Is there a G-factor in Hypnotic Suggestibility? Confirmatory Factor Analysis of the Harvard Group Scale of Hypnotic Suggestibility

Anoushiravan Zahedi * [1, 2, 3], Werner Sommer [1, 4]

[1] Department of Psychology, Humboldt-Universitat zu Berlin, Rudower Chaussee 18, 12489, Berlin, Germany; [2] German Institute of Human Nutrition, Decision Neuroscience and Nutrition, Arthur-Scheunert-Allee 114-116, 14558, Nuthetal, Germany; [3] Neuroscience Research Center, Charité-Universitätsmedizin Berlin, Germany; [4] Department of Psychology, Zhejiang Normal University, Jin Hua, China

Author Note

* Corresponding Author; Institut für Psychologie, Humboldt-Universität zu Berlin, Rudower Chaussee 18, 12489, Berlin, Germany.

E-mail address: zahedian@hu-berlin.de

Acknowledgment

This work was supported by a scholarship to A.Z. by the Deutscher Akademischer Austauschdienst (DAAD).

Abstract

Individuals differ in their responsiveness to (post-)hypnotic suggestions. However, defining and measuring hypnotizability is contentious because standardized scales, such as the Harvard group scale (HGSHS:A), measure a mixture of general-suggestibility and its increase due to hypnotic induction (hypnotic-suggestibility). Exploratory factor analysis (FA) of standardized scales found them to be heterogeneous; however, the number and nature of latent factors are debated. We applied Confirmatory FA to HGSHS:A scores of 477 volunteers and tested several theory-driven models. Scores were best explained by a bifactor model consisting of a G-factor, tapping into hypnotizability, and three grouping factors, measuring specific suggestibilities, each requiring a unique combination of three top-down cognitive functions: cognitive-simulation, sensory-adaptation, and problem-solving. Structural equation modeling revealed that the simulation-adaptation factor (requiring cognitive-simulation and sensory-adaptation), predicts the other suggestibility factors. These results demonstrate the multifaceted structure of hypnotic-suggestibility and underscore the desideratum for developing a more differentiated scale, focusing on simulation-adaption suggestions.

Keywords: Confirmatory factor analysis, Structural equation modeling, Hypnosis, Suggestibility, Hypnotizability, HGSHS:A.

Is there a G-factor in Hypnotic Suggestibility? Confirmatory Factor Analysis of the Harvard Group Scale of Hypnotic Suggestibility

1. Introduction

Historically, hypnosis has been explained by two main alternative accounts. The state approach defines hypnosis as an altered state of consciousness similar to yoga or meditation (Elkins, Barabasz, Council, & Spiegel, 2015). In contrast, the socio-cognitive account (Kirsch & Lynn, 1998; Lynn & Green, 2011; Lynn, Rhue, & Weekes, 1990) emphasizes cognitive, social, and psychological variables involved in responding to hypnotic suggestions (Spanos, 1971; Spanos, Cobb, & Gorassini, 1985). As discussed by Jensen et al. (2015) and Lynn and Green (2011), contemporary theories of hypnosis only partially align with these traditional alternative views (for review, see Zahedi & Sommer, 2021) and no consensus has been reached. Independent of ongoing disputes, theories of hypnosis agree on the existence of substantial within- and between-subject variability in responding to hypnotic and posthypnotic suggestions (Shor & Orne, 1963). However, defining and measuring hypnotizability is a more contentious issue. Two approaches can be distinguished.

First, based on studies (e.g., Braffman & Kirsch, 1999; e.g., Mazzoni et al., 2009; McGeown et al., 2012; Palfi, Parris, McLatchie, Kekecs, & Dienes, 2020; Parris & Dienes, 2013) that have shown strong correlations between responding to suggestions inside and outside of hypnosis ($r = .67$ for behavioral scores; $r = .82$ for subjective scores; Braffman & Kirsch, 1999), Kirsch (1997) concluded that suggestibility and hypnotizability should be separated. He distinguished (I) *suggestibility* as the capability to respond to suggestions regardless of hypnosis, (II) *hypnotic-suggestibility* as the capability to respond to suggestions under the influence of hypnosis, and finally, (III) *hypnotizability* as the increase in suggestibility due to induction of

hypnosis (i.e., the difference between hypnotic-suggestibility and suggestibility). Unfortunately, as yet, this sophisticated definition of hypnotizability has not been translated into a reliable and valid hypnotizability scale.

Alternatively, one might define hypnotizability as what standardized scales of hypnotizability are measuring. Noteworthy, in this definition, hypnotizability is equated with hypnotic-suggestibility in terms of Kirsch's (1997) description of hypnotizability. Two of the most commonly employed hypnotic-suggestibility scales are the Stanford scale of hypnotic susceptibility (SSHS; Weitzenhoffer & Hilgard, 1962) and the Harvard group scale of hypnotic susceptibility (HGSHS:A; Shor & Orne, 1962; Shor & Orne, 1963). These two scales are similar in nature, except that the HGSHS:A is designed for group administration, whereas the SHSS is designed for individual-participant administration. In both scales, a range of suggestions is presented consecutively, and in the end, either participants themselves (in the HGSHS:A) or administrators (in the SHSS) determine how many of the suggestions had been executed. Hence, participants can receive a score between 0-12 in the HGSHS:A and 0-11 in the SHSS. Based on their hypnotic-suggestibility scores, participants are conventionally categorized as high-, medium-, and low-suggestibles. For instance, in the HGSHS:A, participants with $scores \geq 9$, $8 \geq scores \geq 5$, and $4 \geq scores$ are considered high-, medium-, and low-hypnotic-suggestible, respectively. Both scales are very stable over time; for instance, for the HGSHS:A scores, stability coefficients of $.82\ (15 - year\ retest)\ and\ .71\ (25 - year\ retest)$ have been reported (Piccione, Hilgard, & Zimbardo, 1989).

However, these scales are not flawless. The internal consistency of the HGSHS:A is at best just acceptable (e.g., Bongartz, 1985; Peter et al., 2014; Robin, Kumar, & Pekala, 2005; Varga, Farkas, & Mero, 2012). A related problem is the heterogeneity of standardized scales aiming to

measure hypnotic-suggestibility. For instance, several previous studies scrutinized the structure of HGSHS:A scores (e.g., McConkey, Sheehan, & Law, 1980; Oakman & Woody, 1996; Piesbergen & Peter, 2006; Woody, Barnier, & McConkey, 2005) and SSHS scores (e.g., Woody et al., 2005) by conducting exploratory factor analyses (EFA). Based on these studies, a strong consensus has been reached that the HGSHS:A items do not represent only a single factor. However, there is less consensus about the number and nature of latent factors involved. There is a tentative consensus that the HGSHS:A items are measuring at least three latent factors (Tab. 1). The first latent factor is characterized by ideomotor suggestions, such as "soon after thinking of your head falling forward, you feel a tendency to make the movement." The second factor consists of so-called challenge suggestions, such as "your hands feel heavy… too heavy to be lifted". And the third factor is related to cognitive suggestions such as "you will be increasingly aware of a fly that is going round and round about your head." However, the third factor is very loosely defined, and in EFA, sometimes only one suggestion is loading on this factor (e.g., McConkey et al., 1980).

Why is there only a weak consensus about the number and nature of latent factors in the HGSHS:A? All previous studies investigating the structure of hypnotizability scales have utilized EFA (e.g., McConkey et al., 1980; Oakman & Woody, 1996; Piesbergen & Peter, 2006). Even though EFA is an essential and necessary step, it has limitations. These limitations can be overcome by confirmatory factor analysis (CFA) and structural equation modeling (SEM) (Coulacoglou & Saklofske, 2017). (1) In EFA, no explicit theory-driven hypotheses are formulated and tested. Therefore, the interpretation of the derived factors are post hoc and may vary across studies (Coulacoglou & Saklofske, 2017; Harrington, 2009). In contrast, CFA is theory-driven, and factors are defined a priori. Therefore, we hold that the data-driven nature of EFA is the main reason why results from EFA do not converge on the number and theoretical explanations of latent factors.

Consequently, the present study will apply CFA and SEM for investigating the homogeneity and structure of the HGSHS:A (Coulacoglou & Saklofske, 2017; Harrington, 2009) as a canonical example of hypnotic-suggestibility scales. (2) Even though EFAs have shown that variance in the items of the HGSHS:A must be explained by more than one factor, it is unclear whether a model, which, in addition to specific grouping factors, also assumes a general factor (G-factor) of hypnotizability, will explain the data better than a simple multifactor model. Of particular interest for addressing this question is bifactor modeling (Reise, 2012), which can be implemented only in CFA but not EFA. Bifactor models have addressed long-standing questions, from personality psychology (Musek, 2017) to neuroimaging of individual differences (Cooper, Jackson, Barch, & Braver, 2019) and probably most noticeably in psychometric and intelligence research (Eid, Krumm, Koch, & Schulze, 2018). (3) Finally, SEM allows us to explore causal relationships between latent variables. Thus, SEM of HGSHS:A factors can address whether there is a special grouping factor to which other grouping factors regress. Using SEM, we will test whether responsiveness to a certain group of suggestions (as a latent variable) can predict how participants respond to other suggestions at the level of latent factors. Causality pathway testing is only interpretable when hypotheses are theory-driven and formulated a priori and, therefore, only applicable in conjunction with CFA. Summarizing, several important questions, namely, the number and nature of factors, the existence of a G-factor, and causal pathways between latent factors, can be addressed best by conducting CFA and SEM, which is the general aim of the present paper.

A vital prerequisite to formulating relevant hypotheses for CFA is an appropriate theory. Based on the separable categories of suggestions found by EFAs, Woody et al. (2005) concluded that there are multiple hypnotizabilities, and the differences between different categories of

suggestions cannot be attributed to mere the difficulty of items. Consequently, the most straightforward theoretical interpretation of multiple hypnotizabilities is to assume distinguishable cognitive mechanisms underlying the different categories of suggestions. Hence, for a theory to be considered appropriate, it should assume at least two distinguishable cognitive processes as underlying mechanisms for responding to different suggestions. However, only two theories make such assumptions (for review, see Zahedi & Sommer, 2021). In the following, these theories will be discussed shortly. Notably, here, we are neither presenting an exhaustive review of hypnosis theories (for an in-depth review, see Zahedi & Sommer, 2021) nor evaluating them; instead, based on the results of existing EFAs, we try to account for the HGSHS data by means of CFA based on an appropriate theory.

The first theory is the *unified cognitive theory* (Brown & Oakley, 2004), which is grounded on the concept of contention scheduling (Norman & Shallice, 1986). Norman and Shallice (1986) assume two separable control systems to be involved in action production, that is, the supervisory attentional system (SAS) and contention scheduling (CS). The SAS will interfere when the existing response repertoire is not sufficient for handling a situation or task. In these cases, either a new schema (i.e., a stimulus-response contingency) must be created, or a well-established (prepotent) schema should be inhibited in favor of a less-established schema. In situations that need less cognitive control and can be handled by existing response repertoire, different sets of potential "source schemata" may compete with each other, and the schema that first exceeds a certain activation threshold will be selected by CS (Norman & Shallice, 1986). Brown and Oakley (2004) defined two mutually exclusive styles of responding to hypnotic suggestions, namely, constructive and concentrative. In the concentrative style, the SAS will be disabled or decoupled, and therefore, cannot be used for responding to suggestions. Hence, CS will be the only system in charge of

action control. In contrast, in the constructive style, goal-directed imagination, requiring the SAS, will be used for responding to hypnotic suggestions. These two styles, however, are used to characterize two different groups of participants; for instance, "… subjects using a constructive style may report using a range of strategies to create the suggested experience, whereas concentrative responders are more likely to report that the experience simply occurred" (Brown & Oakley, 2004, p. 172). Hence, even though the unified cognitive theory and its two response styles can be used to understand between-subject differences in responding to suggestions, it does not explain within-subject variance. Therefore, one cannot employ the unified cognitive theory for interpreting the existence of multiple hypnotizabilities within individuals.

The second hypnosis theory is the *simulation-adaptation theory of hypnosis* (SATH; Zahedi & Sommer, 2021) which incorporates three concepts. The first concept is cognitive-simulation, suggesting that imagining a stimulus has the same effects as perceiving that stimulus (Farah, 1988). The main difference between imagining and perceiving is that the former, in contrast to the latter, is caused by inner thoughts rather than external stimuli (Hesslow, 2002). For instance, imagining a stimulus not only activates the same brain areas but also causes the same responses as perceiving the corresponding stimulus (for review, see Hesslow, 2002). The second concept is top-down regulation of sensory input (Frank, 2016; Lopresti-Goodman, Turvey, & Frank, 2013), suggesting if mental representations of stimuli are generated and cognitively manipulated, the perception of such stimuli is subjected to top-down adaptation. In other words, forming mental representations of stimuli, between perceiving and responding to them, allows for top-down downregulation of sensory input. For example, if participants are asked to judge whether they need one or two hands for lifting planks of different sizes while they either grasp them or only look at them, the latter condition but not the former will be subjected to top-down sensory-adaptation

(Lopresti-Goodman et al., 2013). The third concept is predictive coding (Friston, 2010), which suggests that any action, that is, motion or perception, will be initiated by forming predictions about the next state of the motor and/or sensory apparatus (Adams, Shipp, & Friston, 2013; Clark, 2013). Next, predictions are propagated downward through cortico-cortical and corticospinal projections to the relevant muscles and sensory units. Notably, down propagating signals are always *predictions* and not *motor commands* (for the treatment of neuromuscular mechanisms, see Adams et al., 2013). If there will be a difference between the state of the system and the prediction, a prediction error is formed (i.e., being in the surprise state). Any self-organizing system aims to reside in the lowest possible energy state. Therefore, in the surprised-state, such systems attempt to minimize prediction errors (Friston, 2010). During volitional movements, predictions have a higher weight in comparison to prediction errors. Consequently, to leave the surprise-state, prediction errors are used in reflex arcs to correct the movement and align it with predictions. Hence, prediction errors are gradually downregulated in reflex arcs during backpropagation and are diminished sufficiently to be not propagated beyond thalamic nuclei (Adams et al., 2013; Brown, Adams, Parees, Edwards, & Friston, 2013). During perception, on the other hand, prediction errors are given a higher weight compared to predictions. Hence, this time, predictions are updated based on prediction errors for coming out of the surprised-state. In predictive coding theory, these two processes are called active and perceptual inference, respectively. Despite its popularity and success in explaining normal perception and action, predictive coding cannot explain, why during hypnosis, participants can execute actions described by suggestions but attribute the action to external sources, that is, perceive them to be caused externally rather than by their own volition (Lynn et al., 1990). In terms of the predictive coding model, to start a movement, predictions must be given a higher weight than sensory feedback from the external

world (i.e., active inference); however, participants will only perceive their response as externally originated if prediction errors have a higher weight than predictions (i.e., perceptual inference) (Brown et al., 2013; Clark, 2013).

SATH employs three top-down cognitive processes to explain hypnotic and posthypnotic suggestion-induced responses; these processes are used to different extents and in various combinations, depending on the individual's capabilities. (1) *Cognitive-simulation* (Farah, 1988; Hesslow, 2002). During hypnosis, participants have two sources of input, imagination and the external world. In other words, besides perceptual input from the external world, stimuli described by suggestions that are imagined by participants provide "perceptual" input from mental imagery. (2) *Sensory-adaptation* (Frank, 2016; Lopresti-Goodman et al., 2013). When sensory input from external stimuli is not aligned with imaginations and suggestions, sensory-adaption may downregulate perceptual input from the external world. Together with predictive coding, cognitive-simulation and sensory-adaptation can explain why hypnotized participants execute responses and attribute them to external sources. During hypnosis, imaginations are given higher weight in comparison to predictions. Hence, predictions are updated based on perceptual input from imaginations (perceptual inference). Simultaneously, predictions are given a higher weight in comparison to sensory input. Therefore, sensory input from the external world is used in reflex arcs to align movements with predictions (active inference) and simultaneously downregulated by top-down sensory-adaptation preventing it from passing beyond thalamic nuclei. Consequently, hypnotized participants exert a suggested motion and simultaneously attribute it to the suggestions (or external sources) rather than to themselves. (3) *Mental practice and problem-solving* (Zahedi, Sturmer, & Sommer, 2020). There are two situations that cannot be explained with cognitive-simulation and sensory-adaption and, as a result, require a further cognitive top-down process. (I)

When suggestions do not provide imagery-provoking descriptions of stimuli, hypnotized participants need to fill in the gap and find an appropriate imagination (i.e., strategy) for cognitive-simulation. (II) In case that suggestions aim to form a new trigger-response contingency, cognitive-simulation can provide an (imagined) exercise environment, where new strategies, either outlined by suggestions or developed by participants themselves, are mentally practiced.

According to SATH, besides the aforementioned top-down cognitive processes, social and psychological factors may also be of great importance. As all the top-down processes above are volitional and goal-directed, participants' expectations, openness, and willingness will determine whether they will be motivated to engage in the responses described in the suggestions.

1.1. Theory-driven Hypotheses for CFA and SEM

In the present study, we used SATH as a framework to formulate theory-driven hypotheses to be tested by CFA and SEM. Based on SATH, a bifactor model should best explain the variance in HGSHS:A scores. The proposed bifactor model consists of a G-factor and three grouping factors, as will be justified next and illustrated in Figure 1.

The first grouping factor covers simulation suggestions, where sensory information is congruent with the portrayed response (Fig. 1A). For instance, consider the suggestion: "stretch your arm and keep it in the air, after a while your hand starts to feel fatigued and it starts to move downward. It is as if a heavyweight has been put on your arm" (Shor & Orne, 1962). Here, sensory input is aligned with the suggestion; if participants stretch their arms, they will feel fatigued after a while. Therefore, the suggestion predicts a sensation that will indeed occur. Nonetheless, if the suggestion is successful, the fatigue will not be attributed to body-internal processes (Kihlstrom, 2008; Lynn et al., 1990) but to the hypnotic condition. This will only happen if participants imagine what is described by simulation suggestions (i.e., a heavyweight is put on their arm). Hence,

cognitive-simulation is necessary for successful exertion of targeted hypnotic-suggestion-induced responses.

The second group of suggestions can be described as simulation-adaptation suggestions (Fig. 1B). In these suggestions, sensory information is incongruent or conflicting with the suggested information. For instance, consider the hand levitation suggestion: your hand feels lighter "as if there's a large helium balloon under [your] palm, or attached with strings to each [one of your] fingertip and [your] wrist … [your] hand and arm will begin to float up" (Hammond, 1998, pp. 43-44). As explained by SATH, when sensory input is not aligned with imagination, it is downregulated through the top-down mechanism of sensory-adaptation. Together, cognitive-simulation and sensory-adaptation can explain why the motion described in the suggestion is executed by the hypnotized participant, but nevertheless, attributed to an external cause. Let's consider the hand levitation suggestion. Here, predictions are generated based on cognitive-simulation of the suggestions' description (i.e., helium balloons are attached to the fingers). Since the predictions (i.e., the hand and arm will be levitated) will receive a lower weight than the imaginary input, they are updated based on prediction errors formed from the comparison of predictions with cognitive-simulation feedback. This explains why the motion is attributed to an external source or cause (i.e., helium balloons). At the same time, predictions will be given a higher weight compared to sensory feedback, which is downregulated by top-down sensory-adaptation. Consequently, the targeted motion is executed, and sensory feedback will be used for adapting the movement with predictions. Evidently, in simulation-adaptation suggestions, besides cognitive-simulation, sensory-adaptation must be incorporated. This fact distinguishes this category from mere simulation suggestions, where only cognitive-simulation is necessary.

The third grouping factor relates to suggestions that require executive functions, for example, in order to find appropriate imagery for cognitive-simulation, forming a new trigger-response contingency, or adapting an existing response to a novel situation. In the most common cases, an executive function suggestion will only describe a goal but no concrete strategy for accomplishing the goal, and participants are responsible for filling the gap. That is, participants have to find an appropriate cognitive strategy and then implement the necessary mechanisms for executing this strategy. Importantly, after finding a suitable strategy for complying with a suggestion, the suggestion will turn into a simulation or simulation-adaptation suggestion (Fig. 1C). Consider, for example, the posthypnotic suggestion item of HGSHS:A after the termination of hypnosis, "when you hear a tapping noise, you will reach down and touch your ankle" (Shor & Orne, 1962). The suggestion has clearly defined the goal but no strategy for implementing it. If participants merely reach down and touch their ankle while believing that they are doing this because the suggestion instructed or commanded them to do so, the action will be attributed to the exertion of direct volitional effort. However, in an alternative scenario, if participants imagine the sound and repeatedly connect it to itching or burning in their ankle, they will exert the portrayed action after hearing the sound but will not attribute it to any volitional effort of their own. Any suggestion can be an executive function suggestion if it fails to imply an applicable strategy for the suggested action. This notion has been confirmed by Galea, Woody, Szechtman, and Pierrynowski (2010). In their study, first, high-hypnotic-suggestible participants were given a suggestion, aiming to induce rigidity and stiffness in their arms. Afterward, participates were asked to move their arms with the implication that they cannot do it. The authors intentionally did not present a relevant strategy, for instance, believing to be paralyzed. As inferred by electromyography, the participants of Galea et al. (2010) came up with different strategies, such as

simultaneously activating agonist and antagonist muscles (biceps and triceps), only activating the antagonist (triceps) but inactivating the agonist, or not activating any muscle.

Another form of executive function suggestion covers situations where participants are directly asked to mentally practice a strategy until it becomes semi-automatic. This form of suggestions has been repeatedly employed in studies that sought to enhance performance in different cognitive tasks (e.g., Iani, Ricci, Baroni, & Rubichi, 2009; Iani, Ricci, Gherri, & Rubichi, 2006; Raz, Fan, & Posner, 2005; Raz, Kirsch, Pollard, & Nitkin-Kaner, 2006; Zahedi, Abdel Rahman, Sturmer, & Sommer, 2019; Zahedi, Luczak, & Sommer, 2020; Zahedi, Stuermer, Hatami, Rostami, & Sommer, 2017; Zahedi, Sturmer, et al., 2020). However, as no relevant suggestion is included in the HGSHS:A or SSHS, we will not discuss it further.

Based on SATH (c.f., Zahedi & Sommer, 2021), we assume that simple multifactor models are not adequate for explaining the variance in hypnotic-suggestibility scales. Instead, two alternatives to the multifactor model presented in Table 1 can be proposed. First, in contrast to other top-down cognitive processes required for responding to suggestions, SATH assumes that cognitive-simulation is needed to successfully execute all types of suggestions presented in the HGSHS:A. This proposition is corroborated by the meta-analysis of Landry, Lifshitz, and Raz (2017), which showed that imagination is the shared characteristic of many different forms of suggestions. Hence, one might expect, treating the cognitive-simulation factor (i.e., a grouping factor) as a G-factor should improve the multifactor model.

The second model focuses on SATH's hypothesis that there are two sources of variability in hypnotic-suggestibility, echoing the proposition of Kirsch (1997), asserting that hypnotic-suggestibility can be decomposed into general suggestibility (capability of responding to suggestions regardless of hypnosis) and hypnotizability (increase in suggestibility due to hypnotic

induction). Hence, besides the specific correlated grouping factors, capturing the different suggestibilities described above (related to different top-down cognitive processes), one may expect that the addition of a general factor to which all HGSHS:A items contribute will improve the model. Evidently, two G-factors are included in this model. That is, the suggestibility G-factor, measured by the correlated grouping factors, and the G-factor of hypnotizability, capturing the psychosocial variables involved in responding to hypnotic suggestions (i.e., hypnotizability). Notably, the G-factor of hypnotizability is unrelated to cognitive processes involved in responding to suggestions (i.e., the suggestibilities).

Finally, an important question regarding the HGSHS:A is whether the outcome of one category of suggestions predicts the outcomes of other categories. Based on SATH, the simulation-adaptation category encompasses items that require a combination of essential underlying mechanisms (i.e., cognitive-simulation and sensory-adaptation) necessary to comply with suggestions. Consequently, the outcome of simulation-adaptation suggestions might predict success in the simulation category and, to a lesser extent, also in the executive function category.

Together, in the current study, for solving several issues related to EFA, we used CFA and SEM to investigate the structural construct of HGSHS:A scores. Since CFA needs a theory, from which relevant hypotheses can be derived, after scrutinizing possible candidates, we chose the SATH and used it as the basis for our CFA and SEM. Based on this theory, a bifactor model was postulated, consisting of three specific grouping factors and a G-factor measured by all items. Further, the internal construct of grouping factors was investigated to establish a causal pathway between categories of suggestions.

2. Methods

2.1. Participants

A sample of 477 participants ($252\,Women, Mean\ age\ =\ 28.7\,years, SD\ =\ 12.6\,years$) was recruited. Several different methods have been used for finding prospective participants; besides inviting local Psychology students, the study was advertised on eBay Kleinanzeigen (https://www.ebay-kleinanzeigen.de) and local radio stations. The study had been approved by the ethics committee of [masked]. Prior to the experiment, signed consents were obtained. Participation was compensated by free assessment of hypnotic-suggestibility or course credits.

Based on the most conservative estimation, the sample size of a CFA should be $N = 20 * free\ parameters$ (Tanaka, 1987). In our study, the number of free parameters in the basic model was 24, which shows our sample was big enough, $N \approx 20 * 24$. Further, based on $1 - \beta > 0.9$, $H0 < 0.05$, and $H1 > 0.1$ (Cohen, 2016), we calculated the required sample size for Root Mean Square Error of Approximation (RMSEA) (Preacher & Coffman, 2006), which yielded a minimum required sample size of, $N = 120$, which again confirms that the analyses conducted here have sufficient power to test presented hypotheses.

2.2. Measurements and Procedure

The HGSHS:A (Shor & Orne, 1962) has 12 suggestions (Table 1) and is designed for administration in group sessions. In the present study, we screened groups of 2-15 volunteers per session. At the beginning of the session, a short description of hypnosis and hypnotizability was given by a certified hypnotizer (A.Z.), as advised in the HGSHS:A manual (Shor & Orne, 1962). Afterward, a recorded German version of HGSHS:A (Bongartz, 1985) was administered while participants were required to follow the suggestions. After the presentation of all suggestions,

participants completed a questionnaire regarding their experience, consisting of two sections. (1) An objective section inquired whether the participant had complied with each of the suggestions, and (2) a subjective section asked how strongly they had experienced the effect of each suggestion. In the objective section, compliance or non-compliance was scored as 1 or 0, respectively, for the first 11 suggestions; the 12th suggestion (i.e., hypnotic amnesia) was scored as 0, if less than 4 items could be remembered, and otherwise as 1. Hence, participants could receive a total score of $0 =< hypnotic\ suggestibility\ =< 12$.

Noticeably, in the objective section, for each suggestion, an observable criterion is presented on which participants should base their responses. For instance, "[w]ould you estimate that an onlooker would have observed that your hand lowered at least six inches (before the time you were told to let your hand down deliberately)?" Participants' responses are independently verified by the hypnotist and administrators present at the session.

2.3. Data analyses

All data analyses were conducted in R (R Core Team, 2013); for CFA and SEM, the lavaan package was used (Rosseel, 2012). As the objective scores of HGSHS:A are binary, the diagonally weighted least squares (DWLS) was used to estimate model parameters, and the full weight matrix (WLSMV) was utilized to compute robust standard errors and mean- and variance-adjusted test statistics. Since the study benefitted from a large sample, WLSMV was preferred for ordinal data in comparison to maximum likelihood (ML) or robust maximum likelihood (MLR) (Li, 2016). Distributions were fitted with VGAM (Yee, 2010; Yee, 2015) and fitdistrplus (Delignette-Muller & Dutang, 2015). In order to evaluate whether the variance in HGSHS:A scores can be attributed merely to the variance in the difficulties of items, the Rasch model was applied before conducting

CFA. Rasch models were calculated by eRM (Mair & Hatzinger, 2007) employing conditional maximum likelihood (CML) and tested using Andersen's likelihood-ratio test (Andersen, 1973).

3. Results and Discussion

3.1. Descriptive results

Figure 2 presents the distribution of HGSHS:A scores; since these parameters are ordinal, they cannot be expected to have a normal distribution. However, the methods used for both CFA and SEM analyses and calculation of estimated loadings in CFA and SEM are robust and insensitive to deviations from normality. Regarding the distribution of HGSHS:A scores, three points should be noted. First, the scores show a beta-binomial distribution with estimated $\mu = 0.52$ and $\rho = 0.099$ [1] (the distribution fit indices are presented in Table 4). Second, it could be argued, as in our study, participants were volunteers, the sampling procedure might have been biased toward more hypnotic-suggestible participants. However, the distribution is not biased (skewed) toward low- or high-hypnotic-suggestibles. Third, the bimodal shape of the distribution is of great interest and will be discussed in section 3.4.

For testing the global fit of the Rasch model, scores were split based on median, and Andersen's likelihood-ratio test results showed bad fit ($\chi^2(11) = 70.802, p < 0.001$)[2]. This result is in agreement with the findings of McConkey et al. (1980) and indicates that the heterogeneity in HGSHS:A cannot be attributed to mere item difficulties.

---

[1] More familiar indices of beta binomial distribution can be calculated by $\alpha=\mu(1-\rho)/\rho$ and $\beta=(1-\mu)(1-\rho)/\rho$.

[2] When χ^2 test is not significant the model fits the data.

3.2. Confirmatory Factor Analyses

For testing the hypotheses outlined in section 1.1, we fitted four models to the data (Fig. 3). First, we tested a single-factor model, with the hypothesis that there might be a single G-factor, which can account for all the variance in the data. Model 2 (i.e., the three-factor model) represents the basic multifactor model introduced in Table 1, where each suggestion is related to only one of the three grouping factors.

Models 3 and 4 represent two bifactor models that were proposed in section 1.1. to enhance the basic multifactor model. In conventional bifactor models, it is necessary to assume orthogonality between latent factors. That is, the correlation between latent factors are constrained to zero (Reise, 2012). Otherwise, interpretation may be complicated (Musek, 2017; Reise, 2012), and models may be unidentifiable (i.e., common anomalies of bifactor models) (Eid, Geiser, Koch, & Heene, 2017; Eid et al., 2018). Anomalies in conventional bifactor models may arise if data were not obtained in a two-level sampling procedure (Eid et al., 2017); that is, if not both, participants, as well as items to participants, have been randomly assigned. Such a two-level sampling is not possible for HGSHS:A since items are fixed. Therefore, based on the suggestions of Eid et al. (2017), we computed bifactor- $(S - 1)$ and bifactor- $(S.I - 1)$ models instead of conventional bifactor models. In the bifactor- $(S - 1)$ model (Model 3), one of the grouping factors, the simulation factor, was conceptualized as a G-factor measured by all HGSHS:A items. The reason for choosing the simulation factor as the reference domain was that cognitive-simulation is a top-down process required for responding to different forms of suggestions. Sensory-adaptation and executive functions, in contrast, are necessary only for a special group of suggestions. In the bifactor- $(S.I - 1)$ model (Model 4), in addition to three *correlated grouping factors*, a G-factor measured by all HGSHS:A items was assumed. In this model, one item is

reserved only for the G-factor. This indicator serves to distinguish the G-factor from the grouping factors. Item 3 of HGSHS:A (Hand Lowering) was chosen as the reference indicator since it is a simple suggestion that arguably relates more to psychosocial variables than any top-down cognitive processes.

Table 2 summarizes the fit indices for all four models, graphically presented in Figure 3. Based on the results shown in Table 2, and considering adjusted thresholds, the G-factor model (Model 1) showed a poor fit to the data. The three-factor model fared better and had a lower standardized root-mean-squared residual (RMSEA) compared to Model 1; however, it still showed only a modest fit. In contrast, both the bifactor- $(S-1)$ (# 3) and bifactor- $(S.I\ -1)$ models (# 4) showed exact fits to the data. The bifactor- $(S.I\ -1)$ model has the lowest χ^2, and standardized root mean squared residual (SPMR), indicating a better fit in comparison to other models, which is expected, as more complex models usually fit the data better.

Even though only the bifactor- $(S-\ 1)$ and bifactor- $(SI.-1)$ models closely fit the data, we compared all four models using the likelihood ratio test (Tab. 3). We reasoned that, since we had a large number of participants, χ^2 might over-reject null hypotheses (*H0*: a given model does not fit the data). All of our models can be tested using the likelihood ratio test, as suggested by Reise (2012). The likelihood ratio tests were conducted hierarchically; that is, a model was only compared to the next simpler model. The results of these tests (Table 3) confirm that the G-factor model (Model 1) is least capable of capturing the variance in the data. Hence, the HGSHS:A is definitely not a homogenous scale as it was initially assumed. Second, the bifactor- $(S\ -\ 1)$ model (Model 3) fits the data significantly better compared to the three-factor model, corroborating our theoretically derived hypothesis that cognitive-simulation is the shared top-down cognitive process employed in all forms of suggestions. Finally, the bifactor- $(S.I\ -1)$ model (Model 4) is

significantly better in comparison to the bifactor- ($S - 1$) model, supporting our hypothesis that hypnotizability and suggestibility can be distinguished.

With regard to the bifactor- ($S.I - 1$) model, one should consider that a model with three correlated first-order factors without a G-factor is equivalent to a second-order G-factor model; that is, a model in which, besides three specific uncorrelated factors, a second-order G-factor is assumed (Eid et al., 2017). Hence, in our bifactor- ($S.I - 1$) model, two G-factors, one hidden and one explicit, are assumed, which capture variance from two different sources. In the introduction, we argued that based on a proposal by Kirsch (1997), which is also incorporated in SATH (Zahedi & Sommer, 2021), hypnotic-suggestibility, as measured by HGSHS:A, can be decomposed into suggestibilities (capability of responding to different types of suggestions in general) and hypnotizability (the increase in suggestibilities due to the hypnotic condition). Consequently, in our bifactor- ($S.I - 1$) model, the correlated grouping factors measure suggestibilities, and the G-factor measures hypnotizability. Although suggestibility and its components can be clearly defined by SATH, hypnotizability is less identifiable and can be related to various psychosocial variables, including willingness and openness (Green & Lynn, 2011; Lynn, Laurence, & Kirsch, 2015), the prior expectations about hypnosis (Kirsch & Lynn, 1997; Terhune, Cleeremans, Raz, & Lynn, 2017), expectations induced by the wordings of suggestions (Lynn, Neufeld, & Matyi, 1987; Matthews, Bennett, Bean, & Gallagher, 1985; Spanos, 1971), rapport with the hypnotist (Lynn et al., 2019), and motivation to respond to suggestions (Jones & Spanos, 1982).

Two critical points related to the loadings in the bifactor- ($S - 1$) and ($SI. - 1$) models should be discussed here. Firstly, in the bifactor- ($S - 1$) model, executive function indicators loaded significantly on the G-factor but not on the executive function factor. This result could be

predicted as in section 1.1., we argued that for executive function suggestions, participants need first to find appropriate mental imagery, and then these suggestions turn into typical simulation or simulation-adaptation suggestions. As each of these suggestions poses a unique problem, it seems that cognitive-simulation (the G-factor) is more salient than problem-solving (the executive function factor) in responding to these suggestions. Secondly, in the bifactor- $(SI. - 1)$ model (Model 4), simulation-adaptation indicators loaded strongly on the simulation-adaption factor but very weakly on the G-factor, indicating that participants responded to simulation-adaptation suggestions less than expected considering their motivations and expectations. This result could also be anticipated as simulation-adaption suggestions require two essential top-down cognitive processes, which makes motivations, expectations, and other psychosocial variables less relevant or irrelevant in responding to them.

3.3. Structural Equation Modeling

Next, we tested our hypotheses that the specific grouping factor, which is parceling the suggestions with all the essential top-down processes required for different suggestibilities, can predict the outcome of the other grouping factors. In other words, we expected that the simulation-adaptation factor can predict the simulation and executive function factors, but not vice versa. We tested this hypothesis in the models with three grouping factors (Models 2 and 4; Fig. 4). Both the executive function factor, as well as the simulation factor, regressed significantly to the simulation-adaption factor, $Ps < .05$, but in none of these models, did simulation-adaptation regress significantly to executive function and simulation factors, $Ps > .1$. Further, in the bifactor- $(SI. - 1)$ model, the correlation between the executive-function and simulation factors became insignificant when they regressed to the simulation-adaption factor, indicating that relationships

between these factors can be reduced to regression of the executive function and simulation factors on the simulation-adaption factor.

These results corroborate SATH's claim that simulation-adaptation suggestions encompass the essential components of suggestibility. In other words, simulation-adaptation suggestions need a balanced interaction between vital top-down cognitive processes employed for responding to different forms of suggestions. That is, imagination (i.e., cognitive-simulation) and top-down downregulation of sensory information (i.e., sensory-adaption) are both required for responding to simulation-adaptation suggestions, distinguishing them from other forms of suggestions. This conclusion has a notable implication for using standardized scales of hypnotic susceptibility in clinical and experimental usage, as will be discussed below.

Figure 5 shows the distributions of factors' scores (for details, see Table 1). As shown in Table 4, all three distributions are beta-binomial. However, only the simulation-adaptation factor showed estimated $\alpha < 1$ and $\beta < 1$, indicating a U-shape distribution. By considering the structure of grouping factors, the fit of the proposed models, and the distribution of factors' scores, we can finally discuss the observed bimodal distribution of HGSHS:A scores (see Fig. 1). Importantly, this bimodality was also observed in many other studies (for review, see Balthazard & Woody, 1989). Three possible explanations may be offered for this distribution. (1) The bimodality in HGSHS:A total scores may be the consequence of overdispersion of a beta-binomial distribution. In other words, the bimodality may be caused by smearing of the peak of the distribution and can be considered as noise. (2) If we consider that the simulation-adaptation factor has a U-shaped beta-binomial distribution and that the simulation-adaptation factor predicts both executive function and simulation factors, we may also suggest that the bimodality in HGSHS:A is forced by the simulation-adaptation distribution. Therefore, hypnotic-suggestibility may be

related to a trait with a U-shape distribution in the population, and assuming a normal distribution does not represent this facet. In other words, there is a higher chance that a person would be either fully capable or entirely incapable of responding to simulation-adaption suggestions rather than in between. Based on the discussion regarding the processes involved in simulation-adaptation suggestions, one may conclude that in a given participant, the interaction between cognitive-simulation (i.e., factor S) and sensory-adaptation (i.e., factor SA) either succeeds in making imaginations dominant over sensory inputs, rendering this participant high-hypnotic-suggestible, or otherwise the participant will be low-hypnotic-suggestible. Importantly, this conclusion does not mean that all hypnotic-suggestible participants are the same. That is, a given participant might have a powerful imagination (i.e., cognitive-simulation) and therefore needs less sensory-adaptation (i.e., top-down downregulation of sensory input), and another might be very talented in sensory-adaptation, and hence, is less dependent on imagination (e.g., Terhune, Cardena, & Lindgren, 2011).

Finally, (3) the observed bimodality might be due to the superimposition of two unimodal distributions. Especially, given that the bifactor- $(SI.-1)$ model was the best fit for the data, one may assume that hypnotizability is related to two underlying mechanisms, namely, suggestibility and hypnotizability, both of which have a normal distribution in the general population. However, as they might have positive versus negative skewness, when they are added together under the general concept of hypnotic-susceptibility, they may cause HGSHS:A scores to have a bimodal distribution. It must be noted that these interpretations need to be experimentally investigated before one can draw any firm conclusion.

4. General Discussion

Our results attest to the advantages of CFA and SEM over EFA. First, since CFA is theory-driven, the results can be interpreted with greater confidence and are not subjected to post hoc interpretations. For instance, the bifactor- $(SI.-1)$ model consists of two parts, a G-factor and the basic three-factor model. Regarding the basic three-factor part, there is an exact correspondence between the theory-driven basic three-factor model and the dominant solution of previous EFAs (e.g., McConkey et al., 1980) (see Tab. 1). However, in previous EFA (e.g., McConkey et al., 1980; Oakman & Woody, 1996; Piesbergen & Peter, 2006), grouping factors and their indicators were data-driven and explained post hoc, distinguishing them from theory-driven (a priori) hypotheses needed for CFA. This issue hindered previous EFAs in agreeing on the most relevant model since several statistically-indistinguishable models fitted the data similarly well. In the current study, on the other hand, the latent factors and their indicators were derived from SATH (Zahedi & Sommer, 2021), and therefore, the uncertainty about the number and nature of latent factors were minimized.

Second, even though multiple studies showed that HGSHS:A scores are best explained by more than one latent factor (McConkey et al., 1980; Woody et al., 2005), to the best of our knowledge, no previous study has applied bifactor modeling to data from HGSHS:A or any other standardized hypnotic susceptibility scale. However, bifactor modeling is of great relevance for analyzing hypnotic-suggestibility according to the proposition of Kirsch (1997) and SATH. In these accounts, standardized scales such as the SSHS or HGSHS:A measure hypnotic-suggestibility, consisting of two components, namely, suggestibility and hypnotizability. Our results show that adding a G-factor, measuring the common variance of all items of the HGSHS:A, to the basic multifactor model significantly enhanced model fit. This G-factor possibly measures global hypnotizability. Hence, the current results corroborate the idea that in addition to the specific

grouping factors, tapping into different suggestibilities, a G-factor, measuring hypnotizability is necessary to fully explain the variance of HGSHS:A scores.

Third, previous studies (McConkey et al., 1980; Woody et al., 2005) have demonstrated multiple hypnotic-suggestibilities but did not investigate the relations or pathways between these latent factors. Our SEM results showed that the outcome of the simulation-adaptation factor, which requires a combination of the two critical top-down processes involved in suggestibility, can predict the outcomes of both the simulation and executive function factors but not vice versa.

Our results regarding the causal pathways have notable implications for the future applications of standardized tests in clinical and experimental settings. Usually, hypnotic suggestions in clinical situations benefit from several properties that are not covered by the HGSHS:A. (1) In clinical applications, there is a cohesive story connecting different suggestions. (2) Clinical suggestions present perspicuous and elaborate strategies, and (3) involve a flexible procedure that can be adapted to client's preferred speed. Because of these differences with clinical applications, the HGSHS:A is not an optimal or even useful predictor of suggestibility outside of screening sessions. This fact is reflected in the problems of standardized hypnotic-suggestibility scales in predicting outcomes of hypnotherapy (e.g., Alladin & Alibhai, 2007; Golden, 2012; Schoenberger, 2000), interventions for pain reduction (e.g., Perri, Rossani, & Di Russo, 2019) or interventions targeting modification of food preferences (e.g., Zahedi, Luczak, et al., 2020). A suitable hypnotic-suggestibility scale for clinical purposes might focus on simulation-adaptation suggestions as (a) these suggestions measure the interaction between cognitive-simulation (i.e., imagination) and sensory-adaptation (i.e., top-down driven downregulation of sensory input), which are the essential elements of suggestibility. (b) In contrast to executive function suggestions, sensory-adaptation suggestions are appropriately elaborated and present a clear strategy. Hence,

they do not rely on participants' problem-solving capabilities to compensate for vague suggestions. (c) In contrast to simulation suggestions, sensory-adaptation suggestions require more than only subjective re-coding of a phenomenon. Therefore, these suggestions should be closer to what participants usually experience outside of screening sessions and better for predicting hypnotic-suggestibility in applied settings.

With regard to experimental applications, standardized hypnotic-suggestibility scales might also not be as useful as deemed. Basic research in hypnosis frequently poses two types of questions. (1) How is hypnotic-suggestibility as a trait related to other traits and/or cognitive functions? This type of experimental research, designated intrinsic (Cox & Bryant, 2008; Wagstaff, 1996), asks, for example, whether hypnotic-suggestibility is related to the capability of concentration, fantasy, or daydreaming (for review, see Lynn et al., 2019). (2) Do (posy)hypnotic suggestions affect perception or performance in specific tasks, such as those measuring cognitive functions (for review, see Kihlstrom, 2013; Kihlstrom, 2014)? Such studies, sometimes referred to as instrumental (Cox & Bryant, 2008; Wagstaff, 1996), ask, for instance, whether (post)hypnotic suggestions can subdue pain perception (e.g., Perri et al., 2019) or facilitate inhibition of irrelevant but intrusive information (e.g., Zahedi et al., 2019).

When questions regarding the influence of hypnosis on any dependent variable are addressed, it is a common practice to use standardized hypnotic-suggestibility scales to categorize participants as low- versus high-hypnotic-suggestibles (for review, see Lifshitz, Aubert Bonn, Fischer, Kashem, & Raz, 2013). Then, for curbing the effects of confounding variables, low- and high-hypnotic-suggestibles are compared, effectively using the former group as controls for the latter (Cox & Bryant, 2008; Wagstaff, 1996). Jensen et al. (2017) discussed why the utilization of low-hypnotic-suggestibles as the only constituent of control groups might be misleading. Our

results support the caveats of Jensen et al. (2017) concerning current screening tools. The variance in hypnotic-suggestibility scores appears to originate from two sources, differences in suggestibility and hypnotizability. Differences in suggestibility are caused by individual differences in the employment of several top-down processes and their interaction, including cognitive-simulation, sensory-adaptation, and problem-solving (Zahedi & Sommer, 2021). Further, differences in hypnotizability are likely influenced by multiple psychosocial factors, which are already discussed in section 3.2. Confirmatory Factor Analyses. Consequently, low-hypnotic-suggestible participants are distinguishable from high-hypnotic-suggestibles due to a combination of cognitive and psychosocial factors. Due to this uncertainty, the utilization of low- hypnotic-suggestibles as controls violates the basic assumption of a control condition. That is, participants in the control and experimental conditions must be identical in possible aspect except for one measurable property. When this assumption does not hold, no conclusion can be drawn from observed differences between the control and experimental groups. Noticeably, the HGSHS:A is not a homogenous instrument as a single G-factor model cannot reasonably fit the data, which underscores our concerns that a given HGSHS:A total score is appropriate for predicting participants' responses to an experimental suggestion.

For intrinsic hypnosis studies, employing available standardized suggestibility scales may be misleading, too. For instance, let's consider the inconsistent relationship between hypnotic-suggestibility and cognitive control. Several well-conducted studies using both psychometric and neural measures have recently shown that without hypnosis, high-hypnotic-suggestibles performed better than low-hypnotic-suggestibles in cognitive tasks (Kirenskaya, Storozheva, Solntseva, Novototsky-Vlasov, & Gordeev, 2019; Srzich et al., 2019). In contrast, Khodaverdi-Khani and Laurence (2016) observed inferior digit span performance in high-hypnotic-suggestibles compared

to low-hypnotic-suggestibles, but there was no significant difference in an N-back task, revealing inconclusive findings with regard to WM performance. Furthermore, in a large sample ($N = 180$), Dienes et al. (2009) found no correlation between hypnotic-suggestibility and cognitive capabilities. The present study indicates why the findings of studies investigating the relation between hypnotic-suggestibility and cognitive control are inconclusive (e.g., Dienes et al., 2009; Faerman & Spiegel, 2021; Ludwig et al., 2013; Parris, 2017; Terhune et al., 2017). Studies investigating the relation between hypnotizability and other processes must take into account the bifactorial nature of the scales employed. Hypnotic-suggestibility scales involve two sources of variance, suggestibility and hypnotizability, which should be considered separately. Our concerns about current scales of hypnotic-suggestibility are shared by many other researchers in the field of hypnosis, questioning the applicability and usefulness of existing scales (Acunzo & Terhune, 2021; Jensen et al., 2017).

Several limitations of our study must be discussed. First, non-significant loadings on some factors are a common anomaly in bifactor modeling (Eid et al., 2017). We also observed this issue in a conventional bifactor model of HGSHS:A scores (see supplementary materials). Therefore, based on the suggestions of Eid et al. (2018) and Eid et al. (2017), we used the bifactor- ($S - 1$) and ($SI. - 1$) models instead of the conventional bifactor model. Even though the anomalies were better controlled in these models, some of the indicators still had nonsignificant loadings (see Fig. 3). Even though these anomalies were interpreted post hoc (see section 3.2. Confirmatory Factor Analyses), they were not expected at the time of postulating these models. However, these anomalies do not call into question our interpretations or conclusions regarding our theory-driven hypotheses.

Further, in our study, we only employed the HGSHS:A for modeling hypnotic-suggestibility scores. In future studies, however, it will be beneficial to measure other top-down cognitive functions, such as top-down sensory-adaptation and cognitive-simulation, postulated by SATH to be important in responding to suggestions. However, one should notice that for measuring some of these top-down cognitive functions, such as sensory-adaptation, there is no standardized test available, and therefore, novel instruments have to be developed first.

In conclusion, in the current study, we conducted a CFA of HGSHS:A based on SATH (Zahedi & Sommer, 2021). Going beyond the results of previous EFAs (e.g., McConkey et al., 1980; Woody et al., 2005), our results showed that (1) the HGSHS:A is best explained by a bifactor model, in which a G-factor tapping into hypnotizability is assumed beside three specific grouping factors measuring different components of suggestibility, namely, simulation, simulation-adaptation, and executive function factors. (2) SEM of causal pathways between latent factors revealed that the outcome of simulation-adaptation suggestions, which require a combination of cognitive-simulation and sensory-adaptation, can predict the outcome of other suggestions. this finding corroborates SATH's claim that these two top-down processes are the essential elements of suggestibility. (3) Finally, our results have several important implications for future applications of the existing standardized scales of hypnotic susceptibility. First, in line with the conclusions of Acunzo and Terhune (2021), the current study shows the need for developing a new scale of hypnotic-suggestibility focused on simulation-adaptation suggestions for clinical and experimental applications. Second, in line with the proposition of Jensen et al. (2017), the current results show the importance of using control groups consisting of participants with similar rather than different hypnotic-suggestibility scores to experimental groups, and a desideratum for revisiting

applications of standardized scales in studies investigating the relation between hypnotizability and other mental processes.

References

Acunzo, D., & Terhune, D. (2021). A critical review of standardized measures of hypnotic suggestibility. *International Journal of Clinical and Experimental Hypnosis*.

Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: active inference in the motor system. *Brain Struct Funct, 218*(3), 611-643. doi:10.1007/s00429-012-0475-5

Alladin, A., & Alibhai, A. (2007). Cognitive hypnotherapy for depression: an empirical investigation. *International Journal of Clinical and Experimental Hypnosis, 55*(2), 147-166. doi:10.1080/00207140601177897

Andersen, E. B. (1973). A goodness of fit test for the rasch model. *Psychometrika, 38*(1), 123-140. doi:10.1007/bf02291180

Balthazard, C. G., & Woody, E. Z. (1989). Bimodality, dimensionality, and the notion of hypnotic types. *International Journal of Clinical and Experimental Hypnosis, 37*(1), 70-89. doi:10.1080/00207148908410534

Bongartz, W. (1985). German norms for the Harvard Group Scale of Hypnotic Susceptibility, Form A. *International Journal of Clinical and Experimental Hypnosis, 33*(2), 131-139. doi:10.1080/00207148508406643

Braffman, W., & Kirsch, I. (1999). Imaginative suggestibility and hypnotizability: an empirical analysis. *Journal of Personality and Social Psychology, 77*(3), 578-587. doi:10.1037//0022-3514.77.3.578

Brown, H., Adams, R. A., Parees, I., Edwards, M., & Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cognitive Processing, 14*(4), 411-427. doi:10.1007/s10339-013-0571-3

Brown, R. J., & Oakley, D. A. (2004). An integrative cognitive theory of hypnosis and high hypnotizability. *The highly hypnotizable person: Theoretical, experimental and clinical issues*, 152-186.

Burnham, K. P., & Anderson, D. R. (2016). Multimodel Inference. *Sociological Methods & Research, 33*(2), 261-304. doi:10.1177/0049124104268644

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181-204. doi:10.1017/S0140525X12000477

Cohen, J. (2016). Statistical Power Analysis. *Current Directions in Psychological Science, 1*(3), 98-101. doi:10.1111/1467-8721.ep10768783

Cooper, S. R., Jackson, J. J., Barch, D. M., & Braver, T. S. (2019). Neuroimaging of individual differences: A latent variable modeling perspective. *Neuroscience and Biobehavioral Reviews, 98*, 29-46. doi:10.1016/j.neubiorev.2018.12.022

Coulacoglou, C., & Saklofske, D. H. (2017). Advances in Latent Variable Measurement Modeling. In *Psychometrics and Psychological Assessment* (pp. 67-88).

Cox, R. E., & Bryant, R. A. (2008). Advances in hypnosis research: methods, designs and contributions of intrinsic and instrumental hypnosis. *The Oxford handbook of hypnosis: Theory, research, and practice, 2008*, 311-336.

Delignette-Muller, M. L., & Dutang, C. (2015). fitdistrplus: AnRPackage for Fitting Distributions. *Journal of Statistical Software, 64*(4). doi:10.18637/jss.v064.i04

Dienes, Z., Brown, E., Hutton, S., Kirsch, I., Mazzoni, G., & Wright, D. B. (2009). Hypnotic suggestibility, cognitive inhibition, and dissociation. *Consciousness and Cognition, 18*(4), 837-847. doi:10.1016/j.concog.2009.07.009

Eid, M., Geiser, C., Koch, T., & Heene, M. (2017). Anomalous results in G-factor models: Explanations and alternatives. *Psychological Methods, 22*(3), 541-562. doi:10.1037/met0000083

Eid, M., Krumm, S., Koch, T., & Schulze, J. (2018). Bifactor Models for Predicting Criteria by General and Specific Factors: Problems of Nonidentifiability and Alternative Solutions. *J Intell, 6*(3). doi:10.3390/jintelligence6030042

Elkins, G. R., Barabasz, A. F., Council, J. R., & Spiegel, D. (2015). Advancing research and practice: The revised APA Division 30 definition of hypnosis. *American Journal of Clinical Hypnosis, 57*(4), 378-385. doi:10.1080/00029157.2015.1011465

Faerman, A., & Spiegel, D. (2021). Shared cognitive mechanisms of hypnotizability with executive functioning and information salience. *Scientific Reports, 11*(1), 5704. doi:10.1038/s41598-021-84954-8

Farah, M. J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review, 95*(3), 307-317. doi:10.1037/0033-295x.95.3.307

Frank, T. D. (2016). Perception adapts via top-down regulation to task repetition: A Lotka-Volterra-Haken modeling analysis of experimental data. *Journal of Integrative Neuroscience, 15*(1), 67-79. doi:10.1142/S0219635216500059

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews: Neuroscience, 11*(2), 127-138. doi:10.1038/nrn2787

Galea, V., Woody, E. Z., Szechtman, H., & Pierrynowski, M. R. (2010). Motion in response to the hypnotic suggestion of arm rigidity: a window on underlying mechanisms. *International Journal of Clinical and Experimental Hypnosis, 58*(3), 251-268. doi:10.1080/00207141003760561

Golden, W. L. (2012). Cognitive hypnotherapy for anxiety disorders. *American Journal of Clinical Hypnosis, 54*(4), 263-274. doi:10.1080/00029157.2011.650333

Green, J. P., & Lynn, S. J. (2011). Hypnotic responsiveness: expectancy, attitudes, fantasy proneness, absorption, and gender. *International Journal of Clinical and Experimental Hypnosis, 59*(1), 103-121. doi:10.1080/00207144.2011.522914

Hammond, D. C. (1998). *Hypnotic induction and suggestion*. Chicago, IL: American Society of Clinical Hypnosis.

Harrington, D. (2009). *Confirmatory factor analysis*: Oxford university press.

Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences, 6*(6), 242-247. doi:10.1016/s1364-6613(02)01913-7

Iani, C., Ricci, F., Baroni, G., & Rubichi, S. (2009). Attention control and susceptibility to hypnosis. *Consciousness and Cognition, 18*(4), 856-863. doi:10.1016/j.concog.2009.07.002

Iani, C., Ricci, F., Gherri, E., & Rubichi, S. (2006). Hypnotic suggestion modulates cognitive conflict: the case of the flanker compatibility effect. *Psychological Science, 17*(8), 721-727. doi:10.1111/j.1467-9280.2006.01772.x

Jensen, M. P., Adachi, T., Tome-Pires, C., Lee, J., Osman, Z. J., & Miro, J. (2015). Mechanisms of hypnosis: toward the development of a biopsychosocial model. *International Journal of Clinical and Experimental Hypnosis, 63*(1), 34-75. doi:10.1080/00207144.2014.961875

Jensen, M. P., Jamieson, G. A., Lutz, A., Mazzoni, G., McGeown, W. J., Santarcangelo, E. L., . . . Terhune, D. B. (2017). New directions in hypnosis research: strategies for advancing the

cognitive and clinical neuroscience of hypnosis. *Neurosci Conscious, 3*(1). doi:10.1093/nc/nix004

Jones, B., & Spanos, N. P. (1982). Suggestions for altered auditory sensitivity, the negative subject effect and hypnotic susceptibility: A signal detection analysis. *Journal of Personality and Social Psychology, 43*(3), 637-647. doi:10.1037/0022-3514.43.3.637

Khodaverdi-Khani, M., & Laurence, J.-R. (2016). Working memory and hypnotizability. *Psychology of Consciousness: Theory, Research, and Practice, 3*(1), 80-92. doi:10.1037/cns0000058

Kihlstrom, J. F. (2008). The domain of hypnosis, revisited. *The Oxford handbook of hypnosis: Theory, research and practice*, 21-52.

Kihlstrom, J. F. (2013). Neuro-hypnotism: prospects for hypnosis and neuroscience. *Cortex, 49*(2), 365-374. doi:10.1016/j.cortex.2012.05.016

Kihlstrom, J. F. (2014). Hypnosis and cognition. *Psychology of Consciousness: Theory, Research, and Practice, 1*(2), 139-152. doi:10.1037/cns0000014

Kirenskaya, A. V., Storozheva, Z. I., Solntseva, S. V., Novototsky-Vlasov, V. Y., & Gordeev, M. N. (2019). Auditory Evoked Potentials Evidence for Differences in Information Processing between High and Low Hypnotizable Subjects. *International Journal of Clinical and Experimental Hypnosis, 67*(1), 81-103. doi:10.1080/00207144.2019.1553764

Kirsch, I. (1997). Suggestibility or hypnosis: what do our scales really measure? *International Journal of Clinical and Experimental Hypnosis, 45*(3), 212-225. doi:10.1080/00207149708416124

Kirsch, I., & Lynn, S. J. (1997). Hypnotic involuntariness and the automaticity of everyday life. *American Journal of Clinical Hypnosis, 40*(1), 329-348. doi:10.1080/00029157.1997.10403402

Kirsch, I., & Lynn, S. J. (1998). Dissociation theories of hypnosis. *Psychological Bulletin, 123*(1), 100-115. doi:10.1037/0033-2909.123.1.100

Landry, M., Lifshitz, M., & Raz, A. (2017). Brain correlates of hypnosis: A systematic review and meta-analytic exploration. *Neuroscience and Biobehavioral Reviews, 81*(Pt A), 75-98. doi:10.1016/j.neubiorev.2017.02.020

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods, 48*(3), 936-949. doi:10.3758/s13428-015-0619-7

Lifshitz, M., Aubert Bonn, N., Fischer, A., Kashem, I. F., & Raz, A. (2013). Using suggestion to modulate automatic processes: from Stroop to McGurk and beyond. *Cortex, 49*(2), 463-473. doi:10.1016/j.cortex.2012.08.007

Lopresti-Goodman, S. M., Turvey, M. T., & Frank, T. D. (2013). Negative hysteresis in the behavioral dynamics of the affordance "graspable". *Atten Percept Psychophys, 75*(5), 1075-1091. doi:10.3758/s13414-013-0437-x

Ludwig, V. U., Stelzel, C., Krutiak, H., Prunkl, C. E., Steimke, R., Paschke, L. M., . . . Walter, H. (2013). Impulsivity, self-control, and hypnotic suggestibility. *Consciousness and Cognition, 22*(2), 637-653. doi:10.1016/j.concog.2013.04.001

Lynn, S. J., & Green, J. P. (2011). The sociocognitive and dissociation theories of hypnosis: toward a rapprochement. *International Journal of Clinical and Experimental Hypnosis, 59*(3), 277-293. doi:10.1080/00207144.2011.570652

Lynn, S. J., Green, J. P., Polizzi, C. P., Ellenberg, S., Gautam, A., & Aksen, D. (2019). Hypnosis, hypnotic phenomena, and hypnotic responsiveness: Clinical and Research Foundations-A 40-Year Perspective. *International Journal of Clinical and Experimental Hypnosis, 67*(4), 475-511. doi:10.1080/00207144.2019.1649541

Lynn, S. J., Laurence, J. R., & Kirsch, I. (2015). Hypnosis, suggestion, and suggestibility: an integrative model. *American Journal of Clinical Hypnosis, 57*(3), 314-329. doi:10.1080/00029157.2014.976783

Lynn, S. J., Neufeld, V., & Matyi, C. L. (1987). Inductions versus suggestions: Effects of direct and indirect wording on hypnotic responding and experience. *Journal of Abnormal Psychology, 96*(1), 76-79. doi:10.1037/0021-843x.96.1.76

Lynn, S. J., Rhue, J. W., & Weekes, J. R. (1990). Hypnotic involuntariness: A social cognitive analysis. *Psychological Review, 97*(2), 169-184. doi:10.1037/0033-295x.97.2.169

Mair, P., & Hatzinger, R. (2007). Extended Rasch Modeling: TheeRmPackage for the Application of IRT Models inR. *Journal of Statistical Software, 20*(9). doi:10.18637/jss.v020.i09

Matthews, W. J., Bennett, H., Bean, W., & Gallagher, M. (1985). Indirect versus direct hypnotic suggestions--an initial investigation: a brief communication. *International Journal of Clinical and Experimental Hypnosis, 33*(3), 219-223. doi:10.1080/00207148508406650

Mazzoni, G., Rotriquenz, E., Carvalho, C., Vannucci, M., Roberts, K., & Kirsch, I. (2009). Suggested visual hallucinations in and out of hypnosis. *Consciousness and Cognition, 18*(2), 494-499. doi:10.1016/j.concog.2009.02.002

McConkey, K. M., Sheehan, P. W., & Law, H. G. (1980). Structural analysis of the Harvard Group Scale of Hypnotic Susceptibility, Form A. *International Journal of Clinical and Experimental Hypnosis, 28*(2), 164-175. doi:10.1080/00207148008409838

McGeown, W. J., Venneri, A., Kirsch, I., Nocetti, L., Roberts, K., Foan, L., & Mazzoni, G. (2012). Suggested visual hallucination without hypnosis enhances activity in visual areas of the brain. *Consciousness and Cognition, 21*(1), 100-116. doi:10.1016/j.concog.2011.10.015

Musek, J. (2017). The Overall Strength of the GFP. In *The General Factor of Personality* (pp. 183-202).

Norman, D. A., & Shallice, T. (1986). Attention to Action. In *Consciousness and Self-Regulation* (pp. 1-18).

Oakman, J. M., & Woody, E. Z. (1996). A taxometric analysis of hypnotic susceptibility. *Journal of Personality and Social Psychology, 71*(5), 980-991. doi:10.1037/0022-3514.71.5.980

Palfi, B., Parris, B. A., McLatchie, N., Kekecs, Z., & Dienes, Z. (2020). Can unconscious intentions be more effective than conscious intentions? Test of the role of metacognition in hypnotic response. *Cortex*. doi:10.1016/j.cortex.2020.11.006

Parris, B. A. (2017). The role of frontal executive functions in hypnosis and hypnotic suggestibility. *Psychology of Consciousness: Theory, Research, and Practice, 4*(2), 211-229. doi:10.1037/cns0000106

Parris, B. A., & Dienes, Z. (2013). Hypnotic suggestibility predicts the magnitude of the imaginative word blindness suggestion effect in a non-hypnotic context. *Consciousness and Cognition, 22*(3), 868-874. doi:10.1016/j.concog.2013.05.009

Perri, R. L., Rossani, F., & Di Russo, F. (2019). Neuroelectric evidences of top-down hypnotic modulation associated with somatosensory processing of sensory and limbic regions. *Neuroimage, 202*, 116104. doi:10.1016/j.neuroimage.2019.116104

Peter, B., Vogel, S. E., Prade, T., Geiger, E., Mohl, J. C., & Piesbergen, C. (2014). Hypnotizability, personality style, and attachment: an exploratory study, part 1-general results. *American Journal of Clinical Hypnosis, 57*(1), 13-40. doi:10.1080/00029157.2014.906152

Piccione, C., Hilgard, E. R., & Zimbardo, P. G. (1989). On the degree of stability of measured hypnotizability over a 25-year period. *Journal of Personality and Social Psychology, 56*(2), 289-295.

Piesbergen, C., & Peter, B. (2006). An investigation of the factor structure of the Harvard Group Scale of Hypnotic Susceptibility, Form A (HGSHS:A). *Contemporary Hypnosis, 23*(2), 59-71. doi:10.1002/ch.311

Preacher, K. J., & Coffman, D. L. (2006). Computing power and minimum sample size for RMSEA. In.

R Core Team. (2013). R: A language and environment for statistical computing.

Raz, A., Fan, J., & Posner, M. I. (2005). Hypnotic suggestion reduces conflict in the human brain. *Proceedings of the National Academy of Sciences of the United States of America, 102*(28), 9978-9983. doi:10.1073/pnas.0503064102

Raz, A., Kirsch, I., Pollard, J., & Nitkin-Kaner, Y. (2006). Suggestion reduces the stroop effect. *Psychological Science, 17*(2), 91-95. doi:10.1111/j.1467-9280.2006.01669.x

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate behavioral research, 47*(5), 667-696.

Robin, B. R., Kumar, V. K., & Pekala, R. J. (2005). Direct and indirect scales of hypnotic susceptibility: resistance to therapy and psychometric comparability. *International Journal of Clinical and Experimental Hypnosis, 53*(2), 135-147. doi:10.1080/00207140590927617

Rosseel, Y. (2012). lavaan: AnRPackage for Structural Equation Modeling. *Journal of Statistical Software, 48*(2). doi:10.18637/jss.v048.i02

Schoenberger, N. E. (2000). Research on hypnosis as an adjunct to cognitive-behavioral psychotherapy. *International Journal of Clinical and Experimental Hypnosis, 48*(2), 154-169. doi:10.1080/00207140008410046

Shor, R. E., & Orne, E. C. (1962). *Harvard group scale of hypnotic susceptibility*: Consulting Psychologists Press.

Shor, R. E., & Orne, E. C. (1963). Norms on the Harvard group scale of hypnotic susceptibility, form A. *International Journal of Clinical and Experimental Hypnosis, 11*, 39-47. doi:10.1080/00207146308409226

Spanos, N. P. (1971). Goal-directed fantasy and the performance of hypnotic test suggestions. *Psychiatry, 34*(1), 86-96. doi:10.1080/00332747.1971.11023658

Spanos, N. P., Cobb, P. C., & Gorassini, D. R. (1985). Failing to resist hypnotic test suggestions: a strategy for self-presenting as deeply hypnotized. *Psychiatry, 48*(3), 282-292. doi:10.1080/00332747.1985.11024288

Srzich, A. J., Cirillo, J., Stinear, J. W., Coxon, J. P., McMorland, A. J. C., & Anson, J. G. (2019). Does hypnotic susceptibility influence information processing speed and motor cortical preparatory activity? *Neuropsychologia, 129*, 179-190. doi:10.1016/j.neuropsychologia.2019.03.014

Tanaka, J. S. (1987). "How Big Is Big Enough?": Sample Size and Goodness of Fit in Structural Equation Models with Latent Variables. *Child Development, 58*(1). doi:10.2307/1130296

Terhune, D. B., Cardena, E., & Lindgren, M. (2011). Dissociative tendencies and individual differences in high hypnotic suggestibility. *Cognitive Neuropsychiatry, 16*(2), 113-135. doi:10.1080/13546805.2010.503048

Terhune, D. B., Cleeremans, A., Raz, A., & Lynn, S. J. (2017). Hypnosis and top-down regulation of consciousness. *Neuroscience and Biobehavioral Reviews, 81*(Pt A), 59-74. doi:10.1016/j.neubiorev.2017.02.002

Varga, K., Farkas, L., & Mero, L. (2012). On the objectivity of the scoring of Harvard Group Scale of Hypnotic Susceptibility. *International Journal of Clinical and Experimental Hypnosis, 60*(4), 458-479. doi:10.1080/00207144.2012.675298

Wagstaff, G. F. (1996). Methodological issues in hypnosis. *Psychological research: Innovative methods and strategies*, 202-220.

Weitzenhoffer, A. M., & Hilgard, E. R. (1962). *Stanford hypnotic susceptibility scale, form C* (Vol. 27). Palo Alto, CA: Consulting Psychologists Press.

Woody, E. Z., Barnier, A. J., & McConkey, K. M. (2005). Multiple hypnotizabilities: differentiating the building blocks of hypnotic response. *Psychological Assessment, 17*(2), 200-211. doi:10.1037/1040-3590.17.2.200

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods, 51*(1), 409-428. doi:10.3758/s13428-018-1055-2

Yee, T. W. (2010). TheVGAMPackage for Categorical Data Analysis. *Journal of Statistical Software, 32*(10). doi:10.18637/jss.v032.i10

Yee, T. W. (2015). *Vector generalized linear and additive models: with an implementation in R*: Springer.

Zahedi, A., Abdel Rahman, R., Sturmer, B., & Sommer, W. (2019). Common and specific loci of Stroop effects in vocal and manual tasks, revealed by event-related brain potentials and posthypnotic suggestions. *Journal of Experimental Psychology: General, 148*(9), 1575-1594. doi:10.1037/xge0000574

Zahedi, A., Luczak, A., & Sommer, W. (2020). Modification of food preferences by posthypnotic suggestions: An event-related brain potential study. *Appetite, 151*, 104713. doi:10.1016/j.appet.2020.104713

Zahedi, A., & Sommer, W. (2021). How hypnotic suggestions work – critical review of prominent theories and a novel synthesis. doi:10.31234/osf.io/mp9bs

Zahedi, A., Stuermer, B., Hatami, J., Rostami, R., & Sommer, W. (2017). Eliminating stroop effects with post-hypnotic instructions: Brain mechanisms inferred from EEG. *Neuropsychologia, 96*, 70-77. doi:10.1016/j.neuropsychologia.2017.01.006

Zahedi, A., Sturmer, B., & Sommer, W. (2020). Can posthypnotic suggestions boost updating in working memory? Behavioral and ERP evidence. *Neuropsychologia, 148*, 107632. doi:10.1016/j.neuropsychologia.2020.107632

Figures and Tables

Figure 1

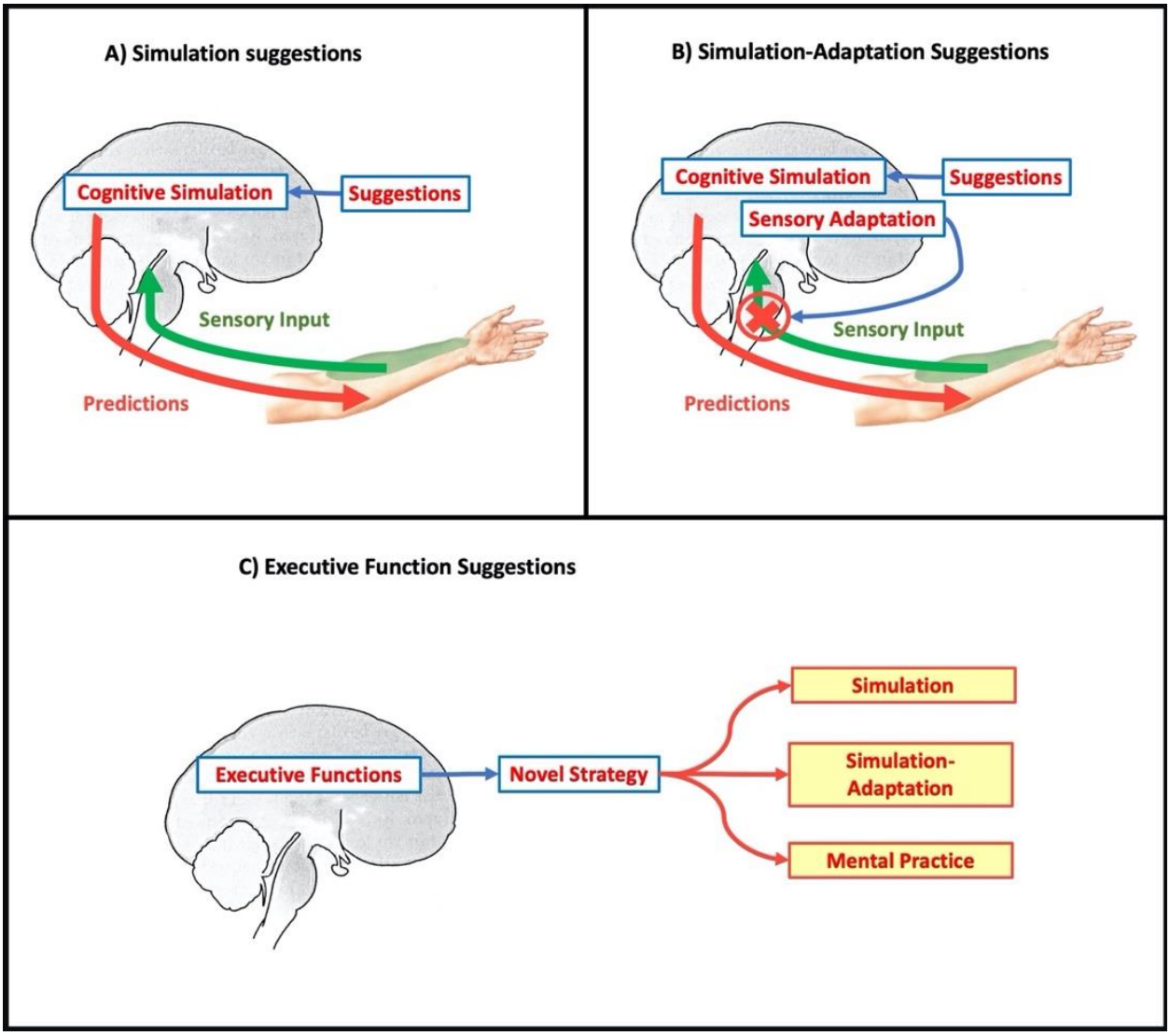

Figure 1. Schematic representation of different kinds of hypnotic and post-hypnotic suggestions and the hypothetical underlying processes; A) simulation, B) simulation-adaptation, and C) executive function suggestions. For details, please see text.

Figure 2

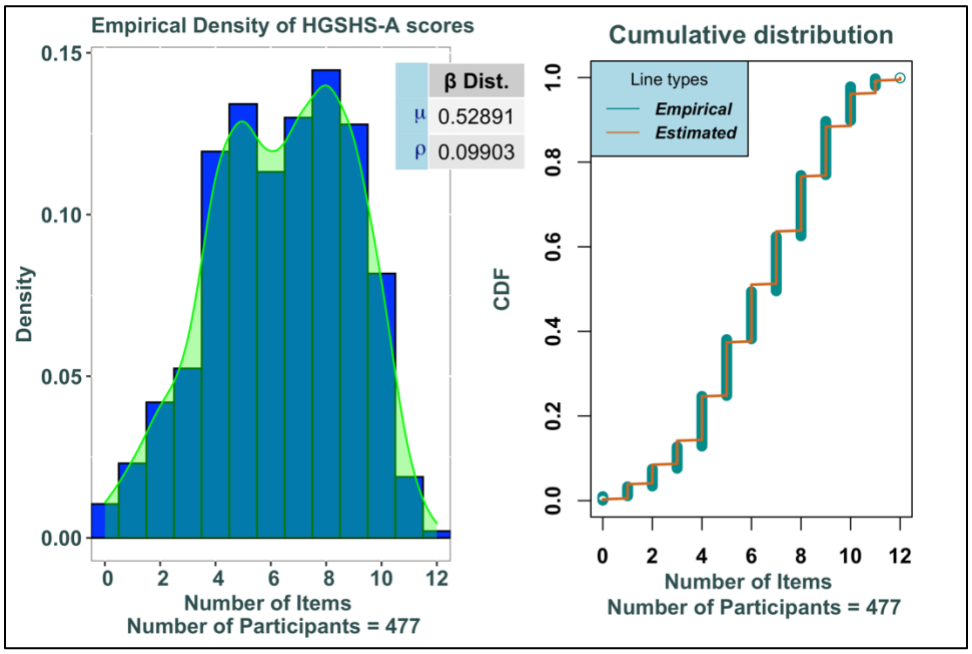

Figure 2. Left: Histogram and density estimation distribution of HGSHS:A scores. Right: Cumulative distribution of HGSHS:A scores.

Figure 3

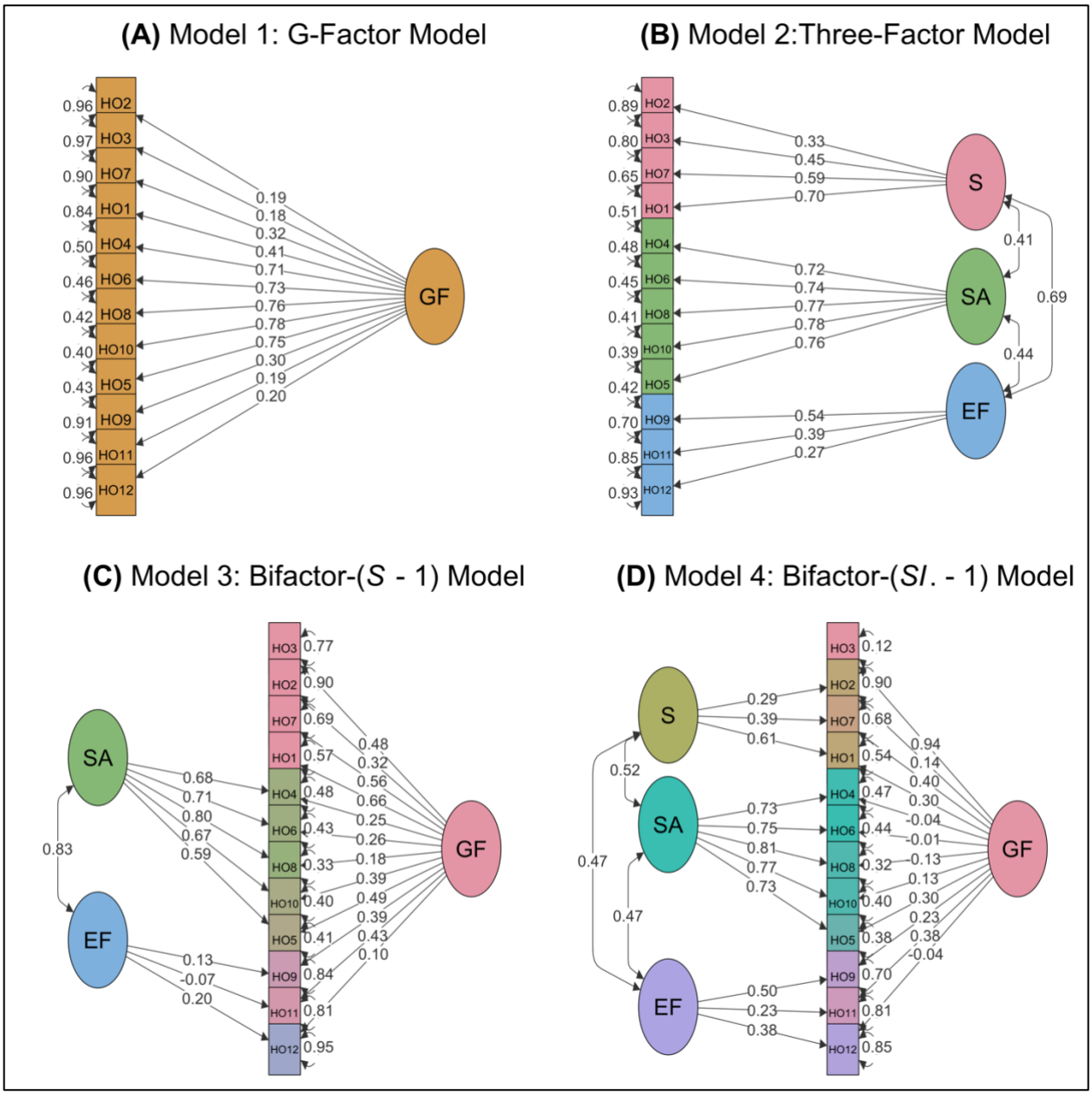

Figure 3. Estimated loadings of the CFA models. (A) G-factor model, (B) three-factor model, (C) bifactor- $(S-1)$ model, and (D) bifactor- $(S.I\ -1)$ model. On the single-headed arrows, standardized factor loadings are given. All loadings in the G-factor and three-factor models are significant, $p < .05$. In the bifactor- $(S\ -\ 1)$ model, the loadings of EF factors and HO12 on GF are not significant; however, other loadings are significant, $p < .05$. In the bifactor- $(SI. - 1)$ model, the loadings of HO2, HO4, HO6, HO8, HO10, and HO12 on the G-factor are not significant; however, other loadings are significant, $p < .05$. Loadings are equivalent to standardized regression coefficients (beta weights), and they are estimated with diagonally weighted least squares. The self-loops show error terms. Squaring these terms gives an estimate of

the variance for each task that is not accounted for by the latent construct. The curved, double-headed arrows indicating correlation coefficients between latent variables, all of which, except the correlation between EF and S in bifactor- $(S - 1)$ model, are significant, $p < .05$.
Note: S: simulation; SA: simulation-adaptation; EF: executive function.

Figure 4

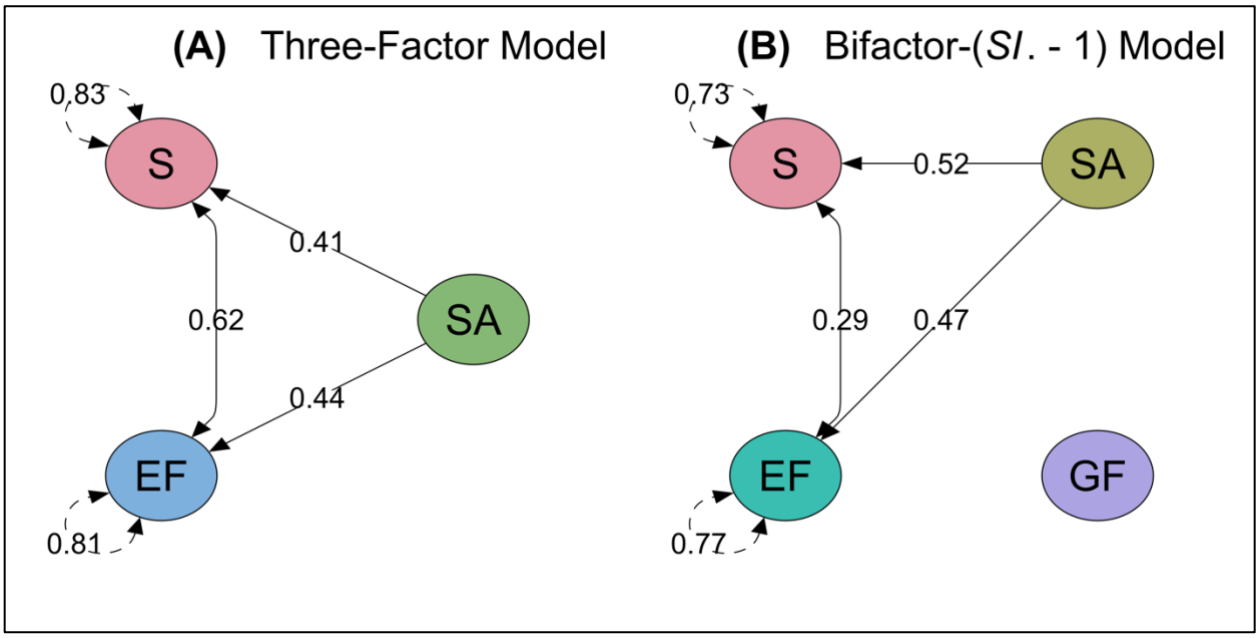

Figure 4. The structural segment of (A) the three-factor, and (B) bifactor- $(SI. - 1)$ models. All correlation and regression coefficients, except the correlation between $EF \sim\sim S$ in the bifactor-$(SI. - 1)$ model, are significant, $p < .05$.

Note: Adding regressions did not change the models, as previously we had assumed correlations between factors. Therefore, we only presented the structural segments of the models.

Figure 5

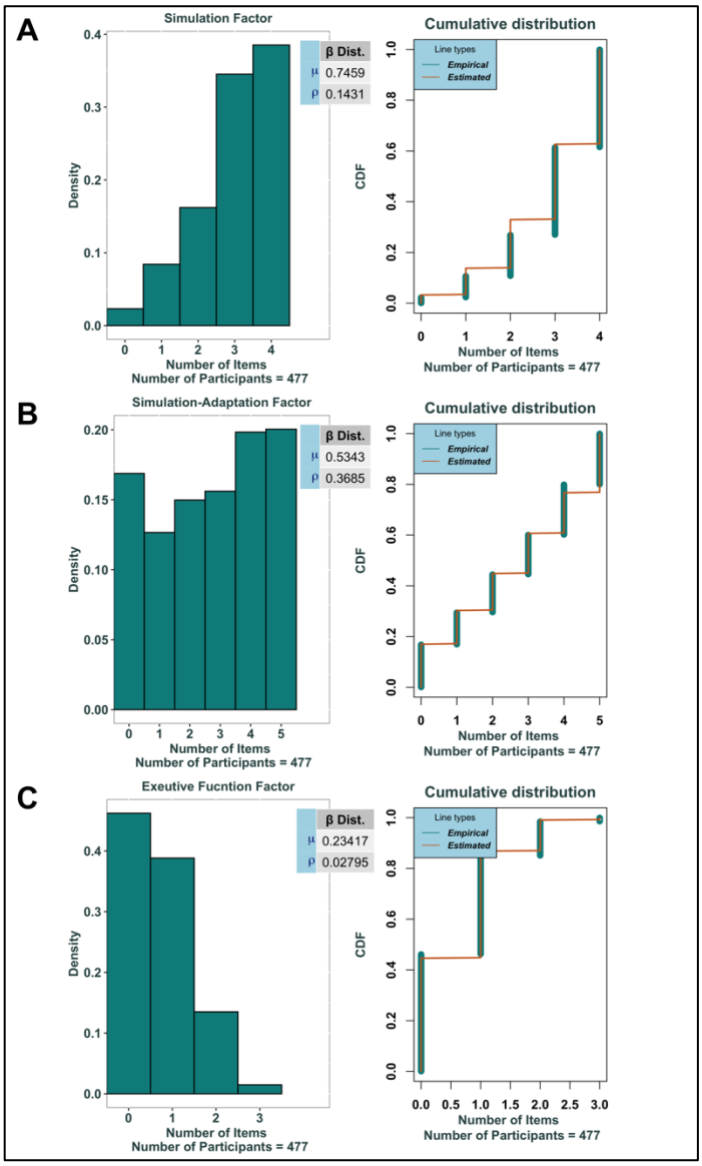

Figure 5. Distributions of the (A) Simulation, (B) Simulation-Adaptation, and (C) Executive Function factor scores. Left: Histograms of distributions. Right: Cumulative distributions of the empirical and estimated data, based on a beta-binomial distribution by the given μ and ρ.

Table 1. Categorization of the suggestions in HGSHS:A into three proposed specific factors and the results of exploratory factor analysis (EFA, McConkey et al., 1980)

Suggestions	Proposed Specific Factor	EFA [a]
1. HEAD FALLING	S	2.Ideomotor
2. EYE CLOSURE	S	2.Ideomotor
3. HAND LOWERING (LEFT HAND)	S	2.Ideomotor
4. ARM IMMOBILIZATION (RIGHT ARM)	SA	1.Challenge
5. FINGER LOCK	SA	1.Challenge
6. ARM RIGIDITY (LEFT)	SA	1.Challenge
7. MOVING HANDS TOGETHER	S	2.Ideomotor
8. COMMUNICATION INHIBITION	SA	1.Challenge
9. EXPERIENCING OF A FLY	EF	3.Cognitive
10. EYE CATALEPSY	SA	1.Challenge
11. POST-HYPNOTIC SUGGESTION	EF	3.Cognitive
12. HYPNOTIC AMNESIA	EF	3.Cognitive

Note: S: simulation; SA: simulation-adaptation; EF: executive function.

Table 2. Fit Indices for the Full Confirmatory Factor Analysis Model and Reduced Models

Model	*df*	χ^2 [a]	RMSEA [90% CI] [b]	SPMR [c]	CFI [d]	TLI [d]
G-factor model	54	142.6 ***	0.059 [0.047-0.079]	0.105	0.934	0.919
Three-factor model	51	80.5 **	0.035 [0.019-0.049]	0.079	0.978	0.972
Bifactor- (S -1) model	45	61.2	0.028 [0.000-0.044]	0.068	0.988	0.982
Bifactor- ($S.I$ -1) model	**40**	**33.7**	**0.000 [0.000-0.024]**	**0.048**	**1**	**1.008**

Note: The endorsed model is indicated in bold. SPMR: standardized root-mean-squared residual; RMSEA: Root Mean Square Error of Approximation; CFI: Bentler's Comparative Fit Index; TLI: Tucker-Lewis Index.

[a] When χ^2 test is not significant the model fits the data. However, as the $N = 477$ was very large, it is expected that H_0 would be over-rejected.

[b] Lower values of RMSEA indicate better fit, with values < .05 indicating a close fit to the data (Xia & Yang, 2019). For $1 - \beta > 0.9$, $H0 < 0.05$, and $H1 > 0.1$ the required sample size is $N > 120$.

[c] Lower values of SRMR indicate better fit, with SRMR < .08 indicating a close fit to the data.

[d] Values > .95 for CFI and TLI indicate a good fit (Xia & Yang, 2019). TLI is not normalized and may have values > 1.

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 3. Results of likelihood ratio tests comparing the four models

Competing models		*df*	$\chi^{2\,a}$
Three-factor model	G-factor model	3	47.9 ***
Bifactor- ($S\ -1$)	Three-factor model	6	18.9 **
Bifactor- ($S.I\ -1$)	Bifactor- ($S\ -1$)	5	22.2 ***

Note: [a] If the χ^2 test is significant the more complex model will be endorsed, if not, the simpler model is endorsed.
** $p < .01$; *** $p < .001$

Table 4. Fit indices for the distributions of the HGSHS:A scores and grouping factors

Distribution	Fitted Distribution	ΔAIC [a]	ΔBIC [a]	Estimated α and β
HGSHS:A scores	**Beta-Binomial**	**0**	**0**	**4.81 - 4.28**
	Poisson	45	42	-
	Negative Binomial	47	48	-
SA scores	**Beta-Binomial**	**0**	**0**	**.99 - .79**
	Poisson	204	200	-
	Negative Binomial	199	199	-
S scores	**Beta-Binomial**	**0**	**0**	**4.46 – 1.52**
	Poisson	328	324	-
	Negative Binomial	326	330	-
EF scores	**Beta-Binomial**	**0**	**0**	**8.14 – 26.63**
	Poisson	11	7	-
	Negative Binomial	13	13	-

Note: The endorsed distribution is highlighted in bold.
S: simulation; SA: simulation-adaptation; EF: executive function.
[a] When $\Delta AIC > 10$, the model with minimum AIC will be endorsed (Burnham & Anderson, 2016). $\Delta AIC = AIC_i - AIC_{minimum}$, and $\Delta BIC = BIC_i - BIC_{minimum}$. The AIC are corrected for the finite sampling (AIC_c).
* $p < .05$; ** $p < .01$; *** $p < .001$