Legal and Criminological Psychology (2009), 14, 149–155


The British Psychological Society

www.bpsjournals.co.uk

# Shortened versions of the Gudjonsson Suggestibility Scale meet the standards

T. Smeets*, J. Leppink, M. Jelicic and H. Merckelbach

Department of Clinical Psychological Science, Maastricht University, The Netherlands

**Purpose.** The Gudjonsson Suggestibility Scale (GSS; Gudjonsson, 1984, 1997) is a well-established forensic tool for measuring interrogative suggestibility. However, one restriction of this tool is that it requires an extensive testing procedure. The present study examined whether shorter versions of the GSS yield similar results as the original GSS procedure.

**Methods.** One group ($N = 20$) was given a shortened version of the GSS that consisted of an immediate recall test and the specific questions. GSS scores in this group were compared with those in a group ($N = 20$) that had the standard procedure which includes a retention interval and immediate and delayed recall tests. A third group ($N = 20$) was administered a shortened procedure in which the 20 GSS questions immediately followed the GSS story. In the fourth group ($N = 20$), participants were given the retention interval, but no recall tests were administered.

**Results.** ANOVA showed no differences in GSS scores amongst the four groups. *Post hoc* power analyses indicated that these non-significant findings were not the result of a power problem and that larger sample sizes are expected to yield comparable results. Further analyses showed that neither the retention delay nor the recall tests affected suggestibility scores.

**Conclusions.** These results suggest that shortened procedures for administering the GSS may be employed in situations where time is a key factor.

The susceptibility of individuals to give in to leading questions and interrogative pressure may present a major risk factor when they are interrogated by the police. A number of reports have documented that interrogative suggestibility may, in the face of seemingly unfavourable evidence, cause people to falsely confess to criminal offences (e.g. Gudjonsson, 1991, 1993, 1995; Gudjonsson & MacKeith, 1990; Kassin & Gudjonsson, 2004). Apart from individual case reports (for detailed analyses of such cases, see Gudjonsson, 2003) and studies with forensic groups (e.g. Gudjonsson, 1991), there are also psychometric studies (e.g. Sigurdsson & Gudjonsson, 1996), indicating

*Correspondence should be addressed to Tom Smeets, Department of Clinical Psychological Science, Maastricht University, PO Box 616, 6200 MD, Maastricht, The Netherlands (e-mail: tom.smeets@psychology.unimaas.nl).

DOI:10.1348/135532507X190207

that interrogative suggestibility is related to the tendency of some suspects to react to interrogative pressure with developing the false belief and subsequent confession that they actually committed the crime of which they are being accused (i.e. coerced-internalized false confessions).

A widely used and well-researched forensic tool for measuring interrogative suggestibility is the Gudjonsson Suggestibility Scale (GSS; Gudjonsson, 1984, 1997). Basically, the GSS procedure consists of reading a story aloud to participants, who are then asked to free recall as much as they can remember. After a 50-minute delay, they are once again asked to free recall the story. Subsequently, participants are probed with 20 questions pertaining to the story, 15 of which are misleading. When the 20 questions have been answered, the participants are clearly and firmly given a *negative feedback* on their performance. Specifically, they are told that they have made a number of errors and that it is therefore necessary to repeat the questions to obtain more accurate answers (for work on the influence of interviewer demeanour in imparting negative feedback, see Bain & Baxter, 2000). On the basis of participants' answers to the misleading items, a total suggestibility score can be calculated (see below for more details). A wide range of empirical studies have employed the GSS to explore the links between interrogative suggestibility, various memory parameters (e.g. confabulation and distortion), intellectual abilities (see, e.g. Gudjonsson, 1987b, 1991; Gudjonsson & Clare, 1995; Henry & Gudjonsson, 2003; Smith & Gudjonsson, 1995) and personality characteristics (e.g. acquiescence, anxiety, dissociation, locus of control, self-esteem; see Forrest, Wadkins, & Larson, 2006; Gudjonsson & Clare, 1995; Merckelbach, Muris, Rassin, & Horselenberg, 2000; Muris, Meesters, & Merckelbach, 2004). Furthermore, some studies have focused on the GSS's ability to detect patterns of malingering (e.g. Baxter & Bain, 2002; Smith & Gudjonsson, 1986; Woolston, Bain, & Baxter, 2006) and on the role of interrogative suggestibility in laboratory-induced false confessions (e.g. Forrest, Wadkins, & Miller, 2002; Horselenberg, Merckelbach, & Josephs, 2003; Horselenberg *et al.*, 2006; Redlich & Goodman, 2003).

From a practical point of view, one factor that potentially limits the applicability of the GSS is the 1-hour testing session it requires. Given the widespread use of the GSS in forensic as well as laboratory settings and its rather lengthy assessment procedure, it might be interesting (particularly for research purposes) to explore whether more condensed administration forms produce interrogative suggestibility scores that are comparable to those obtained with a traditional GSS procedure. It is worth mentioning that the GSS manual (Gudjonsson, 1997) asserts that it is acceptable to administer the 20 specific GSS questions without the retention delay (i.e. after immediate recall), but the delay considerably increases the difficulty of the task. To the best of our knowledge, however, there is no study that systematically investigates this issue by comparing this condensed version with the standard GSS, though it should be noted that a few studies employed a shortened procedure (e.g. Baxter & Bain, 2002; Woolston *et al.*, 2006). With this in mind, the present study investigated whether shortened procedures, in which a retention interval and recall tests were either included or omitted, would yield similar levels of interrogative suggestibility as the standard GSS procedure that includes a retention interval and immediate and delayed recall tests.

According to the 'discrepancy detection' principle (Tousignant, Hall, & Loftus, 1986; see also Schooler & Loftus, 1986), recollections are most vulnerable to distortion if a person does not straight away detect discrepancies between *post hoc* misinformation and the originally encoded information. Discrepancy detection is said to be influenced by the strength of the encoded information and the way in which

the misinformation is presented. The original GSS procedure contains not only the elements that promote discrepancy detection but also the elements that work against such detection. Multiple recall tests will increase the strength of the memory trace through rehearsal, yet the relatively long retention interval will allow for story details to be forgotten and, thus, provide optimal conditions for post-event misinformation to influence individuals' memory reports. Accordingly, we predicted that the net result of these opposing effects in terms of suggestibility scores would be similar to that of a condensed procedure in which participants are not engaged in multiple recall tests and receive the 20 GSS questions directly after story presentation. Because we could not *a priori* rule out the possibility that a relatively long interval between story and GSS questions is essential for obtaining suggestibility effects, we included a fourth group that had a 50-minute time interval between story presentation and GSS questions, but no multiple free recall tests. Thus, the present study sought to investigate whether the retention interval and the recall tests are essential components of the GSS. By comparing these condensed GSS procedures with the standard procedure, our results are informative for clinicians and researchers who want to obtain suggestibility scores in a situation that has time constraints.

## Methods

### *Participants*

Our sample consisted of 80 young healthy undergraduate students (19 men; 61 women) with a mean age of 21.1 years ($SD = 2.79$). Test protocols were approved by the Standing Ethics Committee of the Psychology Faculty of Maastricht University. All participants gave informed consent and received a small financial compensation for completing the experiment.

### *Materials*

The Gudjonsson Suggestibility Scales (Gudjonsson, 1997) essentially consist of two instruments (i.e. GSS-1 and its parallel form GSS-2) specifically designed to assess individuals' interrogative suggestibility. The GSS has been extensively validated (e.g. Clare, Gudjonsson, Rutter, & Cross, 1994; Gudjonsson, 1984, 1987a; Richardson & Smith, 1993; Tully & Cahill, 1984) and the present study employed a Dutch version of the GSS-1 (see Merckelbach, Muris, Wessel, & van Koppen, 1998). Basically, the GSS is administered as follows. Participants are first read out the GSS-1 story at a fairly slow pace. They then are asked to write down as much as they can remember about the story (i.e. immediate free recall). After a 50-minute retention interval, the participants are again asked to write down whatever they can remember of the story (i.e. delayed free recall). Finally, 20 specific questions (15 of which are misleading) are administered twice, with the experimenter giving the participants explicit negative feedback on their performance in between both administrations. From the data thus obtained, the following scores can be derived:

(1) *Immediate recall*: Reflects the amount of distinct ideas (*range*: 0–40) recalled immediately after the story was read out.
(2) *Delayed recall*: The number of distinct ideas recalled after the retention interval (*range*: 0–40).
(3) *Yield 1*: The extent to which individuals give in to the leading specific questions prior to having been provided with negative feedback (*range*: 0–15).

(4) *Yield 2*: Identical to yield 1, except that it represents the yield score after the negative feedback was given (*range*: 0–15).
(5) *Shift*: The number of distinct changes in responses to the 20 questions after presentation of the negative feedback (*range*: 0–20).
(6) *Total suggestibility*: The sum of yield 1 and shift scores *range*: 0–35).

### *Design and procedure*

All participants were tested individually and gave written informed consent. After arrival in the laboratory, participants were told that various aspects of their memory were going to be tested. Consistent with GSS guidelines, the actual test of interest (i.e. the GSS) was imbedded amongst other (neuropsychological) tests (e.g. Digit Span, see Lezak, 2004). The participants were randomly assigned to one out of the four groups. In the *standard* group ($N = 20$), the GSS was administered in accordance with the procedure outlined in Gudjonsson (1997; cf. supra). The *no delay/immediate recall* group ($N = 20$) consisted of participants who listened to the GSS-1 story and subsequently took an immediate recall test, followed by the 20 GSS questions (i.e. without retention interval). Participants in the *no delay/no recall* group ($N = 20$) were invited to answer the 20 GSS questions immediately after the GSS-1 story had been read to them. Thus, this group had neither multiple recall tests nor a retention interval. In the *delay/no recall* group ($N = 20$), the GSS was administered as follows: participants were presented with the GSS-1 story, followed by the standard retention interval. However, no immediate or delayed recall test was given so that the participants were prevented from explicitly recalling and rehearsing the story before they were given the 20 GSS questions. Groups did not differ with respect to the proportion of men vs. women [$\chi^2(3, N = 80) = 0.76$; $p = .86$; Cramer's $V = .097$], yet the no delay/immediate recall group was slightly younger than the other groups [$F(3, 76) = 4.19$; $p < .05$; $\eta_p^2 = .14$]. Mean age and gender distribution for each group is displayed in Table 1.

## Results and discussion

Table 1 shows mean immediate and delayed recall, yield 1, yield 2, shift and total suggestibility scores. GSS scores fell well within the norms for adults in the general population (see Gudjonsson, 1997). To test the group differences, one-way analyses of variance (ANOVA) with group (standard vs. no delay/immediate recall vs. no delay/no recall vs. delay/no recall) as between-subject factor was carried out.[1] Inspection of Table 1 suggests that the delay/no recall and the no delay/no recall groups had higher total suggestibility scores, yet the ANOVA showed that there were no significant group differences for yield 1 [$F(3, 76) < 1.0$; $p = .40$; $\eta_p^2 = .038$], yield 2 [$F(3, 76) < 1.0$; $p = .92$; $\eta_p^2 = .007$], shift [$F(3, 76) < 1.0$; $p = .73$; $\eta_p^2 = .017$] or total suggestibility scores [$F(3, 75) < 1.0$; $p = .52$; $\eta_p^2 = .029$].

To check whether our non-significant results were due to a lack of statistical power, we conducted *post hoc* power analyses using GPower (Faul & Erdfelder, 1992; for a full

[1] *As the no delay/immediate recall group was slightly younger than the other groups, we also conducted ANOVAs with age included as a covariate. These analyses, however, generated highly similar results in that no between-group differences emerged for yield 1, yield 2, shift and total suggestibility scores.*

**Table 1.** Psychometric properties and mean yield 1, yield 2, shift and total suggestibility scores for each of the four groups, and immediate (no delay/immediate recall and standard group) and delayed (standard group only) recall scores. Standard deviations are given in parentheses.

| | Standard ($N = 20$) | No delay/immediate recall ($N = 20$) | No delay/no recall ($N = 20$) | Delay/no recall ($N = 20$) |
|---|---|---|---|---|
| Age (years) | 21.60 (1.60) | 19.35 (1.84) | 21.60 (2.48) | 22.00 (3.96) |
| Number of men/women | 6/14 | 4/16 | 4/16 | 5/15 |
| Immediate recall | 19.90 (4.52) | 22.83 (6.63) | – – – | – – – |
| Delayed recall | 20.00 (5.07) | – – – | – – – | – – – |
| Yield 1 | 4.55 (2.69) | 3.65 (1.93) | 4.80 (2.76) | 4.70 (1.92) |
| Yield 2 | 5.10 (3.29) | 5.20 (3.19) | 5.40 (3.59) | 5.75 (2.22) |
| Shift | 3.20 (1.94) | 3.65 (2.13) | 3.65 (1.79) | 3.90 (2.02) |
| Total suggestibility | 7.75 (3.26) | 7.30 (3.21) | 8.45 (3.32) | 8.60 (2.70) |

description, see Erdfelder, Faul, & Buchner, 1996) with power ($1 - \beta$) set at 0.80 and $\alpha = .05$, two-tailed. This showed us that sample sizes would have to increase up to $N = 296$, 1,668, 660 and 388 for yield 1, yield 2, shift and total scores, respectively, in order for group differences to reach statistical significance at the .05 level. Thus, it is unlikely that our negative findings can be attributed to a limited sample size.

To delineate the net contributions of the retention delay and recall tests, we also conducted ANOVAs with retention delay (delay vs. no delay) and recall tests (included vs. no recall tests) as between-subject factors. ANOVA showed that there were no main effects of retention delay or recall tests, or a significant retention delay × recall tests interaction for yield 1 (all $F$ values $< 1.52$; all $p$ values $> .22$), yield 2 (all $F$ values $< 1.0$; all $p$ values $> .54$), shift (all $F$ values $< 1.0$; all $p$ values $> .43$) or total suggestibility (all $F$ values $< 2.04$; all $p$ values $> .16$).

Our findings indicate that at least in a sample of undergraduates, GSS scores depend neither on whether multiple free recall tests are administered nor on whether the standard 50-minute retention interval is applied. Thus, the present study suggests that GSS scores are relatively unaffected by condensed forms of administration. Our finding that GSS scores obtained with shortened procedures do not differ from those obtained by the standard GSS procedure is relevant to large-scale laboratory investigations of the relationship between interrogative suggestibility and other constructs of interest (e.g. Gudjonsson & Clare, 1995; Merckelbach *et al.*, 2000).

On the basis of the discrepancy detection principle, one would have predicted that the delay/no recall group would have been the most vulnerable group. In this group, there was a relatively long retention period together with the absence of multiple free recall tests. However, when compared with standard conditions, suggestibility scores in this group were not significantly raised, indicating that at least in healthy undergraduate samples, GSS scores are relatively insensitive to variations in retention intervals and free recall tests. An obvious limitation of our study is that it relied on a homogeneous sample of undergraduate students. It might well be the case that in such samples, suggestibility scores are relatively resistant against all kinds of manipulations (multiple testing, retention intervals and so on). Therefore, future studies should focus on whether shortened GSS procedures can also be used in forensic settings where forensic experts sometimes do not possess ample time to fully examine a suspect's mental abilities.

## Acknowledgements

This research was supported by the Netherlands Organization for Scientific Research (NWO) grant 452-02-006 awarded to Dr Marko Jelicic.

## References

Bain, S. A., & Baxter, J. S. (2000). Interrogative suggestibility: The role of interviewer behaviour. *Legal and Criminological Psychology*, *5*, 123–133.

Baxter, J. S., & Bain, S. A. (2002). Faking interrogative suggestibility: The truth machine. *Legal and Criminological Psychology*, *7*, 219–225.

Clare, I. C. H., Gudjonsson, G. H., Rutter, S. C., & Cross, P. (1994). The inter-rater reliability of the Gudjonsson suggestibility scale (Form 2). *British Journal of Clinical Psychology*, *33*, 357–365.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments and Computers*, *28*, 1–11.

Faul, F., & Erdfelder, E. (1992). GPOWER: A priori, post-hoc, and compromise power analyses for MS-DOS [Computer program]. Bonn, FRG: Bonn University, Dep. of Psychology.

Forrest, K. D., Wadkins, T. A., & Larson, B. A. (2006). Suspect personality, police interrogations, and false confessions: Maybe it is not just the situation. *Personality and Individual Differences*, *40*, 621–628.

Forrest, K. D., Wadkins, T. A., & Miller, R. L. (2002). The role of preexisting stress on false confessions. *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. (1987a). A parallel form of the Gudjonsson suggestibility scale. *British Journal of Clinical Psychology*, *26*, 215–221.

Gudjonsson, G. H. (1987b). The relationship between memory and suggestibility. *Social Behaviour*, *2*, 29–33.

Gudjonsson, G. H. (1991). The effects of intelligence and memory on group differences in suggestibility and compliance. *Personality and Individual Differences*, *12*, 503–505.

Gudjonsson, G. H. (1993). Confession evidence, psychological vulnerability and expert testimony. *Journal of Community and Applied Social Psychology*, *3*, 117–129.

Gudjonsson, G. H. (1995). I'll help you boys as much as I can: How eagerness to please can result in a false confession. *Journal of Forensic Psychiatry*, *6*, 333–342.

Gudjonsson, G. H. (1997). *The Gudjonsson suggestibility scales manual*. London, UK: Psychology Press.

Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions: A handbook*. New York: Wiley.

Gudjonsson, G. H., & Clare, I. C. H. (1995). The relationship between confabulation and intellectual ability, memory, interrogative suggestibility and acquiescence. *Personality and Individual Differences*, *19*, 333–338.

Gudjonsson, G. H., & MacKeith, J. A. C. (1990). A proven case of false confession: Psychological aspects of the coerced-compliant type. *Medicine, Science and the Law*, *30*, 329–335.

Henry, L. A., & Gudjonsson, G. H. (2003). Eyewitness memory, suggestibility, and repeated recall sessions in children with mild and moderate intellectual disabilities. *Law and Human Behavior*, *27*, 481–505.

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.

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.

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.

Lezak, M. D. (2004). Digit span. In M. D. Lezak, D. B. Howieson & D. W. Loring (Eds.). *Neuropsychological assessment* (4th ed., pp. 751–775). London: Oxford University Press.

Merckelbach, H., Muris, P., Rassin, E., & Horselenberg, R. (2000). Dissociative experiences and interrogative suggestibility in college students. *Personality and Individual Differences*, *29*, 1133–1140.

Merckelbach, H., Muris, P., Wessel, I., & van Koppen, P. J. (1998). The Gudjonsson suggestibility scale (GSS): Further data on its reliability, validity, and metacognition correlates. *Social Behavior and Personality*, *26*, 203–210.

Muris, P., Meesters, C., & Merckelbach, H. (2004). Correlates of the Gudjonsson suggestibility scale in delinquent adolescents. *Psychological Reports*, *94*, 264–266.

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.

Richardson, G., & Smith, P. (1993). The inter-rater reliability of the Gudjonsson suggestibility scale. *Personality and Individual Differences*, *14*, 251–253.

Schooler, J. W., & Loftus, E. F. (1986). Individual differences and experimentation: Complementary approaches to interrogative suggestibility. *Social Behaviour*, *1*, 105–112.

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.

Smith, K., & Gudjonsson, G. H. (1986). Investigation of the responses of fakers and non-fakers on the Gudjonsson suggestibility scale GSS. *Medicine, Science and the Law*, *26*, 66–71.

Smith, P., & Gudjonsson, G. H. (1995). Confabulation among forensic inpatients and its relationship with memory, suggestibility, compliance, anxiety, and self-esteem. *Personality and Individual Differences*, *19*, 517–523.

Tousignant, J. P., Hall, D., & Loftus, E. F. (1986). Discrepancy detection and vulnerability to misleading postevent information. *Memory and Cognition*, *14*, 329–338.

Tully, B., & Cahill, D. (1984). *Police interviewing of mentally handicapped persons: An experimental study*. London, UK: The Police Foundation of Great Britain.

Woolston, R., Bain, S. A., & Baxter, J. S. (2006). Patterns of malingering and compliance in measures of interrogative suggestibility. *Personality and Individual Differences*, *40*, 453–461.

*Received 5 September 2006; revised version received 15 February 2007*