Gudjonsson Suggestibility Scales.

**Purpose:** "Developed in order to measure objectively the vulnerability or proneness of people [to suggestive influence and/or] to give erroneous accounts when interviewed," particularly in forensic contents.

**Population:** Ages 6 and over.

**Publication Date:** 1997.

**Scores:** 7: Immediate Recall, Delayed Recall, Yield 1, Yield 2, Shift, Total Suggestibility, Confabulation.

**Administration:** Individual.

**Forms:** 2 parallel forms: GSS1, GSS2.

**Price Data:** Available from publisher.

**Time:** Administration time not reported.

**Author:** Gisli H. Gudjonsson.

**Publisher:** Taylor & Francis [England].

Review of the Gudjonsson Suggestibility Scales by MARC JANOSON, President, Forensic Psychology PC, New York, NY, and BRUCE FRUMKIN, Director, Forensic and Clinical Psychology Associates, South Miami, FL:

**DESCRIPTION.** The Gudjonsson Suggestibility Scales (GSS) is presented as a memory test. A short narrative paragraph containing 40 facts is read to the person being tested with that person being asked to try to remember everything he or she can about the story. He or she is then asked to state everything that can be remembered about the story. Unless the person being tested has very poor recall initially, after a 50-minute delay, he or she is again asked to recall what they can about the story.

After the recall portion(s) of the test, the test taker is asked 20 standardized questions about the story, 15 of which have been specifically designed as subtly leading (i.e., they lead the subject toward an inaccurate reporting of what they believe they remember about the story). The extent to which a test taker yields to the 15 leading questions comprises the Yield 1 score of the GSS. All participants are clearly and firmly told, "You have made a number of errors. It is therefore necessary to go through the questions once more, and this time try to be more accurate" (manual, p. 11). The examiner is then able to assess how much the person "yields" to the 15 questions after being pressured, the scoring of which leads to the Yield 2 score. The extent to which a test taker shifts from the original response, right or wrong, to a different response after pressure, comprises the Shift score. The Yield 1 (0 to 15) and Shift (0 to 20) are combined for a Total Suggestibility score. These scores can be compared to various normative groups on a number of dimensions including age, legal status, and intellectual ability.
The GSS comprises two parallel forms, the GSS 1 and the GSS 2. They are identical in structure except for the narrative paragraph and the questions asked about the paragraph.

DEVELOPMENT. Gisli Gudjonsson developed the GSS in order to measure, subtly yet objectively, the construct of interrogative suggestibility. Interrogative suggestibility is the extent to which an individual comes to accept messages or information communicated during formal questioning, essentially coming to believe the information presented as true. As measured by the Yield and Shift scores, information is obtained about the degree to which an individual yields to leading or misleading questions and gives in to negative feedback or pressure.

In addition to use for research, the GSS was developed for clinical use, such as assessing the psychological vulnerability of a defendant or witness to yielding to leading questions and to shifting from one response, right or wrong, to a different response, under pressure. This use has applications in providing data to the court regarding an individual's susceptibility to providing false information during police questioning, which is highly relevant when the trier of fact is assessing the validity of a confession or witness statements. The measure also has applications when a court is determining the voluntariness of a confession or Miranda rights waiver.

TECHNICAL. The scoring of the Yield and Shift scales is highly nondiscretionary and generally clear cut. Interscorer reliability for the suggestibility scales (Yield 1, Yield 2, Shift, Total GSS Score) ranges from .949 to .992 for the GSS 1 and .989 to .996 for the GSS 2 (Richardson & Smith, 1993).

In light of the nature of the GSS, where individuals are likely to remember some of the narrative paragraph over time, test-retest reliability scores have not been obtained for the individual scales. Instead, temporal consistency scores have been obtained, comparing the GSS 1 with the GSS 2 for a variety of populations over different time frames. All the correlations for suggestibility were highly significant. Using a forensic population retested the same day, the correlation was .92 for the Total GSS score and ranged from .80 to .90 on the individual scales (Yield 1, Yield 2, and Shift). Another forensic group retested from one day to 18 months later had a correlation for Total GSS score of .83 and .74 to .78 on the individual scales.

Grisso (1986) reviewed the early validation studies on the GSS 1 and concluded "Construct validation research with the GSS has placed the forensic examiner in a good position to use the GSS scores when considering questions of an examinee's decreased resistance to suggestion or subtle pressure in interrogations by law enforcement officials" (p. 147). Since that initial review, additional research has been done. Suggestibility has been shown to correlate with a number of cognitive variables. Gudjonsson (2003, p. 360-412) summarized the relevant research. There is a negative relationship of suggestibility scores to intelligence and memory. Poor narrowness, evaluative anxiety, state anxiety, and avoidance coping strategies correlated with suggestibility. Research has also shown that although adolescents do not "yield" to leading questions any more than adults, they are more likely to have higher Shift scores (i.e., change a response when provided with pressure or negative feedback). Sleep deprivation is also correlated with suggestibility. Mental illness per se does not correlate with suggestibility. Significantly, research has shown that false confessors have higher GSS scores than forensic patients and those who have maintained their innocence.

COMMENTARY. The GSS was developed using normative data from Great Britain and Iceland. Yet this test is quite appropriate for use with populations from the United States. The reader should be reminded that London, like many big cities in the United States, is a multicultural city.
There is perhaps no reason to believe that those detainees residing in London would score as a group much differently than comparable populations in any multiculturally diverse American cities although American norms would be quite useful. Moreover, there were few differences in performance between those residing in Iceland and those in Great Britain. Relatively little research has been performed on cross-cultural factors and the GSS. Although Gudjonsson, Rutter, and Clare (1995) found that Afro-Caribbean police detainees scored significantly higher on all GSS 2 scores compared to their Caucasian counterparts, such factors as intelligence, memory, and anxiety produce more of an effect in suggestibility scores than ethnicity. Even without data normed on an American population, the GSS provides excellent behavioral data relating to how an individual responds when given leading questions and pressured with negative feedback. Certainly norms from American subjects would enhance the perceived applicability of this test in the United States.

The GSS can be used in a variety of forensic, clinical contexts. Although it might be argued that the test is only relevant in situations in which a defendant has potentially produced a coerced-internalized false confession (has faulty memory for events surrounding an offense but is led to believe by police through leading questions and/or pressure that in fact he or she committed the crime), the GSS has far more applicability when episodic or autobiographical memory is an issue during police questioning. The GSS measures behavioral responses to leading questions and negative feedback, the same processes that occur in many interrogations. Although the GSS does not provide a direct measure of compliance (which does not require personal acceptance of the information provided or request made), research has shown a correlation between GSS scores and that construct as measured by the Gudjonsson Compliance Scale (GCS; Gudjonsson, 2003). The correlations for Yield 1, Shift, and Total GSS score were .40, .53, and .54, respectively.

When testing is performed in a forensic context, the clinician must address issues of response distortion or malingering. The GSS is particularly resistant to exaggeration or feigning of interrogative suggestibility. First, test takers believe they are being given a memory test. Also, a study by Baxter and Bain (2002) demonstrated that even when individuals were informed that the test measures suggestibility and were told to feign suggestibility on the test, only the Yield 1 score was susceptible to faking.

SUMMARY. Gudjonsson has successfully produced objective tests (GSS 1 and GSS 2) to help assess interrogative suggestibility and related constructs in the context of police questioning of suspects and witnesses. The GSS enables comparison of a person's suggestibility to normative groups and provides behavioral samples relevant to those behaviors a defendant may have exhibited when confronted with leading questioning or negative feedback during a Miranda waiver or confession (see Frumkin, in press). The GSS should not be used to assess whether a Miranda waiver or confession was voluntary or whether a confession was false.

The reviewers note that the GSS has limitations. Its normative data are based upon populations in Great Britain and Iceland. Its simplicity invites misuse whereby clinicians put undue weight on individual scores without viewing the data as one piece of what needs to be a comprehensive assessment to address issues pertinent to Miranda waiver and confessions. It is also not meant to provide data suggesting whether or not a confession is true of false. Someone may have high GSS scores, be highly suggestive, and be susceptible to influence, yet still have committed the offense for which he or she confessed. Nevertheless, psychologists now have a unique, objective, standardized test to help in their assessment in confession-related forensic cases.

REVIEWERS' REFERENCES


Review of the Gudjonsson Suggestibility Scales Report by ROMEO VITELLI, Private Practice, Hamilton, Ontario, Canada:

DESCRIPTION. The Gudjonsson Suggestibility Scale (GSS1) and its parallel form (GSS2) were designed to provide professionals with a concise and self-contained test of interrogative suggestibility and verbal recall for use in research, forensic, and clinical applications. Although originally developed for use in forensic contexts, the GSS scales may be valuable tools in any interview setting. Interrogative suggestibility is defined in the GSS manual with a quote from an earlier article as: "the extent to which, within a closed social interaction, people come to accept messages communicated during formal questioning" (Gudjonsson & Clark, 1986, p. 84). Research using early measures of the GSS (Gudjonsson, 1984) has identified two critical components of suggestibility: Yield (tendency to give in to leading questions) and Shift (tendency to shift responses due to interpersonal pressure), and the GSS1 and GSS2 have been designed to measure both components. In addition to interrogative suggestibility measures, the GSS1 and GSS2 provide objective measures of response distortions and fabrications (the two components associated with confabulation), as well as measures of immediate and delayed verbal recall. To complement the information provided by the GSS1 and GSS2, the Gudjonsson Compliance Scale (GCS), a 20-item, self-report measure of compliance using a true/false format is also provided.

The GSS1 and GSS2 are designed to be presented to the test taker as a test of verbal recall. In each version of the GSS, a brief story is presented. It may be either read by the examiner or presented using the audiotape provided with the test package. Immediate and delayed (usually after 50 minutes) recall is measured verbatim, and 20 questions relating to the story are given. Following negative feedback, the 20 questions are re-administered to score for Yield and Shift measures. The test manual provides comprehensive instructions for scoring suggestibility, recall, and confabulation measures. The GCS can be administered either using the printed test sheet or read to the subject if literacy problems prevent standard administration. Scoring the GCS is done using the provided scoring key. The test author gives specific cautions concerning the use of the
GCS for individuals whose IQ falls below 70 or who might have difficulty understanding the test items.

For the purpose of interpretation, normative data are provided for adult, adolescent, and forensic populations. GSS norms are also provided for use with persons suffering from intellectual disabilities. The test author specifically notes that the GSS1, GSS2, and GCS represent only one component in a comprehensive forensic assessment and cannot be used exclusively to make conclusions about witness validity.

DEVELOPMENT. Although the formats of the GSS1 and GSS2 are identical with respect to administration and scoring criteria, the nature of the narrative paragraphs and corresponding interrogative questions differ in specific content with the GSS1 story having a forensic context (a robbery) whereas the GSS2 story has a nonforensic content (a couple saving a boy from an accident). The GSS2 was developed to provide a nonforensic narrative to complement the forensic objectives of the GSS1.

The GCS was originally developed using 28 true/false items selected to measure compliance in an interrogation setting. The original 28 GCS items were administered to a sample of 164 subjects and factor-analyzed to identify 20 items with a high loading on the compliance factor (the manual presents the loadings for the final 20 GCS items). The GCS is presented in two forms: Form D as the standard administration and Form E for use as part of the interview process.

TECHNICAL. Information on the normative samples and the standardization process is provided in the GSS user's manual. Normative data were collected using samples from a variety of different clinical populations from various studies over the years. Information on the sample demographics for the normative and clinical samples are given, and statistics are provided for adults, juvenile offenders, forensic populations, and intellectually disabled offenders. Separate norms for male and female respondents are not provided. Percentile rankings are provided for most of the norms. Analyses of the role of age, forensic status, and clinical diagnosis in test responding are also described.

Interscorer reliability analyses for the GSS1 and GSS2 are provided with coefficients falling in the .95-.99 range for measures of verbal recall and suggestibility with slightly lower results for measures of confabulation (.72-.80 range). Due to the format of the scales, test-retest reliability coefficients are not provided for the GSS1 and GSS2 although "temporal consistency" involving correlations of scores for respondents who had completed both the GSS1 and GSS2 are provided. All correlations are highly significant with correlations for Shift being consistently lower for Yield 1 than Yield 2. Coefficient alpha reliabilities appear to be slightly higher for the GSS2 than the GSS1. Test-retest reliability data for the GCS are provided for a sample of hospitalized forensic patients over a 1- to 3-month time interval (.88 between the two time periods). Standard error of measurement scores for each of the GSS subscales are provided, determined by the relationship between the GSS1 and GSS2 scores when both scales are administered. Given the parallel nature of the GSS measures, the length of time between administrations of both tests has been found to impact on the variability of the scores with longer intervals between test sessions being more appropriate in forensic contexts.

The test manual provides a series of validation studies although the test author also refers readers to the author's comprehensive review of validation studies of the GSS1 and GSS2 (Gudjonsson, 1992).
Factor-analytic studies of the Yield and Shift items of both tests have found strong correspondence to the two-factor model on which the tests were based. Changes in scoring procedures have yielded stronger item loadings and have been incorporated into the current test and norms. Factor analysis of the GCS items has yielded a three-factor model reflecting avoidance of confrontation, eagerness to please, and general compliance.

Suggestibility as measured by the GSS has been linked to a number of cognitive variables in the hypothesized direction for adults, children, and adolescents. The effects of avoidant coping strategies, anxiety, and poor assertiveness have also been found to correlate significantly with Yield and Shift suggestibility scores. Other factors such as sleep deprivation, intellectual disabilities, mental illness, and the role that they play in susceptibility to leading questions are also discussed.

Research assessing the predictive validity of the GSS1 and GSS2 has identified a strong negative correlation between measures of Shift and Yield and interrogative experience. GSS1 suggestibility measures have also been found to relate to level of accuracy in police interviews. Despite the problematic nature of research into "false confessions," research using the GSS1 has identified a linear relationship differentiating "false confessors," "resisters" (respondents denying their involvement in crimes of which they were accused), and forensic patients. False confessors were also found to score more highly on the GCS than other prisoners in a prison sample.

Although confabulation measures were found to be less reliable than suggestibility measures, validation studies have suggested that personality factors may play a greater role than psychiatric diagnosis. Overall, the studies cited and the validation process meets the psychometric standards for test validation as specified in the Standards for Educational and Psychological Testing (AERA, APA, & NCME, 1999).

COMMENTARY. The increasing use of psychological testimony to investigate issues of suggestibility in obtaining confessions has resulted in the Gudjonsson scales being used in research and forensic settings in countries around the world as well as becoming the focus of an impressive body of research into interrogative suggestibility and compliance.

There continues to be a need for the collection of normative data in the U.S. and in other countries to counterbalance the predominantly U.K.-based norms compiled to date. As well, further research is required to validate the confabulation measures and to investigate the role of cultural influences in interrogative suggestibility and the applicability of the GSS1 and GSS2 with pre-adolescent populations. Another line of investigation should focus on the vulnerability of the Gudjonsson scales to the effects of malingering given its increasing usage in criminal cases and the possibility of "coaching" designed to provide the appearance of suggestibility.

Despite these concerns, the Gudjonsson scales are well-designed and well-validated and can be viewed as a valuable addition to the clinical tools that forensic psychologists can draw upon.

SUMMARY. The GSS1, GSS2, and GCS are well-developed instruments based on extensive research into suggestibility, compliance, and confabulation. They are designed to provide professionals with concise measures of interrogative suggestibility, verbal recall, and compliance for use in research, forensic, and clinical applications. The parallel format provides needed flexibility to accommodate client needs and to address forensic and nonforensic applications across clinical populations. Guidelines and scoring examples are provided to enable consistent
interpretation of test results. The test manual outlines the development and validation of the GSS1, GSS2, and GCS and their use to identify potential problems in witness testimony based on numerous research studies that are cited in the bibliography.

REVIEWER'S REFERENCES


