these findings, with 74% of the participants being blind to both manipulations. In this study, follow-up scores were obtained twice, once about 10 minutes after baseline testing and then again after a week. Parallel to the earlier study, the target item scores were higher in blind participants than control item scores at follow-up testing, while there was no such difference at baseline testing. There were no such differences for non-blind participants at any of the three testing times. Both studies (Merckelbach et al., 2011a, 2011b) lend support to the hypothesis that symptom manipulations can have long-term effects on participants' symptom strength perception. Interestingly, blind participants scored higher on target and control items from the very beginning, indicating that a non-zero symptom level introduces a degree of ambiguity that increases the likelihood of being blind for manipulations.

Indeed, recent research transferring the choice blindness paradigm to eyewitness identifications supports the idea that ambiguity plays an important role for blindness phenomena. Specifically, the participants of Sagana, Sauerland, and Merckelbach (manuscript submitted for publication) watched four mock crime videos and then identified the actors (four per video) from individual target present lineups. Following their forced-choice identification decision, participants were presented with their selection and gave reasons for their decision. Two of the 16 identifications were manipulated, with participants ending up with a non-chosen lineup member. As expected, and in line with the ambiguity notion, retrospective detection was high (94–100%) when no time interval was inserted between making the identification decision and being confronted with the manipulated outcome (Experiments 2a–c). However, retrospective detection deflated to 61% when a 48 h interval was introduced (Experiment 3). So it seems that when memory becomes fuzzier, and therefore the situation more ambiguous, participants are more likely to be blind to manipulations. Importantly, this study also provided additional evidence for the long-term effects of the manipulations: specifically, identification accuracy was decreased for manipulated but not non-manipulated lineups three to five months after the initial test.

DOI: 10.1002/bsl

Taken together, these studies (Merckelbach et al., 2011a, 2011b; Sagana et al., manuscript submitted for publication) help us to understand the scope of choice blindness phenomena. First, they demonstrate that choice blindness is not limited to *preferences* (Hall et al., 2010; Johansson et al., 2005), but can also occur for *intensity* and *frequency ratings* as well as *objective decisions*. Accordingly, these studies provide evidence that (choice) blindness is a phenomenon that also applies to evaluations that rely on long-term episodic memory. Second, the results of the follow-ups (Merckelbach et al., 2011a; Sagana et al., manuscript submitted for publication) suggest that manipulations not only may go unnoticed during an interview, but also may actually have long-term effects on participants: some participants seem to internalize the new answer, as indicated by their symptom escalation and wrongful identification decisions on a later test.

Internalization of false statements induced by misinformation is not only typical for choice blindness in the context of symptom or lineup evaluations. It is also a key feature of interrogation situations (Kassin & Gudjonsson, 2004). With this in mind, we wanted to link the choice blindness phenomenon to the false confession literature. Confessions are one of the most potent forms of evidence in criminal law (Kassin & Neumann, 1997). False confessions have often been found to be a result of coercive and deceptive interrogation methods (see, e.g., Kassin, 2005). While presenting false evidence to a suspect is unlawful in most European countries, a 1969 ruling of the US Supreme Court (*Frazier v. Cupp*, 1969) explicitly allowed the presentation of misleading evidence in police interrogations. Deceptive tactics have since then been an important part of the widely taught, controversial Reid technique for criminal interrogations (Inbau, Reid, Buckley, & Jayne, 2001). However, misleading suspects by confronting them with false evidence during an interrogation has been proven to be one of the causes of false confessions (Kassin et al., 2010).

To test the effect of false incriminating evidence experimentally, Kassin and Kiechel (1996) introduced the so-called computer-crash paradigm. Allegedly as a test of reaction time, participants were instructed to type letters on a computer keyboard that were read aloud by a confederate. Participants were warned not to press the ALT key, because this would cause a computer crash and data loss. One minute into the task, the computer automatically crashed. In the ensuing interrogation, the experimenter falsely accused the participant of having caused the crash by pressing the ALT key. Overall, 69% of the 75 participants signed a confessions form. When introducing the confederate's false witness statement as evidence, the rate of confessions was significantly higher than when no false evidence was presented (range 89–100% versus 35–65% in the no witness conditions). Judging from participants' private statements towards another confederate, 55% of the participants who were confronted with false evidence actually internalized their confessions, compared with 6% in the group without false witness evidence. The memory corrupting effect of false evidence has also been observed in other studies (e.g., Horselenberg, Merckelbach, & Josephs, 2003; van Bergen, Jelicic, & Merckelbach, 2008), and some authors have argued that it is primarily carried by false evidence promoting a state of memory distrust, in which people become more willing to confess to things that they did not do (van Bergen, Horselenberg, Merckelbach, Jelicic, & Beckers, 2010).

Obviously, false confessions are by no means always caused by internalization and memory corruption. Research suggests that another mechanism that can lead innocent suspects to confess because of false evidence is the feeling of being trapped by the weight of the (false) evidence provided during the interrogation. In such cases suspects

DOI: 10.1002/bsl

may perceive a confession as the only way out of the situation (Gudjonsson & Sigurdsson, 1999; Moston, Stephenson, & Williamson, 1992; Perillo & Kassin, 2011). Both mechanisms can play a role in choice blindness phenomena. Consider an interrogation situation in which an interrogator – intentionally or unintentionally – keeps a record of answers different from those given by the suspect. Confronted with these changes in a later interrogation, the suspect might accept these (false) answers, either as a result of internalization and memory corruption or due to a feeling of being trapped. In the current studies, we wanted to apply the choice blindness paradigm to an interviewing situation. Specifically, we explored whether people would admit to norm-violating behaviors as a result of misinformation given about their own previous accounts. To this end, we adopted a research paradigm akin to the choice blindness paradigm and to the symptom misinformation procedure employed by Merckelbach et al. (2011a, 2011b). Specifically, we questioned participants about their own accounts of the frequency of certain norm-violating behaviors in their personal history after we had covertly inflated or deflated their previous answers. Given the evidence on choice blindness (Hall et al., 2010; Johansson et al., 2005), blindness concerning one's own psychological symptoms (Merckelbach et al., 2011a, 2011b), and one's eyewitness identification decisions (Sagana et al., manuscript submitted for publication), we expected to find a significant proportion of our participants to be blind for our manipulations. Analogous to Sagana et al. (manuscript submitted for publication), we also tested the effect of ambiguity by varying the retention interval between filling in the questionnaire and being interviewed about it. In Experiment 1, participants were questioned within minutes after giving their answers; in Experiment 2, a one-week interval was inserted. Accordingly, we expected higher rates of choice blindness with the longer retention interval. Additionally, Experiment 2 included a four-week follow-up to test for long-term effects of the manipulations. We expected that blind participants would tend to adopt the misinformation during follow-up sessions.

EXPERIMENT 1

Methods

Participants

A total of $N = 67$ participants (12 men, 55 women; $M_{\text{age}} = 20.7$ years, $SD_{\text{age}} = 2.8$, age range 18–38 years) took part in the experiment. All participants were undergraduates at Maastricht University, most of them (96%) at the Faculty of Psychology and Neuroscience. Participants received course credits or a small financial reimbursement for their participation. The study was approved by the standing ethical board of the Faculty of Psychology and Neuroscience, Maastricht University.

Design

Severity of norm-violating behaviors (low versus moderate versus high) was manipulated as a between-subjects variable, while the status of the items (manipulated target items versus non-manipulated control items) was a within-subjects variable.

DOI: 10.1002/bsl

The detection rate of the manipulated items was the dependent variable. Two different measures of detection were used, namely concurrent and retrospective. The concurrent detection rate includes all participants who noticed the manipulation immediately after it had taken place. Retrospective detection additionally includes all those participants who reported that they had noticed the change of the item ratings in the post-test questionnaire.

Materials

Questionnaire about History of Norm-Violating Behaviors (QHNVB). This Dutch questionnaire consisted of 18 items concerning norm-violating behaviors (see the Appendix). Nine items were taken from Jelicic, Merckelbach, Timmermans, and Candel (2004), and we added nine more. In a pilot study, $N = 15$ psychology students (eight women, age 18–25) rated the severity of the 18 behaviors on a five-point Likert scale (range 1–5). Based on the results, we created three conditions in which (1) two less severe (items 6 and 12), (2) two moderately severe (items 7 and 13), or (3) two highly severe (items 8 and 14) behaviors were manipulated. The severity scores of these behaviors can be found in Table 1. The means of each severity level differed significantly from the means of the two other severity levels, all $F(1, 14) \geq 10.42$, $p \leq .006$.

We used two different versions of the QHNVB. In version 1, participants ($n = 24$) had to answer whether or not they had ever engaged in the behavior described by the item (*yes–no* format). The answers to the manipulated items were hence changed from yes to no or from no to yes, depending on the answers given by the participant. In version 2 (shown in the Appendix), participants ($n = 43$) indicated whether they had displayed the behavior *never*, *seldom*, *sometimes*, or *often* (four-point format, coded 0–3). Here, the target items were increased by two scale points (i.e., from never to sometimes and from seldom to often). In those instances in which participants had scored the critical items with *sometimes* or *often*, the manipulation consisted of decreasing the items by two full scale points (i.e., they were recoded as *never* or *seldom*, respectively).

The answer format had no impact on either concurrent or retrospective detection for both manipulated items, all $chi^2(1, N = 67) \leq 3.68$, $p \geq .081$, $|phi| \leq .23$ (Fisher's exact test due to expected counts < 5). We shall therefore not discuss this response-format manipulation any further.

Marlowe–Crowne Social Desirability Scale (MCSDS). The MCSDS (Crowne & Marlowe, 1960) is a 33-item true/false self-report inventory that aims at assessing an individual's need for approval. Examples of items are "I like to gossip at times" and

Table 1. Severity scores of the manipulated items on the Questionnaire about History of Norm-Violating Behaviors

Item		*M*	*SD*
6	Cheat in high-school exam	2.20	1.15
12	Keep silent about too much change	2.40	1.30
7	Use public transport without ticket	3.27	1.16
13	Park in disabled parking	3.47	0.99
8	Shoplifting	4.53	0.52
14	Bike theft	4.35	0.64

Note: Items 6 and 12, 7 and 13, or 8 and 14 were manipulated together.

DOI: 10.1002/bsl

"I never resent being asked a favor". In the current study, we administered the MCSDS to examine whether blindness to manipulations is associated with the tendency to act according to social demands, meaning that people might fail to report that they did notice the change in manipulated items due to their need for approval from the experimenter.

Gudjonsson Compliance Scale (GCS). The GCS (Gudjonsson, 1989) is a self-report inventory consisting of 20 true/false statements such as "People with a lot of authority make me feel uncomfortable" and "I try to please others". It is used to measure the tendency of people to conform to other people's requests. The instrument was included to control for the possibility that participants were not blind to the manipulations, but rather complied with the presumed expectations of the experimenters. The same holds for the GSS.

Gudjonsson Suggestibility Scale (GSS). We used the Dutch research version (Merckelbach, Muris, Wessel, & van Koppen, 1998) of the GSS (Gudjonsson, 1984, 1997). It consists of 20 questions concerning a short story that was previously read to the participants, of which 15 are suggestive. After answering all questions, participants are given negative feedback on their performance and all questions are repeated once again. *Yield1* and *yield2* scores indicate the number of suggestive items to which a participant agreed during the first and second rounds of questions, respectively. The *shift* score indicates how many answers were changed between the two questioning rounds. Finally, the *total score* consists of the sum of the yield1 and shift score.

Abbreviated Big Five Questionnaire. This questionnaire consists of 15 items (Furnham, McManus, & Scott, 2003) that were drawn from the original 60-item version (NEO-FFI; McCrae & Costa, 1997). It provides measures of Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness, which are measured by three items each. We included this test to explore how personality traits relate to blindness for our manipulations.

Post-test questionnaire. The post-test questionnaire was adjusted from Johansson et al. (2008). Specifically, participants were first asked whether they had noticed anything strange during the experiment, then whether they had noticed anything odd with the answers to the QHNVB. If participants had noticed something strange, they were asked to explain. Explanations indicating that the manipulations had been detected were counted as retrospective detection. If, however, no signs of detection were revealed, participants were told that we planned a follow-up study in which the answers would sometimes be manipulated and were asked if they would notice such a change. The answer to this question is referred to as the degree of *blindness for one's blindness* in the results section. Finally, participants were explicitly asked if they had noticed such a manipulation in the current experiment. If this question was answered with yes, this was also coded as retrospective detection.

Procedure

Participants signed the informed consent form and were then asked to fill in the QHNVB. Next, they were given two Sudoku puzzles to solve. In the meantime, the experimenter manipulated two of the given answers as described in the QHNVB section above. When the experimenter came back to the laboratory after about 10 minutes, he asked the participant questions about the two manipulated items and about two non-manipulated control items. Here is an example of questions asked: "You answered that

DOI: 10.1002/bsl

you have *often* cheated on a high school test. Could you elaborate on that? Why did you do that?". Subsequently, the experimenter would ask "Why do you think other people would cheat on a high school test?". Depending on participants' answers, follow-up questions were asked to be certain about whether or not the manipulation had been discovered. All in all, participants were asked about three or four questions per item. The non-manipulated items were items 1 and 10. The order of questioning was always control item 1, first manipulated item (item 6, 7, or 8), control item 10, and second manipulated item (item 12, 13, or 14).

Hereafter, participants worked on part 1 of the GSS ($n = 48$) or filled in the abbreviated Big Five questionnaire ($n = 19$). The MCSDS and the GCS followed for all participants. Those participants who had the GSS administered then worked on the second part of the GSS. Finally, participants filled in the post-experimental questionnaire.

Results and Discussion

An alpha level of .05 was used for all inferential analyses. For comparisons of means, we report d as a measure of effect size. For non-parametric analyses of 2×2 contingency tables we report phi; for 3×2 contingency tables we report Cramer's V.

Blindness for Manipulations

The first manipulation was concurrently detected by 82.1% ($n = 55$) of the participants, the second manipulation by 88.1% ($n = 59$) of the participants. Three participants did not detect either of the two manipulations concurrently. In total, of the $2 \times 67 = 134$ manipulations, 114 (85.1%) were detected. Note that 92 (68.7%) of the manipulations involved inflations (i.e., upscaling) of the initial answers given and 42 (31.3%) involved deflations (downscaling). This did not have an impact on concurrent or retrospective detection rates, all $p \geq .431$, phi $\leq .12$.

One may argue that the reported detection rates are inflated since participants become suspicious after the first detection is made (Johansson et al., 2005). To avoid such cascading detection effects, we discarded all second manipulation items when the first one had been detected. With this method, 64 out of 79 manipulations were detected (81.1%). This distribution differed significantly from that expected, Cochran's $Q(1) = 23.05$, $p < .001$. Accordingly, we ran all analyses for these 79 trials as well as with the full 134 trials. The results for the 79 trials generally paralleled the results of the 134 trials. We shall therefore report the results for the full 134 trials.

Retrospective detection was 89.6% ($n = 60$) for the first manipulated item and 91.0% ($n = 61$) for the second manipulated item. Only one participant detected neither manipulation retrospectively, 12 detected one, and 54 detected both manipulations.

To explore whether severity of the manipulated items had any effect on detection rates, we computed four chi-square tests with severity (low, moderate, high) as independent variable and concurrent and retrospective detection as dependent variables. Due to some expected cell counts being less than 5, we computed Fisher's exact test. No significant effects were found, all $p \geq .240$, Cramer's $V \leq .14$.

Following Merckelbach et al. (2011a, 2011b), we established whether blind participants had higher initial scores on the QHNVB for the four-point version of the QHNVB. (No such analyses were run for the yes–no version due to the small number of detections within that sample and the resulting lack of power.) No significant effects

DOI: 10.1002/bsl

were found for concurrent or retrospective detection of the first manipulated item, all $|t|(41) \leq 0.52$, $p \geq .606$, $|d| \leq 0.21$. Contrary to our expectations, the mean baseline item rating for detectors was (non-significantly) higher ($M = 1.06$, $SD = 0.94$) than that of non-detectors ($M = 0.88$, $SD = 0.64$). For retrospective detection, the mean baseline item rating for detectors was (non-significantly) smaller ($M = 1.00$, $SD = 0.93$) than that of non-detectors ($M = 1.20$, $SD = 0.45$), as expected. Participants who were concurrently or retrospectively blind to the second manipulation, however, did score significantly higher (all $M = 1.00$, $SD = 0.71$) than non-blind persons on the QHNVB (all $M = 0.29$, $SD = 0.61$), $t(41) = 2.40$, $p = .021$, $d = 0.90$. These results lend some support to the hypothesis of Merckelbach et al. (2011a, 2011b) that a non-zero answer level introduces a degree of ambiguity that increases the likelihood of being blind. This finding is particularly worrying when translating it to the context of interrogations and the risk of false confessions. It suggests that suspects who already have a criminal record are ultimately more likely to falsely confess than those who do not. An interesting topic for future research would be whether partial acceptance makes people more prone to accept other false information offered during the course of an interrogation.

Blindness for One's Blindness

Of the participants who did not concurrently detect the first manipulation ($n = 12$), 91.7% thought they would be able to do so (retrospective detection: 85.7%, $n = 7$). Of the participants who did not concurrently detect the second manipulation ($n = 8$), 62.5% thought they would be able to do so (retrospective detection: 50.0%, $n = 6$).

Compliance, Suggestibility, and Social Desirability Measures and Abbreviated Big Five Questionnaire

Although t tests are most suited to test for individual differences, we reverted to correlational analyses due to the small number of blind participants. Concurrent and retrospective blindness (scored as 0, 1, or 2 detections) were not significantly associated with social desirability, compliance, or the GSS yield, GSS shift, and GSS total score, all $|r| \leq .17$, $p \geq .247$. These results contradict the idea that blindness for manipulations of one's own responses can be attributed to social demands, compliance, or suggestibility. This is in agreement with earlier studies that did not find an association between blindness and social desirability (Merckelbach et al., 2011a) as well as compliance (Sauerland, Sagana, & Otgaar, 2012). To our knowledge, this is the first study to show that blindness is also not associated with suggestibility.

To explore the possible link between blindness and personality traits, some participants ($n = 19$) filled in the Abbreviated Big Five Questionnaire. No significant correlations between blindness and four of the five factors were found, all $|r| \leq .25$, $p \geq .306$. There was a significant correlation of Openness with concurrent and retrospective detection, though, all $|r|(17) \geq -.59$, $p \leq .007$, indicating that more detection occurred when participants had lower Openness scores. This result is somewhat counterintuitive, as one would expect that Openness should be associated with higher detection rates, not lower ones. On the other hand, Openness to Experience is a higher order trait involving the lower order trait of fantasy proneness. Previous work has shown that those with elevated fantasy proneness levels also have heightened false confession rates

DOI: 10.1002/bsl

(Horselenberg et al., 2006), probably because it is easier for these people to see the plausibility of misinformation.

To summarize, blindness for one's own history of norm-violating behaviors seems to be much weaker than choice blindness for facial attractiveness or other preferential choices (Hall et al., 2010; Johansson et al., 2005, 2008). The base rate of blindness found here was also smaller than reported for psychological symptoms (Merckelbach et al., 2011a, 2011b). One reason for the differences found in blindness between the current and the psychological symptom study could be that perhaps different levels of ambiguity were achieved. For example, in the latter study, participants answered 90 questions, using a five-point scale, and were interviewed about ten items, two of which were manipulated. In the current study, participants answered 18 questions, using a two- or four-point scale, and were interviewed about four items, again, two of which were manipulated. Thus the difference between the given and the manipulated answer was slightly less extreme in the psychological symptom study. Further, the proportion of manipulated items, in proportion to the total number of questions and also in proportion to the number of items that were discussed during the interview, was much larger in the current study.

Be that as it may, it cannot be sugarcoated that about 20% of the manipulations remained undetected concurrently and about 10% retrospectively. A rate of 10–20% of suspects who go along with investigator-introduced manipulations constitutes a serious challenge to the probative value of confessions and raises issues regarding interrogation practice-related policy.

One limitation of Experiment 1 is that there were only minutes between filling in the QHNVB and being confronted with one's (manipulated) answers. Therefore, we inserted a one-week interval in Experiment 2. This situation has more ecological validity, as a suspect is commonly interrogated repeatedly on different days (Johnson & Drucker, 2009; Kassin et al., 2007; Wagenaar, 2002). This should also increase ambiguity and thus lead to increased blindness rates in Experiment 2, compared with Experiment 1.

Furthermore, we included a follow-up session after four weeks in Experiment 2. Specifically, we contacted participants via email four weeks after the experiment had taken place and asked them to fill in the QHNVB once again. This allowed us to draw conclusions about the long-term effects of the manipulations. Drawing from Merckelbach et al. (2011a, 2011b) and Sagana et al. (manuscript submitted for publication), we expected that answers to manipulated items should shift more often between baseline testing and follow-up testing, compared with answers to non-manipulated items.

EXPERIMENT 2

Methods

Participants

A total of $N = 40$ participants were tested. Three had to be excluded from the analysis, because the experimenter had accidentally manipulated the wrong item ($n = 2$), and because one participant turned out to be underage. The sample thus included $N = 37$

DOI: 10.1002/bsl

participants (14 men, 23 women; $M_{age} = 21.8$ years, $SD_{age} = 2.1$, age range 18–26). All participants were undergraduates of University College Maastricht ($n = 32$) or Maastricht University ($n = 5$) and received a candy bar for their participation. Participants were randomly assigned to one of the three conditions (low severity, $n = 14$; moderate severity, $n = 11$; high severity, $n = 12$), with gender being counterbalanced. The study was approved by the standing ethical board of the Faculty of Psychology and Neuroscience, Maastricht University.

Procedure

The procedure was similar to Experiment 1 with the exception that there were two sessions as well as a follow-up session about four weeks after the second session. In session 1, participants filled in version 2 (four-point scale) of the QHNVB and the abbreviated Big Five questionnaire. Since the default language at the University College is English and many students at this College do not speak Dutch, we used the English version of the QHNVB. Session 2 was scheduled one week (±2 days, $M = 167.4$ h) later. Here, participants were interviewed about their answers given to the QHNVB. They also filled in the GCS, the MCSDS, and the post-test questionnaire. Five to six weeks after the experiment had taken place, participants were contacted via email and asked to fill in the QHNVB again.

Results and Discussion

Blindness for Manipulations

Of the $2 \times 37 = 74$ manipulations, 47 (63.5%) were detected. The first manipulation was concurrently detected by 64.9% ($n = 24$) of the participants, the second manipulation by 62.2% ($n = 23$) of the participants. As previously reported (Johansson et al., 2005), blind participants could often provide a rationale for their alleged answer (confabulation). Seven participants did not detect any of the two manipulations, of whom three also failed to detect the mismatch retrospectively. Note that we did not ask participants separately about their retrospective detection of the two manipulated items. Rather, retrospective blindness was scored dichotomously (i.e., blind or non-blind rather than detected 0, 1, or 2 items). Thus, as expected, the concurrent detection rates in Experiment 2 (63.5%) were smaller than those observed in Experiment 1 (85.1%), $\text{chi}^2(1, N = 208) = 11.45$, $p < .001$, Cramer's $V = 0.25$. The retrospective detection rates were almost identical, however (9.7% versus 8.1%). Across all 74 manipulations, 61 (82.4%) involved upscaling of the initial answer given and 13 (17.6%) involved downscaling. This did not have an impact on concurrent detection rates, $p \geq .136$, $\text{phi} \leq .25$.

As in Experiment 1, we tested for possible inflated detection rates due to cascading effects. When discarding second manipulated items when the first one was detected, 30 out of 50 manipulations were detected (60.0%). This distribution did not differ significantly from that expected, Cochran's $Q(1) = 0.16$, $p = .691$. It is therefore unproblematic to continue our analyses with the full 74 trials.

It is interesting to look deeper into some of the reactions that came from detectors when they were confronted with the manipulated answers. There were 35 trials (74.5%) during which participants did not continue until they had been given the

DOI: 10.1002/bsl

chance to change "their" answer on the questionnaire. In another 7 trials (14.9%), participants accepted the answer only after they had been shown "their" answer on the questionnaire. Finally, in 5 trials (10.6%), participants expressed some doubts, but continued. The last two categories are a cause of concern when translated to the context of interrogations. While these participants actually detected the manipulation (and were accordingly coded as detectors), they nevertheless accepted it. This behavior can be seen as analogous to a compliant confession in the context of an interrogation. We shall return to this point in the general discussion.

To explore whether the severity of the manipulated items plays a role in blindness, we computed two 2×3 chi-square tests with severity (low, moderate, high) as independent variable and concurrent detection as dependent variable. No analyses were run for retrospective detection because of the small number of participants who failed to detect the change retrospectively. Due to some expected cell counts being less than 5, we computed Fisher's exact test. No significant effect was found for the first manipulated item, $chi^2(2, N=37)=2.65$, $p=.270$, Cramer's $V=0.27$. There was, however, a significant effect for the second manipulated item, $chi^2(2, N=37)=6.58$, $p=.033$, Cramer's $V=0.43$. Post-hoc 2×2 chi-square tests showed that there was a significant difference in the number of detections when comparing the most severe and the least severe norm-violations, $chi^2(1, N=26)=6.00$ $p=.021$, $phi=0.48$. For more severe norm-violations there were only 2 out of 12 blind participants (16.8%); for the least severe norm-violations there were 9 out of 14 participants (64.3%). None of the other comparisons reached significance, $p \geq .111$.

The fact that there was a difference in detections as a function of severity for the second, but not the first, manipulated item could have to do with the exact items we used. For example, there might have been less ambiguity for the second manipulated item, compared with the first one. Inspection of the data shows that the baseline scores of manipulated item 2 were smaller on average ($M=1.54$, $SD=0.73$) than those of manipulated item 1 ($M=2.05$, $SD=0.78$), $t(36)=3.07$, $p=.004$, $d=0.50$. As mentioned before, it seems straightforward that potential for misleading information is greater with behavior that has a relatively high prevalence rate, compared with a relatively low prevalence rate (Merckelbach et al., 2011a, 2011b). It speaks in favor of this interpretation that participants who were blind to the second manipulation had initially scored higher ($M=1.86$, $SD=0.66$) than non-blind persons ($M=1.35$, $SD=0.71$), $t(35)=2.16$, $p=.038$, $d=0.61$). No such difference was found for the first manipulated item, $t(35)=-1.68$, $p=.103$, $d=-0.51$. This is consistent with Experiment 1 and provides further support to the hypothesis of Merckelbach et al. (2011a, 2011b) that a non-zero level increases the likelihood of being blind. Furthermore, this reinforces our concern raised above regarding the use of deceptive information by investigators in the course of an interrogation.

Blindness for One's Blindness

Of the participants who did not concurrently detect the first manipulation ($n=13$), 38.5% thought they would be able to do so. For the second manipulation ($n=14$), 50.0% gave this answer. For retrospective detection, one out of the three blind participants (33.3%) thought s/he would be able to detect the manipulation. As in Experiment 1, blindness for one's own blindness was found. The effect was, however, smaller. This could be expected, since participants were probably more aware of their own proneness to be blind to changes due to the inserted interval of 1 week.

DOI: 10.1002/bsl

Follow-up Testing

One could argue that non-detection does not reflect blindness, but simply poor memory. That is, participants can simply not remember what answers they gave. While it is unlikely that participants would forget whether or not they have, for example, stolen a bike, some participants may have decided to embellish[1] their answers on the QHNVB. These participants may later have had problems with remembering on which level of embellishment they had decided. If this were the case, one would expect shifts in the given answers from the initial testing to the follow-up testing for both manipulated and non-manipulated items.

Fourteen (37.8%) participants responded to our follow-up email. We inspected their responses to the two non-manipulated (control) items about which participants had been interviewed (items 1 and 10) as well as to the two manipulated items. The results are displayed in Figure 1. We found no shifts, that is, deviations from the initial score, in the responses to items 1 and 10. Five shifts out of a possible 28 (17.9%) did, however, take place for the two manipulated items (1 referring to the first manipulation, 4 to the second one). These shifts referred to two upwards as well as two downwards manipulations and, surprisingly, one of the shifts did not concur with the direction of the manipulation. Note that none of these five shifting participants were retrospectively blind for the manipulation. Three had been blind concurrently, whereas two had detected the manipulation on the item on which they had now shifted.

These results speak against the interpretation that the blindness for manipulations that we observed is merely a memory problem. Rather, participants seem to be quite reliable in their statement regarding their history of norm-violating behaviors, as long as their answers are not manipulated. If their responses are manipulated, however, these manipulations can have long-term effects on participants' statements regarding the frequency of them having been involved in certain norm-violating behaviors. Importantly, this is regardless of whether or not participants retrospectively detected the manipulations. We are aware of the fact that we are looking at a very small sample here.

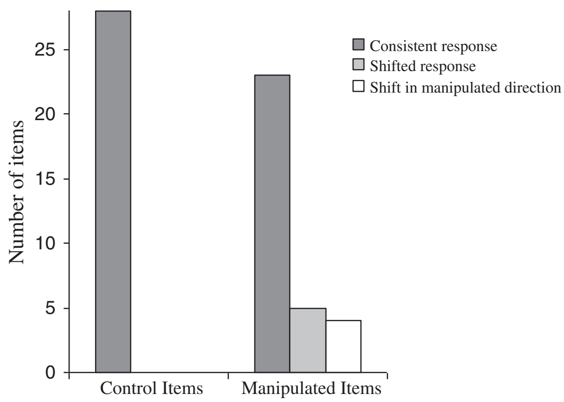

Figure 1. Number of answer shifts occurring for two manipulated and two control items after a four-week follow-up ($N = 14$).

[1] We would like to thank José Wijnands for bringing this point to our attention.

DOI: 10.1002/bsl

Nevertheless, we believe that the results give some interesting insights.

Compliance, Social Desirability, and Abbreviated Big Five Questionnaire

Parallel to Experiment 1, concurrent blindness (scored as 0, 1, or 2 detections) did not correlate significantly either with social desirability or with compliance, all $r \leq |.30|$, $p \geq .071$. When treating participants who detected only one manipulation as non-detectors, t test revealed a marginally significant effect for compliance, $t(35) = -1.85$, $p = .051$, $d = -0.67$. The effect was contrary to the expectations, though, with blind participants being less compliant ($M = 8.15$, $SD = 3.08$) than non-blind participants ($M = 10.00$, $SD = 2.37$). Furthermore, there was a significant effect for social desirability, $t(34) = 2.35$, $p = .025$, $d = 0.79$. As one would expect, social desirability scores were higher for blind participants ($M = 17.35$, $SD = 4.79$) compared with non-blind ones ($M = 13.63$, $SD = 4.66$).

We found no correlations between concurrent blindness and the five personality factors, all $r \leq |.20|$, $p \geq .225$. Analogously, no significant effects were revealed when analyzing the data with t tests. Thus, personality traits as measured by the Abbreviated Big Five Questionnaire do not seem to be a relevant factor when it comes to blindness. With regards to Openness, this result contradicts Experiment 1.

GENERAL DISCUSSION

In two experiments, we tested the hypothesis that a phenomenon akin to choice blindness can occur in the context of people's accounts of their own history of norm-violating behaviors. To this end, participants indicated how frequently they had been involved in certain norm-violating behaviors in the past. As expected, a substantial proportion of participants was blind for manipulated increases or decreases of their own statements. The current studies thus demonstrate that choice blindness phenomena are relevant not only with regards to medical practitioners or lineup administrators, but also for interrogators. The degree of concurrent blindness varied as a function of the time interval inserted between filling in the questionnaire and being interviewed about it. As expected, a longer time interval resulted in reduced concurrent detection, whereas the degree of retrospective detection did not change. Essentially, participants should still remember how often they had been involved in certain norm-violating behaviors, regardless of how long ago they made statements about this. The longer time interval, however, seems to have introduced a level of ambiguity about the answers, leading to a higher threshold to query about answers presented by the experimenter. Once participants were apprised of the fact that some answers had been manipulated, however, they felt free to express their doubts. This is, to some extent, good news: while concurrent detection referring to norm-violating behaviors seems to be prone to factors that introduce ambiguity, retrospective detection seems to be relatively constant, at least across our two experiments.

The results of our follow-up test invalidate the hypothesis that participants can simply not remember their answers. Rather, retest consistency was found to be 100% for non-manipulated items, whereas some participants shifted their answers for manipulated items in a follow-up about a month after the initial testing. This speaks for long-term effects of misinformation manipulations on statements as obtained in the

DOI: 10.1002/bsl

current study and should be of concern for legislations in which deception during interrogations is common practice.

Our results have important implications for the treatment of official statement records. These documents play a crucial role in legal proceedings. The implicit assumption is that these records reflect the account of the suspect, eyewitness, or victim in a correct way. In many jurisdictions, the persons concerned are given the opportunity to re-read their accounts after making a statement, and authors have argued that this is beneficial as it increases statement consistency (e.g., Magner, Markham, & Barnett, 1996). However, the upshot from the current findings is that suspects or eyewitnesses may be blind at a low, but non-trivial, rate when confronted with legal documents that contain elements that they did not introduce themselves.

Two of our findings could raise doubts about the relevance of our findings for the field of false confessions: the relatively small proportion of retrospective blindness and the finding that blindness can vary as a function of the severity of the manipulated behavior. Both can be rebutted, we believe. First, although blindness in the interrogation room may be a rare phenomenon, it might concern a large number of cases in absolute numbers, since thousands of suspects and eyewitnesses are interviewed each year. Furthermore, inspection of a mock jury study (Kassin & Sukel, 1997, Experiment 2) leads to the conclusion that after-the-fact corrections (i.e., detecting in retrospect, but not concurrently) matter little in criminal proceedings. In the study by Kassin and Sukel, mock jurors' verdict was more often in favor of the prosecution (i.e., "guilty") when a confession was presented, compared with a control condition where no confession was presented. This was the case even though mock jurors accepted the confession as involuntary and explicitly denied any influence on their verdict. Hence, even if defendants' incriminating statements are retracted and dismissed on the basis that they were caused by police misinformation, the influence of such dismissal on the jury might be negligible. This holds especially in legal systems in which the presentation of false evidence is considered legitimate (e.g., the United States).

Turning to the second point, one may assume that more severe (rather than less severe) behaviors are frequently the subject of confession-related false convictions. However, interrogations are a process in which law enforcement agents try to either obtain an explicit confession or extract incriminating statements. Statements of the latter kind, for example about the defendant having been at the crime scene, do not directly refer to the crime itself but can, nevertheless, serve as potent evidence against an innocent defendant in court. These, presumably less severe statements can be conceived as similar to less severe norm-violations. Confabulations or lack of explicit reaction (as found in Experiment 2) following deliberate misinformation by the police might be interpreted as silent admission of guilt or as contradictive, increasing law enforcement agents' suspicion. Hence, the effect of the severity of allegations on detection does not necessarily negate the importance of the phenomenon for false convictions. Also note that the effect of severity was found for Experiment 2 only. Further research will have to find ways to test whether indirectly incriminating statements such as the one described above are subject to choice blindness.

We believe that the current results have research and policy related implications. Future studies should particularly focus on three aspects of false confession related blindness. The first revolves around the question at which level of severity manipulated statements fall short of being noticed, while they still have the potential to serve as incriminating evidence. Second, it would be interesting to investigate whether participants show blindness not only with

DOI: 10.1002/bsl

regard to norm-violating behaviours, but also to alibi-related statements. Research could thus not only investigate prevalence statements, but also statements related to the details of certain locations and times. Third, case studies that illustrate blindness phenomena during real-life interrogations could inform the field.

Looking at the practice of law enforcement, our findings corroborate several warnings that have been issued by other scholars in the field of law and psychology (Kassin, 2008; Kassin & Kiechel, 1996). Investigators should beware of deceiving the suspect and confronting him or her with misinformed statements. As shown, innocent suspects can easily fail to detect the misinformation, and stay blind to the mismatch between their original and their new, induced testimony. Investigators should especially refrain from this practice in overly long interrogations (Kassin et al., 2010). As demonstrated, participants are especially prone to deceptions if time introduces ambiguity about what has been said. This issue becomes even more relevant when taking into account that persons with a criminal record are more likely to become suspects, given that their photos more frequently appear in mug shots. The current results have demonstrated that the presence of a non-zero level is among the factors that increase people's proneness to blindness. Likewise, our findings reemphasize the need for camera recording during the interrogation, making it possible to discern original testimony from investigator-induced statements (Kassin et al., 2010). We would therefore like to reinforce the call for abolishing the use of willful manipulations in police interrogations to reduce the occurrence of false confessions (Horselenberg et al., 2003; Kassin, 2008; Kassin & Kiechel, 1996).

ACKNOWLEDGEMENTS

The authors would like to thank Judith Tahalele for her help in collecting data.

REFERENCES

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*, 349–354. DOI: 10.1037/h0047358

Derogatis, L. R., Lipman, R. S., & Covi, L. (1973). SCL-90: An outpatient psychiatric rating scale – preliminary report. *Psychopharmacology Bulletin*, *9*, 13–28.

Frazier v. Cupp, 394 U.S. 731 (1969).

Furnham, A., McManus, C., & Scott, D. (2003). Personality, empathy and attitudes to animal welfare. *Antrozoös*, *16*, 135–146. DOI: 10.2752/089279303786992260

Gudjonsson, G. H. (1984). A new scale of interrogative suggestibility. *Personality and Individual Differences*, *5*, 303–314. DOI: 10.1016/0191-8869(84)90069-2

Gudjonsson, G. H. (1989). Compliance in an interrogative situation: A new scale. *Personality and Individual Differences*, *10*, 535–540. DOI: 10.1016/0191-8869(89)90035-4

Gudjonsson, G. H. (1997). The Gudjonsson Suggestibility Scales manual. Hove, UK: Psychology.

Gudjonsson, G. H., & Sigurdsson, J. F. (1999). The Gudjonsson Confession Questionnaire – Revised (GCQ-R). Factor structure and its relationship with personality. *Personality and Individual Differences*, *27*, 953–968. DOI: 10.1016/S0191-8869(98)00278-5

Hall, L., Johansson, P., Tärning, B., Sikström, S., & Deutgen, T. (2010). Magic at the marketplace: Choice blindness for the taste of jam and the smell of tea. *Cognition*, *117*, 54–61. DOI: 10.1016/j.cognition.2010.06.010

Horselenberg, R., Merckelbach, H., & Josephs, S. (2003). Individual differences and false confessions: A conceptual replication of Kassin and Kiechel (1996). *Psychology, Crime and Law*, *9*, 1–8. DOI: 10.1080/10683160308141

Horselenberg, R., Merckelbach, H., Smeets, T., Franssens, D., Peters, G. J. Y., & Zeles, G. (2006). False confessions in the lab: Do plausibility and consequences matter? *Psychology, Crime and Law*, *12*, 61–75. DOI: http://dx.doi.org/10.1080/10683100042000303076

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2001). Criminal interrogation and confessions (4th ed.). Gaithersberg, MD: Aspen.

Jelicic, M., Merckelbach, H., Timmermans, M., & Candel, I. (2004). De Nederlandstalige versie van de Psychopathic Personality Inventory [Dutch version of the Psychopathic Personality Inventory]. *De Psycholoog*, *39*, 604–608.

Johansson, P., Hall, L., & Sikström, S. (2008). From change blindness to choice blindness. *Psychologia: An International Journal of Psychology in the Orient*, *51*, 142–155. DOI: 10.2117/psysoc.2008.142

Johansson, P., Hall, L., Sikström, S., & Olsson, A. (2005). Failure to detect mismatches between intention and outcome in a simple decision task. *Science*, *310*, 116–119. DOI: 10.1126/science.1111709

Johnson, M. B., & Drucker, J. (2009). Two recently confirmed false confessions: Byron A. Halsey and Jeffrey M. Deskovic. *The Journal of Psychiatry and Law*, *37*, 51–72.

Kassin, S. M. (2005). On the psychology of confessions: Does innocence put innocents at risk? *American Psychologist*, *60*, 215–228. DOI: 10.1037/0003-066X.60.3.215

Kassin, S. M. (2008). False confessions. Causes, consequences, and implications for reform. *Current Directions in Psychological Science*, *17*, 249–259. DOI: 10.1111/j.1467-8721.2008.00584.x

Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions, risk factors, and recommendations: Looking ahead. *Law and Human Behavior*, *34*, 49–52. DOI: 10.1007/s10979-009-9188-6

Kassin, S. M., & Gudjonsson, G. H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest*, *5*, 33–67.

Kassin, S. M., & Kiechel, K. L. (1996). The social psychology of false confessions: Compliance, internalization, and confabulation. *Psychological Science*, *7*, 125–128. DOI: 10.1111/j.1467-9280.1996.tb00344.x

Kassin, S. M., Leo, R. A., Meissner, C. A., Richman, K. D., Colwell, L. H., Leach, A. M., & La Fon, D. (2007). Police interviewing and interrogation: A self-report survey of police practices and beliefs. *Law and Human Behavior*, *31*, 381–400. DOI: 10.1007/s10979-006-9073-5

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. DOI: 10.1023/A:1024871622490

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. DOI: 10.1023/A:1024814009769

Magner, E. S., Markham, R., & Barnett, C. (1996). Would reading an account of an event refresh your memory? *Journal of Applied Psychology*, *81*, 769–776. DOI: 10.1037/0021-9010.81.6.769

McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, *52*, 509–516. DOI: 10.1037/0003-066X.52.5.509

Merckelbach, H., Jelicic, M., & Pieters, M. (2011a). Misinformation increases symptom reporting – a test–retest experiment. *Journal of the Royal Society of Medicine*, *2*. DOI: 10.1258/shorts.2011.011062

Merckelbach, H., Jelicic, M., & Pieters, M. (2011b). The residual effect of feigning: How intentional faking may evolve into a less conscious form of symptom reporting. *Journal of Clinical and Experimental Neuropsychology*, *33*, 131–139. DOI: 10.1080/13903395.2010.495055

Merckelbach, H. Muris, P., Wessel, I., & van Koppen, P. (1998). The Gudjonsson Suggestibility Scale (GSS): Further data on its reliability, validity, and metacognition correlates. *Social Behavior and Personality*, *26*, 203–210. DOI: 10.2224/sbp.1998.26.2.203

Moston, S., Stephenson, G. M., & Williamson, T. M. (1992). The effects of case characteristics on suspect behavior during police questioning. *British Journal of Criminology*, *32*, 23–40.

Perillo, J. T., & Kassin, S. M. (2011). Inside interrogation: The lie, the bluff, and false confessions. *Law and Human Behavior*, *35*, 327–337.

Sagana, A., Sauerland, M., & Merckelbach, H. This is the person you selected: about being blind for one's own eyewitness identification decisions. Manuscript submitted for publication.

Sauerland, M., Sagana, A., & Otgaar, H. (2012). Theoretical and legal issues related to choice blindness for voices. *Legal and Criminological Psychology*. DOI: 10.1111/j.2044-8333.2012.02049.x

van Bergen, S., Horselenberg, R., Merckelbach, H., Jelicic, M., & Beckers, R. (2010). Memory distrust and acceptance of misinformation. *Applied Cognitive Psychology*, *24*, 885–896. DOI: 10.1002/acp.1595

van Bergen, S., Jelicic, M., & Merckelbach, H. (2008). Interrogation techniques and memory distrust. *Psychology, Crime and Law*, *14*, 425–434. DOI: 10.1080/10683160701822533

Wagenaar, W. A. (2002). False confessions after repeated interrogation: The Putten Murder case. *European Review*, *10*, 519–537. DOI: 10.1017/S106279870200042X

APPENDIX
Questionnaire about history of norm-violating behaviors

For the following questions, please indicate which of the four answer options is most applicable to your own behavior. Please tick the applicable box.

1. I have deflated the tires of another person’s bicycle.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
2. I have driven through a red traffic light.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
3. I have tortured a bug.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
4. I have driven away from an accident in which I was involved (with or without my fault).
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
5. I have stolen some kitchen utensils from a student cafeteria.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
6. I have cheated on a test in high school.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
7. I have used public transportation (bus, train, subway, tram) without a valid ticket.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
8. I have committed small-scale shoplifting.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
9. I have bought a bike that I knew was stolen.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
10. I have been in trouble with the police.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
11. I have stolen money from the purse of my parents.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
12. I have gotten too much money back in a supermarket and kept the money.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
13. I have parked my car on a parking space reserved for people with disabilities.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
14. I have stolen a bike.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
15. I cheated my insurance in whatever way.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
16. I have left some junk in the park, after I have had a picnic there.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
17. I have used physical violence against others in a conflict.
☐ Never ☐ Seldom ☐ Sometimes ☐ Often
18. I have destroyed public or other people’s property (bike, phone box, bus shelter).
☐ Never ☐ Seldom ☐ Sometimes ☐ Often

DOI: 10.1002/bsl