



## Review

# Revisiting the domain of suggestion: A meta-analysis of suggestibility across different contexts

Madeline V. Stein<sup>a,b</sup>, Afik Faerman<sup>c</sup>, Trevor Thompson<sup>d</sup>, Irving Kirsch<sup>e</sup>, Steven J. Lynn<sup>f</sup>, Devin B. Terhune<sup>a,b,\*</sup>

<sup>a</sup> Department of Psychology, Goldsmiths, University of London, United Kingdom

<sup>b</sup> Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, United Kingdom

<sup>c</sup> Department of Psychiatry and Behavioral Sciences, Stanford University, United States of America

<sup>d</sup> Centre for Chronic Illness and Ageing, University of Greenwich, United Kingdom

<sup>e</sup> Program in Placebo Studies, Harvard Medical School, United States of America

<sup>f</sup> Department of Psychology, Binghamton University, United States of America

## ARTICLE INFO

## Keywords:

Forensic  
Hypnosis  
Hypnotizability  
Interrogative  
Suggestibility  
Verbal suggestion

## ABSTRACT

Trait responsiveness to verbal suggestions (suggestibility) is relevant to a diverse array of clinical and experimental psychological phenomena. An unresolved question is whether different forms of suggestibility, such as direct verbal suggestibility and indirect (interrogative and sensory) suggestibility, comprise a uniform, superordinate trait or distinct abilities with overlapping characteristics. We conducted a pre-registered meta-analysis in accordance with MOOSE and PRISMA guidelines to quantify associations between diverse measures of suggestibility. Fifty-five eligible suggestibility scale correlation pairs from 18 studies were subjected to random effects meta-analyses and meta-regression analyses. The analyses revealed a scale congruency effect such that direct-direct ( $k = 37$ ) and indirect-indirect ( $k = 5$ ) suggestibility scale pairs were characterized by strong correlations, whereas incongruent scale pairs (direct-indirect;  $k = 13$ ) exhibited near-zero correlations. These results corroborate proposals that direct and interrogative suggestibility scales measure discrete suggestion effects, and trait responsiveness to suggestion does not comprise a uniform set of abilities.

## 1. Introduction

Verbal suggestions are communications for involuntary changes in experience and behaviour (Kirsch, 1999) and are common in various social interactions, from clinical interventions (Varga, 2011) and forensic interviews (Ridley & Gudjonsson, 2013) to participant-experimenter interactions in experimental research (Barber, 1976; Holman et al., 2015). Members of the general public display marked variability in their responsiveness to verbal suggestions, and numerous lines of evidence point to trait-like stability of response patterns, a cognitive-perceptual trait known as suggestibility (Fassler et al., 2008; Piccione et al., 1989; Rasch & Cordi, 2023).

Elevated suggestibility seems to reliably confer greater responsiveness to a diverse array of interventions and proneness to alterations in different features of awareness and perception. Suggestibility is associated with greater response to suggestion-based interventions such as the use of hypnosis to reduce pain (Millington et al., 2021; Montgomery et al.,

2011; Thompson et al., 2019) and may be relevant to a broader array of clinical interventions, including psychotherapy and psychedelics (Enck & Zipfel, 2019; Kirsch, 1990; Szigei et al., 2024). There is preliminary, albeit mixed, evidence that suggestibility is associated with greater responsiveness to nocebo hyperalgesia (Corsi & Colloca, 2017; Stein et al., 2025) and placebo hypoalgesia (Huber et al., 2013; Lund et al., 2015; Parsons et al., 2021). Elevated suggestibility is also associated with symptom reporting (Stein et al., 2023; Yard et al., 2008) and may function as a risk factor for mass psychogenic illness (Sapkota et al., 2020) and dissociative psychopathology (Bell et al., 2011; Wieder et al., 2021; Wieder et al., 2023). By contrast, in forensic contexts, interrogative suggestibility is a reliable predictor of false confessions (Otgaar et al., 2021), false memories (Rassin, 2022), and distortions in eyewitness testimony (Zaragoza & Lane, 1994). Given its pronounced societal impact, understanding the structure and features of suggestibility is of paramount importance in a variety of clinical, experimental, and legal contexts.

\* Corresponding author at: 16 De Crespigny Park, London SE5 8AB, UK.

E-mail address: [devin.terhune@kcl.ac.uk](mailto:devin.terhune@kcl.ac.uk) (D.B. Terhune).

<https://doi.org/10.1016/j.paid.2025.113181>

Received 22 October 2024; Received in revised form 3 February 2025; Accepted 18 March 2025

Available online 27 March 2025

0191-8869/© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Different forms of suggestibility are typically measured through the use of work-sample instruments comprising a series of verbal suggestions for changes in motor control, memory, or perception followed by self-report or behavioral tests of responsiveness (Acunzo & Terhune, 2021; Gheorghiu et al., 1975; Gudjonsson, 1984; Woody & Barnier, 2008). It has long been theorized that suggestibility is not a uniform trait but rather comprises multiple distinct abilities linked through family resemblance (Polczyk, 2016). A common demarcation line is that between *direct verbal suggestibility* (Oakley et al., 2021), often referred to as *imaginative suggestibility* (Braffman & Kirsch, 1999), and *indirect suggestibility* (Polczyk, 2016) (see Table 1). Direct suggestibility encompasses both hypnotic and non-hypnotic responsiveness to direct verbal suggestions, which are typically overt (Oakley et al., 2021). By contrast, according to this distinction, indirect suggestibility corresponds to responsiveness to suggestive influences that are hidden (Polczyk, 2016); this includes both *interrogative suggestibility*, an individual's tendency to endorse misleading information such as in eyewitness interviews (Gudjonsson, 2013), as well as *sensory suggestibility*, an individual's perceptual responsiveness to misleading information about a sensory stimulus (Polczyk & Pasek, 2006). A somewhat alternative formulation (Lynn et al., 1993; Matthews et al., 1993) proposes that direct suggestions are attributional and predictive but not necessarily overt (e.g., a clear statement about an experience an individual will have) whereas indirect suggestions are more permissive and ambiguous and not necessarily covert (e.g., a more permissive statement about possible experiences an individual might have). Measures of sensory suggestibility (e.g., the *Sensory Suggestibility Scale*; Gheorghiu et al., 1975) include directly phrased suggestions that are covertly administered, similar to a typical placebo procedure. This raises the question of whether sensory suggestibility is better understood as a form of direct or indirect suggestibility.

Preliminary evidence demonstrates that these different measures of suggestibility index mostly discrete abilities (Polczyk, 2016), although there is ongoing debate regarding the structure of suggestibility. One perennial point of contention is whether hypnotic suggestibility reflects a distinct form of suggestibility (Kirsch, 1997; Tasso et al., 2020). Tasso et al. (2020) recently observed that hypnotic suggestibility did not reliably relate to measures of non-hypnotic suggestibility and concluded that it reflects a trait distinct from suggestibility. However, the included non-hypnotic suggestibility scales primarily assessed indirect suggestibility and all had very poor psychometric properties (e.g., Eysenck & Furneaux, 1945). By contrast, hypnotic suggestibility tends to reliably correlate with non-hypnotic direct verbal suggestibility scales with good psychometric properties (Braffman & Kirsch, 1999; Hilgard & Tart, 1966; Wieder & Terhune, 2019). Similarly, it remains unclear whether, and to what extent, interrogative suggestibility relates to sensory suggestibility or direct verbal suggestibility (Gudjonsson, 1987; Polczyk & Pasek, 2006). Interrogative scales primarily measure compliance (Gudjonsson, 2013) and are negatively correlated with intelligence (Gudjonsson, 1988), whereas direct scales tend to only be weakly related to compliance (Polczyk & Pasek, 2006) and are unrelated to

intelligence (Geiger et al., 2014). These apparent dissociations are also present in patient populations: dissociative disorder patients reliably display elevated direct verbal suggestibility (Wieder et al., 2022) but exhibit typical response patterns on interrogative suggestibility measures (Vissia et al., 2016); by contrast, psychosis patients display elevated interrogative suggestibility (Peters et al., 2012) but typical direct verbal suggestibility (Frischholz et al., 1992). These disparate results suggest that these different trait measures reflect discrete forms of suggestibility.

The present pre-registered meta-analysis sought to delineate whether suggestibility is a uniform trait that is stable across different contexts and modes of assessment. In particular, we were motivated to assess the hypothesized distinction between direct and indirect suggestibility (Polczyk, 2016). Toward this end, we systematically integrated research examining the correlation between different suggestibility scales using random-effects meta-analysis. Our primary goal was to quantify the magnitude of the correlations between different forms of suggestibility. In addition, in order to identify the factors that contribute to variability in these correlations across different studies, we also examined whether these correlations were moderated by different pre-specified measurement variables, including context, scale type, and methodological quality.

## 2. Method

### 2.1. Transparency and openness promotion

We adhered to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) (Page et al., 2021) and Meta-analysis of Observational Studies (MOOSE) guidelines for systematic reviews (Brooke et al., 2021). The data included in the meta-analysis are freely available in the original papers. This meta-analysis was pre-registered on the Open Science Framework (osf.io/dq9au).

### 2.2. Inclusion criteria

The following inclusion criteria were applied for papers: published in English; published after 1958, corresponding to the introduction of the first modern standardized hypnotic suggestibility scale (Acunzo & Terhune, 2021; Laurence et al., 2008; Woody & Barnier, 2008); report at least one correlation coefficient between two suggestibility scales and the corresponding sample size; at least one of the scales must have been administered in a non-hypnotic context (i.e., without a preceding hypnotic induction); and the scales must be identified as measures of suggestibility or a related variable that involves the administration of suggestions (e.g., hypnotizability, hypnotic susceptibility, phenomenological control). Due to an omission in the pre-registration (we included no reference to age inclusion criteria), we deviated from the pre-registration and only included studies with an adult sample (age > 17).

**Table 1**  
Demarcation of suggestibility types according to Polczyk (2016).

Domain	Characteristics	Suggestion phrasing	Subtype	Representative scale
Direct verbal suggestibility	Overt administration of suggestions	“When you open your eyes, you will see a circle on the screen.” [positive visual hallucination]	Non-hypnotic suggestibility	BSS
		“Your arm is incredibly stiff and rigid, and cannot be bent.” [arm paralysis]	Hypnotic suggestibility	HGHS:A
Indirect suggestibility	Covert administration of suggestions	“He was wearing a red shirt, right?” [false memory]	Interrogative suggestibility	GSS
		“You might feel a light tingling sensation on your fingertips after touching such a warm light.” [positive somatosensory hallucination]	Sensory suggestibility	SSS

Notes. BSS = Barber Suggestibility Scale; HGHS:A = Harvard Group Scale of Hypnotic Suggestibility: Form A; GSS = Gudjonsson Suggestibility Scale; SSS=Sensory Suggestibility Scale.

### 2.3. Search strategy

PubMed and PsycINFO databases were searched independently by one reviewer (MVS) in November 2021 (updated in March 2023 and May 2024) for eligible papers indexed from 1959 onwards. The search string consisted of “(hypno\* OR suggest\*) AND (suggestibility)”; eligible papers from the two searches were integrated into a single database. The reference lists of all eligible papers and relevant review papers were manually searched to identify any additional studies that met eligibility criteria. One author group was contacted twice over a period of one month regarding methodological ambiguities but did not respond to queries, resulting in exclusion.

### 2.4. Study selection

Two reviewers (MVS, AF) independently screened titles and abstracts of articles returned by initial searches; articles that did not meet eligibility criteria were rejected. The full texts of the remaining articles were independently reviewed by the same two reviewers, who then compiled a final list of articles. Disagreements at either stage were discussed with a third reviewer (DBT). Two author groups were contacted to provide data to be included: one author group no longer had access to necessary data, and another was non-responsive to inquiries, resulting in exclusion.

### 2.5. Data extraction

The two reviewers independently coded and extracted data from all eligible studies using an extraction form that included the following: (i) study details (title, year, author funding source); (ii) demographics (age, sex, education, clinical status); (iii) study design details (temporal delay between scale administrations; clinical context; inclusion of a hypnotic induction; live or audio recording of scale administration; group or individual context); (iv) descriptive statistics (*Ms*, *SDs*, *Ns*) and correlation coefficient(s); and (v) scale types (direct or indirect suggestibility [Polczyk, 2016]) and percentage overlap of suggestion content between scale pairs, as calculated by the reviewers through a standardized form. Whenever multiple suggestibility-pair correlations were extracted from a study, the foregoing coding was performed on all pairs. The raters had 79% agreement ( $n = 13$ ) and resolved discrepancies through discussion with a third reviewer (DBT), while MVS and DBT independently extracted a subset of papers ( $n = 3$ ) from the March 2023 search with an 85% agreement rate and resolved discrepancies through discussion. Of the 84 variables extracted, the lowest agreement rates (65–71%) pertained to administration order; suggestion type (cognitive-perceptual; inhibitory motor; ideomotor; posthypnotic amnesia); counterbalancing; scale psychometric properties; and sample education.

### 2.6. Study quality

An 11-item scale was developed to assess study quality (see Supplementary Materials). Items were adapted from previous meta-analyses (Thompson et al., 2019; Wieder et al., 2021; Wieder et al., 2022) and included items based on Cochrane criteria (Higgins et al., 2008) and PRISMA (Page et al., 2021) recommendations. Study methodological quality was evaluated in this manner because conventional risk of bias measures (e.g., Higgins et al., 2022) are more well-suited to randomized controlled trials and omit important criteria relevant to this research project (e.g., experimenter and participant blinding). The two reviewers independently rated each item categorically (0 = criterion not met, 1 = criterion met); discrepancies were resolved through discussion with DBT, and a summed total score was computed for each study. Agreement between raters was 78% (Cohen's kappa = 0.51) and 82% (kappa = 0.62) for the first and second subsets, respectively (see above), reflecting acceptable inter-rater reliability. The lowest agreement ranged from 60 to 70% and concerned adequate description of study inclusion criteria

and sample demographics; rigor of suggestibility scale (reliability and validity); and inclusion of a correction for compliance (see Supplementary Materials).

### 2.7. Meta-analysis and meta-regression

All analyses were performed using Jamovi (The jamovi project, 2023) and R (R Core Team, 2023) using the metafor package (Viechtbauer, 2010). Individual study effect sizes included correlation coefficients that were transformed to z-scores using Fisher's  $r$ -to- $z$  transformation ( $Z_r$ ) and analysed using random-effects meta-analysis with the DerSimonian-Laird method. We also report raw correlation coefficients for ease of interpretation. Heterogeneity was computed through  $I^2$  and  $\tau^2$ , with  $I^2 > 50\%$  interpreted to reflect moderate or greater heterogeneity and  $\tau^2$  describing the standard deviation of effect sizes across populations, with larger values indicating greater heterogeneity. We assessed publication bias by examining funnel plots of effect sizes against standard errors for asymmetry and tested for funnel plot asymmetry using Egger's bias test (Egger et al., 1997), where  $p < .05$  is indicative of asymmetry. We also estimated an asymmetry-corrected effect sizes using the trim-and-fill method (Duval & Tweedie, 2000).

In order to account for heterogeneity in the data, we performed subgroup analyses, when five or more studies could be included, and a series of meta-regression analyses using binary and continuous moderators when there was a minimum of two studies per moderator level and ten studies total, respectively. Moderator variables with  $>2$  levels were decomposed into simpler 2-level moderators. The included moderators were collectively agreed upon by the research team based on practical and theoretical considerations and were pre-specified in our pre-registration; the moderators included 22 categorical variables (0 = absent/no, 1 = present/yes, unless otherwise specified): use of a hypnotic induction for one scale; both scales measured in a hypnotic context (e.g., mention of hypnosis in relation to one scale); suggestibility scale type (0 = indirect, 1 = direct); sample type; administration type (0 = recording, 1 = live); temporal delay (in days) between scale administrations (0 = no delay; 1 = delay of 1 or more days); scale type agreement (0 = type incongruence [e.g., direct-indirect], 1 = congruence [e.g., direct-direct]); scale psychometric properties (0 = poor, 1 = acceptable); measurement of the same scale (0 = different; 1 = same); eight of the 11 individual methodological items (due to our criteria of two studies per level of the moderator); and five continuous variables: year of scale publication; temporal delay (in days) between scale administrations; percentage of suggestion item overlap in scale pair; and methodological quality total rating.

## 3. Results

### 3.1. Study inclusion

A PRISMA diagram presenting study selection can be found in Fig. 1. The final sample of 16 papers reported 18 studies from which 55 suggestibility scale correlation pairs were extracted ( $M = 3.11$  pairs/study,  $SD = 7.12$ ; range = 1–13) (see Table 2). The correlation pairs corresponded to associations between two direct verbal suggestibility scales ( $k = 37$ ), 92% of which included a hypnotic induction for one scale; one indirect and one direct verbal suggestibility scale ( $k = 13$ ), all of which included a hypnotic induction; and two indirect suggestibility scales ( $k = 5$ ). Of the pairs that included one or more indirect measures ( $k = 18$ ), 80% included an interrogative suggestibility scale, and 20% included a sensory suggestibility scale.

### 3.2. Methodological quality criteria

Binary methodological quality criteria scores were summed for a total methodological quality score per correlation pair (theoretical range: 0–11) (see Supplementary Table 1). Studies tended to meet more

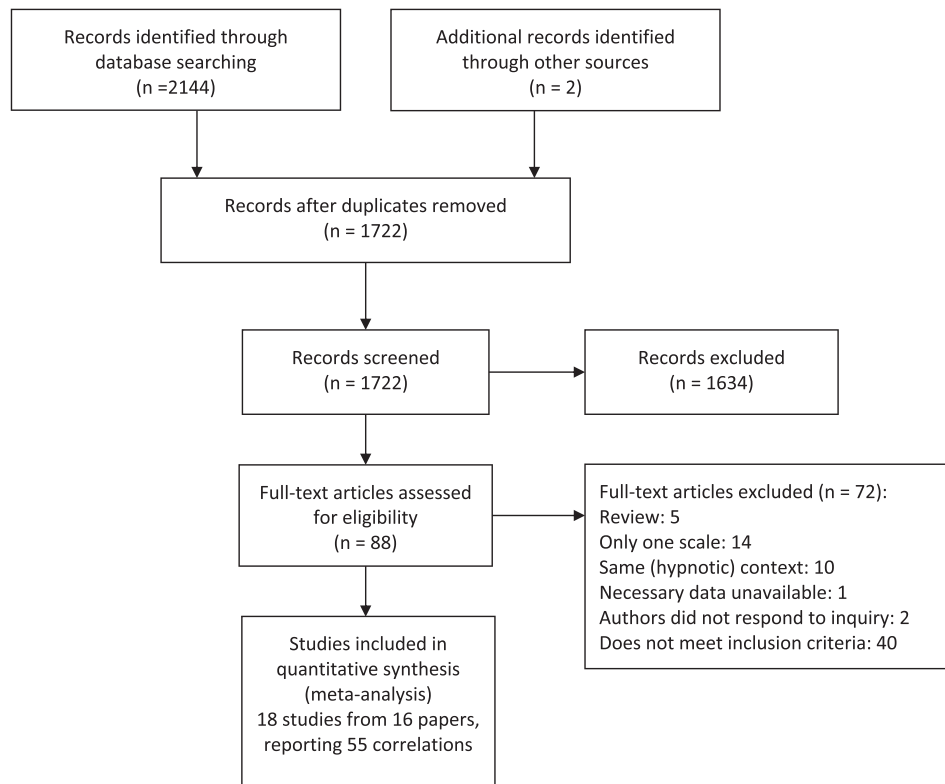


Fig. 1. PRISMA flowchart of study selection process for suggestibility correlation pairs.

than half of the criteria ( $M \pm SD$ :  $7 \pm 1.24$ ; range: 4–8). Only one study (6%) described the inclusion/exclusion criteria in adequate detail, two studies (12%) included or reported a blind experimenter, and one study (6%) included a measure to correct scale scores for compliance.

### 3.3. Meta-analysis of correlations between suggestibility scales

#### 3.3.1. All correlation pairs

A random-effects meta-analysis of correlation coefficients (Fisher's  $r$ -to- $Z$  [ $Z_r$ ]) between suggestibility scale pairs ( $k = 55$ ) revealed an overall significant association,  $Z_r = 0.58$  [95% CI: 0.48, 0.68],  $Z = 11.61$ ,  $p < .001$ , corresponding to a large positive correlation,  $r = 0.50$  [0.43, 0.56] (Fig. 2). However, as can be seen in Fig. 2, there was considerable heterogeneity in effect sizes,  $I^2 = 93\%$ ,  $\tau^2 = 0.22$ , with correlations ranging from  $Z_r = -0.11$  ( $r = -0.11$ ) to  $Z_r = 1.59$  ( $r = 0.92$ ). Due to the substantial variability in the number of inter-dependent correlation pairs across studies (1–12; see Table 1), we re-performed the primary analyses using the mean correlation per study ( $k = 16$ ). The results were comparable to those of the analysis on all pairs (see Supplementary Materials).

#### 3.3.2. Suggestibility subgroup correlation pairs

To identify possible sources of heterogeneity in effect sizes (Fig. 2), we first analysed subgroups of suggestibility scale pair correlations according to whether scales included direct or indirect suggestions. As with the main analyses, we re-performed all primary analyses using the mean correlation per study (see Supplementary Results). Direct-direct suggestibility scale pairs ( $k = 37$ ) exhibited a large, positive correlation,  $Z_r = 0.72$  [0.65, 0.79],  $Z = 20.68$ ,  $p < .001$ , corresponding to  $r = 0.62$ , with moderate heterogeneity,  $I^2 = 72\%$ ,  $\tau^2 = 0.03$ . Similarly, indirect-indirect suggestibility scale pairs ( $k = 5$ ) were highly positively correlated,  $Z_r = 1.06$  [0.49, 1.62],  $Z = 3.69$ ,  $p < .001$ , corresponding to  $r = 0.73$ , albeit with pronounced heterogeneity  $I^2 = 94\%$ ,  $\tau^2 = 0.62$ . This heterogeneity is plausibly attributed to the inclusion of a single study

that included a sensory suggestibility scale (all other indirect-indirect pairs included only interrogative scales). Indeed, the correlation was greater when the analysis was restricted to interrogative-interrogative scale pairs ( $k = 4$ ),  $Z_r = 1.27$  [0.95, 1.58],  $Z = 7.80$ ,  $p < .001$ ,  $r = 0.86$ , with a substantial reduction in heterogeneity,  $I^2 = 69\%$ ,  $\tau^2 = 0.07$ . By contrast, mixed scale pairs (direct-indirect [ $k = 13$ ]), including direct-interrogative ( $k = 9$ ), were weakly, albeit significantly, positively correlated,  $Z_r = 0.07$  [0.01, 0.12],  $Z = 2.48$ ,  $p = .013$ ,  $r = 0.07$ , and  $Z_r = 0.07$  [0.01, 0.14],  $Z = 2.16$ ,  $p = .030$ ,  $r = 0.07$ , respectively, with no evidence for heterogeneity,  $I^2s = 0\%$ ,  $\tau^2s = 0.00$ – $0.01$ . Taken together, these results suggest a clear scale congruence effect: suggestibility scale correlations were large in magnitude (and larger than the aggregate correlation in all pairs) when the two scales included the same types of suggestions (direct or indirect), albeit with moderate-to-high heterogeneity. By contrast, when suggestion type was discordant (direct-indirect), scale pair correlations were negligible and near-zero in magnitude, albeit marginally significant, with no heterogeneity. These results corroborate proposals that direct and indirect suggestibility scales measure discrete suggestion effects (Polczyk, 2016) but highlight the need for further analyses to identify the variables that moderate these correlations.

#### 3.3.3. Publication bias

Egger's test was used to assess publication bias and revealed significant funnel plot asymmetry,  $Z = 2.41$ ,  $p = .016$ , in the total sample of scale pairs ( $k = 55$ ) (Fig. 3), suggesting evidence consistent with publication bias. Using the trim-and-fill method, we re-estimated the effect size, which was reduced, albeit still significant, and moderate in magnitude,  $Z_r = 0.48$  [0.34, 0.59], corresponding to  $r = 0.45$ , a decrease of 0.05. By contrast, we did not observe evidence for funnel plot asymmetry in direct-direct ( $k = 37$ ),  $Z = 0.26$ ,  $p = .79$ , or direct-indirect ( $k = 13$ ),  $Z = -1.03$ ,  $p = .30$ , scale pairs (see Supplementary Fig. 2) (indirect-indirect pairs were not analysed due to the small number of correlation pairs;  $k = 5$ ). These results are in line with possible

**Table 2**  
Characteristics of included samples reporting correlations among suggestibility scales.

Source	Sample type	n (% female)	Age (M (SD))	Direct verbal suggestibility scale(s)	Indirect suggestibility scale(s)	Correlation pair (s)
Barber et al., 1962	Non-clinical	30 (33%)	24 (–)	BSS (hypnotic), BSS (non-hypnotic)	–	1
Barber et al., 1964	Clinical	64 (67%)	–	BSS (hypnotic), BSS (non-hypnotic)	–	4
Braffman & Kirsch, 1999	University students	92 (68%)	18.35 (1.04)	CURSS (hypnotic), CURSS (non-hypnotic)	–	2
David & Brown, 2002	University students	31 (54%)	21 (–)	Romanian versions of HGSHS:A, CIS	Romanian version of GSS 2	3
Gudjonsson, 1987a	University students	28 (50%)	27 (7.7)	–	GSS 1, GSS 2	1
Gudjonsson, 1987b	Clinical	32 (15%)	31 (11.9)	–	GSS 1, GSS 2	1
Gudjonsson, 1987c	Clinical	30 (6%)	31 (10.3)	–	GSS 1, GSS 2	1
Hilgard et al., 1966a,b	University students	20 (–)	–	SHSS:C (hypnotic), SHSS:C (non-hypnotic)	–	2
Hilgard et al., 1966c,d,e	University students	15 (–)	–	SHSS:C (hypnotic), SHSS:C (non-hypnotic)	–	3
Lush et al., 2021a	University students	123 (81%)	19.8 (3.9)	SWASH, PCS	–	1
Lush et al., 2021b	University students	123 (81%)	19.7 (1.7)	SWASH, PCS	–	1
Malinoski & Lynn, 1999	University students	133 (68%)	19.5 (1.1)	HGSHS:A	GSS 2	3
Milling et al., 2003	University students	167 (62%)	–	CURSS (hypnotic), CURSS (non-hypnotic)	–	1
Milling et al., 2005	University students	60 (70%)	–	CURSS (hypnotic), CURSS (non-hypnotic)	–	1
Milling et al., 2010 Study 1	University students	173 (47%)	18.95 (1.05)	CURSS (hypnotic), CURSS (non-hypnotic)	–	3
Milling et al., 2010 Study 2	University students	211 (66%)	18.95 (1.05)	CURSS (hypnotic), CURSS (non-hypnotic)	–	3
Milling et al., 2010 Study 3	University students	143 (44%)	18.95 (1.05)	CURSS (hypnotic), CURSS (non-hypnotic)	–	3
Pires et al., 2013	University students	51 (84%)	21.4 (5.7)	–	Portuguese version of GSS 1, GSS 2	1
Polczyk, 2016	Non-clinical	118 (58%)	22.2 (2)	Polish versions of HGSHS:A, BSS	Polish versions of GSS 2, SSS	13
Spinhoven et al., 1991	Mixed	265 (65%)	34.4 (12.7)	Dutch versions of SHSS:A, CIS	–	1
Weitzenhoffer et al., 1961	University students	30 (–)	–	Modified SHSS:B, Modified SHSS:A	–	5
Wieder & Terhune, 2019	University students	58 (–)	–	Modified HGSHS:A, BSS-C	–	1

Notes: Lowercase letters denote different samples within a single study; – = not reported; BSS=Barber Suggestibility Scale; BSS-C=Brief Suggestibility Scale-Composite, CIS=Creative Imagination Scale; CURSS=Carleton University Responsiveness to Suggestion Scale; GSS 1 & 2 = Gudjonsson Suggestibility Scale 1 & 2; HGSHS:A = Harvard Group Scale of Hypnotic Suggestibility: Form A; PCS=Phenomenological Control Scale; SHSS:A/B = Stanford Hypnotic Susceptibility Scale: Form A/B; SSS=Sensory Suggestibility Scale; SWASH=Sussex Waterloo Scale of Hypnotisability.

publication bias in the full sample of correlation pairs, but this effect is potentially attributable to heterogeneity of effect sizes according to scale pair concordance.

### 3.3.4. Meta-regression analyses

Owing to the moderate-to-pronounced heterogeneity in the total sample and scale pair subgroups, we undertook a series of meta-regression analyses to identify the variables that moderate the magnitude of scale pair correlations.

Our first set of meta-regression analyses compared correlation pairs across different scale subgroups. Indirect-indirect scales were marginally more highly correlated than direct-direct scales ( $k = 42$ ),  $\Delta Z_r = 0.27$  [0.02, 0.52],  $Z = 2.12$ ,  $p = .033$ , with moderate heterogeneity,  $I^2 = 79\%$ ,  $\tau^2 = 0.04$ . Scale correlations were also significantly greater between direct-direct scales than between direct-indirect scales ( $k = 50$ ),  $\Delta Z_r = 0.65$  [0.55, 0.76],  $Z = 11.72$ ,  $p < .001$ , again, with moderate heterogeneity  $I^2 = 64\%$ ,  $\tau^2 = 0.02$ . Similarly, correlations were greater for indirect-indirect pairs than for direct-indirect pairs ( $k = 18$ ),  $\Delta Z_r = 0.92$  [1.17, 0.66],  $Z = 7.08$ ,  $p < .001$ , with moderate heterogeneity,  $I^2 = 77\%$ ,  $\tau^2 = 0.03$ . These effects remained stable when we only included direct-interrogative and direct-sensory scale pairs among the latter group (see

Supplementary Results).<sup>1</sup> We repeated these analyses after omitting the indirect-indirect correlation pair between interrogative and sensory suggestibility, which did not substantially change the results (see Supplementary Results). Taken together, these results complement the foregoing subgroup analyses and demonstrate that suggestibility scale correlations are greater when the suggestion type (direct vs. indirect) is congruent than when it is not.

Our next set of meta-regression analyses examined whether correlations were moderated by methodological quality in the total sample and scale subgroups (Table 3). Methodological quality total scores did not significantly moderate suggestibility scale correlations in the full sample or scale subgroups. The number of methodological quality items that could be analysed as moderators differed across analyses due to sample size requirements (see Methods). In the total sample, only clear reporting of where participants were recruited from was a significant positive moderator. Similarly, in direct-direct correlation pairs, reporting of how and from where participants were recruited were significant, albeit weak, positive moderators. The former effect was similarly

<sup>1</sup> We were unable to contrast the interrogative-sensory suggestibility scale correlation against other scale pairs due to a limited sample size ( $k = 1$ ).

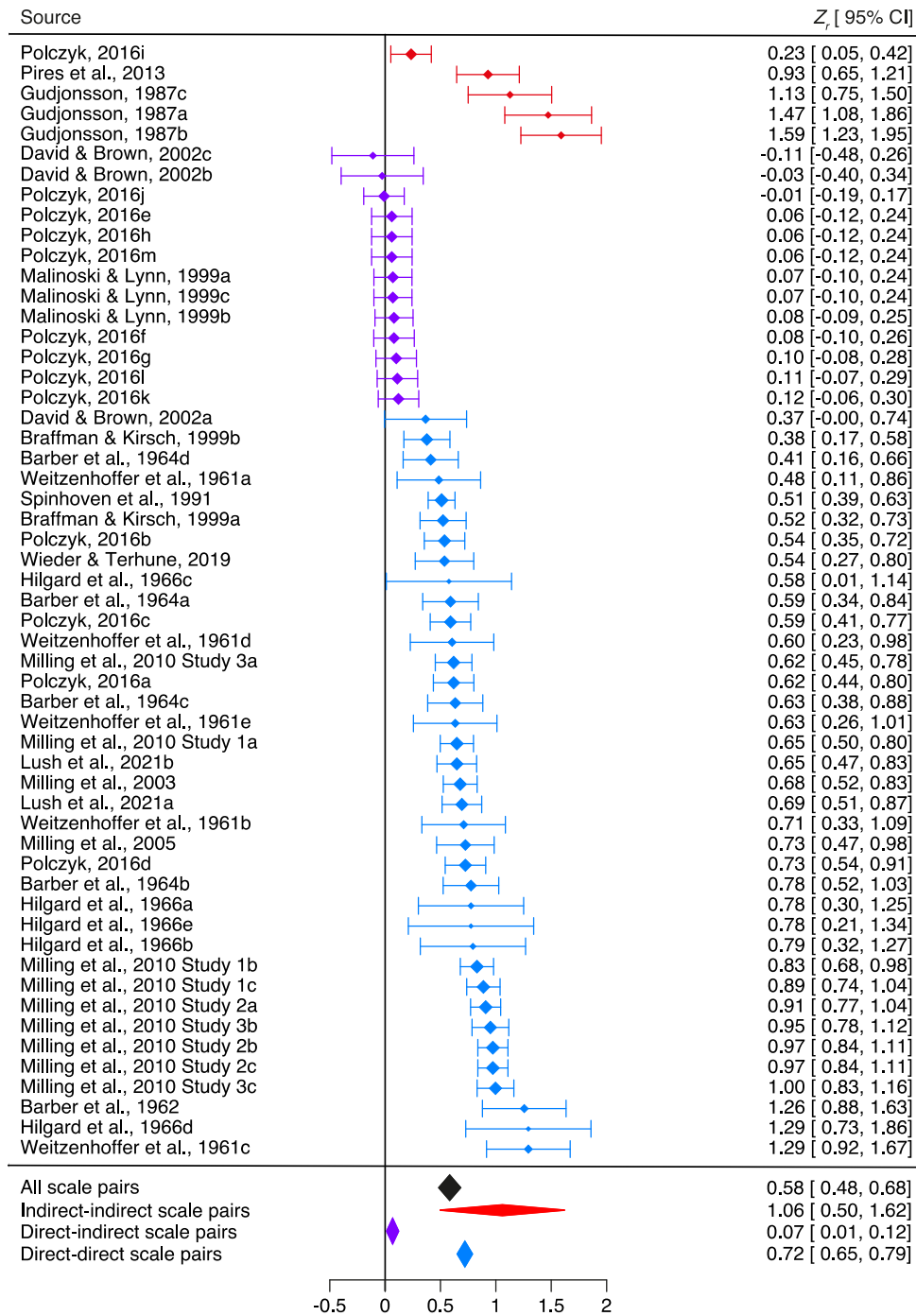


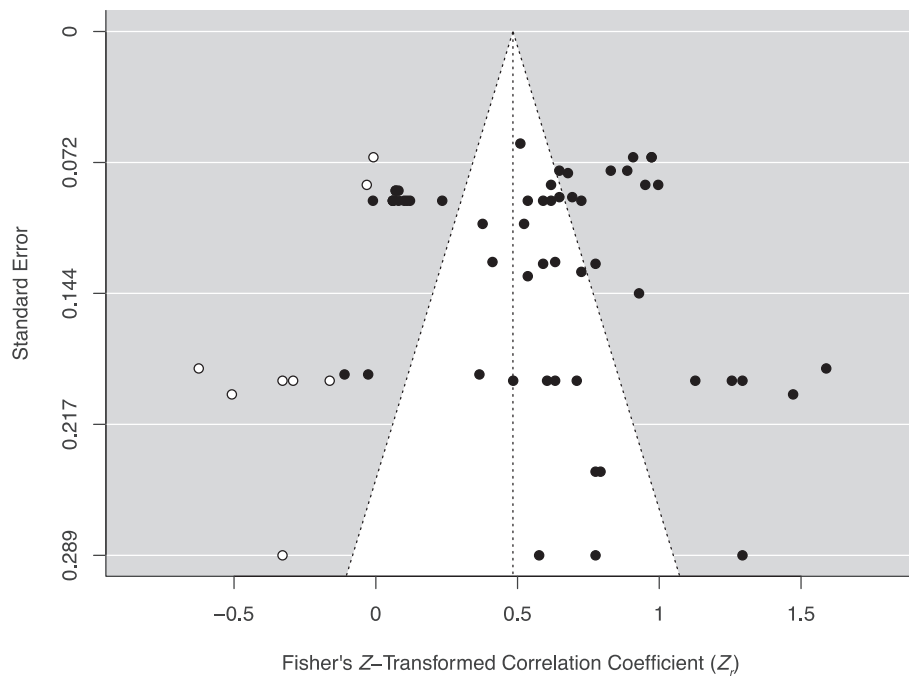
Fig. 2. Forest plot of suggestibility scale pair correlations ( $Z_r$ ) [95% CIs] for all pairs and scale pair subgroups. Notes. Marker sizes reflect study weights, with smaller and larger markers denoting smaller and larger weights, respectively.

observed in indirect-indirect pairs,  $\Delta Z_r = 0.83$  [0.11, 1.54],  $Z = 2.72$ ,  $p = .023$ .<sup>2</sup> Taken together, these results indicate that studies that clearly reported recruitment information tended to display higher correlations between suggestibility scales and more broadly suggest that the correlations between suggestibility scales are unlikely to be driven by poor methodological quality.

The final set of meta-regression analyses examined whether different

<sup>2</sup> Our meta-regression analyses within the indirect-indirect subgroup were limited due to our criterion of two moderators per level. We were only able to analyse one moderator within this subgroup.

methodological features of the studies moderated suggestibility scale correlations in the total sample and scale subgroups (Table 3). In the total sample, scale correlations were moderated by the interval between scale administrations, and the hypnotic context, such that correlations were greater when administered on the same day or if both were administered in a hypnotic context, respectively. Scale type agreement (i.e., both scales being either direct or indirect) was the largest moderator of correlations and indicated that correlations were larger when scales used congruent scale types. Suggestibility scale pair correlations were also greater when the two administrations included the same scale compared with different scales. Similarly, the percentage overlap of suggestion content was also a significant, albeit weak, positive



**Fig. 3.** Funnel plot of scale-pair correlation coefficients ( $Z_r$ ) as a function of standard error in all suggestibility pairs ( $k = 55$ ). Filled circles denote individual study effect sizes, and empty circles denote estimated missing  $Z_r$  values attributable to potential publication bias imputed using the trim-and-fill method.

moderator of scale pair correlations, indicating that scales that included similar suggestions yielded slightly larger inter-scale correlations. Among direct-direct scale pairs ( $k = 37$ ), the inclusion of a hypnotic induction was the only significant negative moderator, indicating that correlations were larger when both scales omitted an induction. Conversely, the live administration of a scale, use of the same scale, same-day administration, and suggestion content overlap were significant positive moderators within the direct-direct subgroup. All moderators were non-significant for direct-indirect scale pairs and indirect-indirect scale pairs (see Supplementary Table 2). Cumulatively, these results confirm that scale type congruence is associated with larger suggestibility scale correlation, while also highlighting the roles of contextual and psychometric factors.

#### 4. Discussion

We conducted a random effects meta-analysis of correlations between standardized suggestibility scales in order to better characterize the structure of suggestibility. There was a positive association between suggestibility scales, although correlations were larger between scales that employ congruent suggestion types (direct-direct and indirect-indirect) compared to those with discordant suggestion types (direct-indirect). Moderation analyses highlighted the importance of context and mode of assessment when measuring suggestibility and indicated that inter-scale correlations were stronger when the scales were more similar in suggestion content. These results broadly align with the proposal that standardized suggestibility scales index two mostly discrete abilities: direct verbal suggestibility and indirect suggestibility, the latter of which primarily comprises interrogative suggestibility (Polczyk, 2016). This measurement dissociation points toward the need for refinements in the conceptualization of suggestibility within research and clinical contexts.

The strong inter-scale correlations among direct-direct and indirect-indirect suggestibility scales reinforce their stability as cognitive-perceptual traits (Piccione et al., 1989). Our results confirm the strong association between hypnotic and non-hypnotic suggestibility (Braffman & Kirsch, 1999; Hilgard & Tart, 1966) and challenge the view that hypnotic suggestibility represents a unique cognitive capacity

beyond responsiveness to direct verbal suggestions (Tasso et al., 2020). The same held for interrogative suggestibility scales, which exhibited strong inter-scale correlations, although the high correlations for this suggestibility subtype are potentially artefactual of the uniform use of a single scale in this research domain (Gudjonsson, 1984). The stability of these suggestibility domains underscores the importance of measuring domain-relevant forms of suggestibility (direct verbal suggestibility and interrogative suggestibility) in clinical, experimental, and forensic applications of suggestion.

The stark contrast in correlations between direct and interrogative scale pairs highlights the differential nature of suggestion-type effects within these two domains. As can be seen in Fig. 4, direct verbal suggestibility scales exhibit robust domain-specific associations (Braffman & Kirsch, 1999). By contrast, the two types of indirect suggestibility scales (interrogative and sensory suggestibility) were only weakly associated, although this should be treated with caution given the small number of studies assessing this association. This leaves open the question of whether interrogative and sensory suggestibility are sufficiently strongly related to warrant subsuming them under the hypothesised category of indirect suggestibility (Polczyk & Pasek, 2006). Beyond these domain-specific effects, the correlations across direct and indirect suggestibility scales were uniformly weak, which aligns with the hypothesis that these scales are indexing different cognitive abilities (Polczyk, 2016). Although hypnotic suggestibility significantly correlated with interrogative suggestibility, these associations are plausibly driven by context effects (e.g., same day and context of administration; Council, 1993). More broadly, the observed measurement dissociation between direct and interrogative suggestibility calls into question the broad applicability of the term *suggestibility* as traditionally understood within the psychological literature (Halligan & Oakley, 2014). This is especially important given that general parlance pertaining to this term tends to align more with interrogative than direct verbal suggestibility. Accordingly, the present results clearly argue for a more nuanced usage of the term: researchers and practitioners should always specify the subtype of suggestibility being measured or hypothesised in order to avoid ambiguity. Going a step further, insofar as these distinct forms of suggestibility are often confused with one another (see Wieder et al., 2023), the present results warrant renewed attention to the theoretical

**Table 3**  
Meta-regression analyses of suggestibility scale pair correlations in all scale pairs and scale subgroup pairs.

Moderator	Suggestibility scale pair ( <i>k</i> ) correlation difference $\Delta Z_r$ [95% CIs]			All scale pairs			Direct-direct			Direct-indirect		
	All scale pairs (55)	Direct-direct (37)	Direct-indirect (13)	Z	<i>p</i>	<i>I</i> <sup>2</sup>	Z	<i>p</i>	<i>I</i> <sup>2</sup>	Z	<i>p</i>	<i>I</i> <sup>2</sup>
Clear study objectives	–	–	–	–	–	–	–	–	–	–	–	–
Clear sample origin	0.37 [0.18, 0.56]	0.16 [0.01, 0.32]	–0.02 [–0.13, 0.09]	3.94	<.001	89%	2.07	.039	66%	–0.34	.73	0%
Clear sample recruitment	0.16 [–0.02, 0.35]	0.18 [0.05, 0.31]	–0.02 [–0.13, 0.09]	1.75	.080	90%	2.76	.006	64%	–0.29	.77	0%
Clear inclusion/exclusion criteria	0.02 [–0.36, 0.39]	–0.13 [–0.35, 0.08]	–	0.11	.91	92%	–1.17	.24	72%	–	–	–
Experimenter blinding	0.11 [–0.26, 0.48]	–0.04 [–0.25, 0.16]	–	0.58	.55	92%	–0.38	.70	72%	–	–	–
Rigorous suggestibility scale A	–0.16 [–0.46, 0.14]	0.01 [–0.18, 0.15]	–	–1.01	.31	92%	0.06	.95	73%	–	–	–
Rigorous suggestibility scale B	0.15 [–0.12, 0.43]	–0.08 [–0.31, 0.15]	–0.01 [–0.13, 0.11]	1.12	.26	91%	–0.68	.49	73%	–0.21	.83	0%
Clear sample characteristics	–0.21 [–0.45, 0.04]	–0.02 [–0.18, 0.14]	–	–1.99	.096	92%	–0.23	.81	73%	–	–	–
Outcome data available	–0.17 [–0.54, 0.18]	–0.03 [–0.26, 0.21]	–	–0.95	.34	92%	–0.24	.81	73%	–	–	–
Hypnotic induction (one scale)	–0.06 [–0.31, 0.18]	–0.58 [–1.08, –0.08]	–	–0.52	.60	91%	–2.31	.021	71%	–	–	–
Hypnotic context (one scale)	0.23 [0.04, 0.42]	0.05 [–0.12, 0.22]	–0.02 [–0.13, 0.09]	2.39	.016	90%	0.65	.51	72%	–0.34	.73	0%
Same day administration	0.27 [0.07, 0.47] <sup>a</sup>	0.17 [0.02, 0.32] <sup>c</sup>	–	2.62	.009	85%	2.23	.025	65%	–	–	–
Live administration of scales	–0.00 [–0.30, 0.30] <sup>b</sup>	0.20 [0.01, 0.37]	–	–0.01	.99	90%	2.03	.042	70%	–	–	–
Scale type agreement	0.69 [0.56, 0.82]	–	–	10.35	<.001	74%	–	–	–	–	–	–
Same suggestibility scale	0.38 [0.23, 0.54]	0.19 [0.07, 0.32]	–	4.79	<.001	86%	2.99	.003	64%	–	–	–
Use of healthy sample	–0.23 [–0.56, 0.07]	–	–	–1.49	.14	92%	–	–	–	–	–	–
Good psychometric properties	–0.14 [–0.36, 0.07]	0.10 [–0.05, 0.23]	–	–1.29	.19	92%	1.29	.20	70%	–	–	–
Suggestion content overlap (%)	0.01 [0.01, 0.01]	0.003 [0.001, 0.005]	0.001 [–0.00, 0.00]	8.36	<.001	81%	2.94	.003	65%	0.41	.68	0%
Temporal delay (in days)	–0.00 [–0.01, 0.01] <sup>a</sup>	–0.003 [–0.01, 0.00] <sup>c</sup>	–	–0.87	.38	90%	–1.72	.076	67%	–	–	–
Methodological quality	0.05 [–0.02, 0.12]	0.03 [–0.03, 0.08]	–0.01 [–0.05, 0.04]	1.51	.13	90%	0.73	.46	69%	–0.31	.75	0%

Notes.  $\Delta Z_r$  = Fisher's *r*-to-*z* correlation difference; *I*<sup>2</sup> = larger values denote more heterogeneity; for all binary moderators, 0 = variable absent, 1 = variable present. - = Unable to analyse due to criterion of two correlations per level of the binary moderator; sample sizes vary within subgroups because of missing data. Indirect-indirect scale pairs were not analysed due to an insufficient number of studies (*k* = 5).<sup>a</sup> *k* = 34; <sup>b</sup> *k* = 52; <sup>c</sup> *k* = 25. Significant results are in bold.

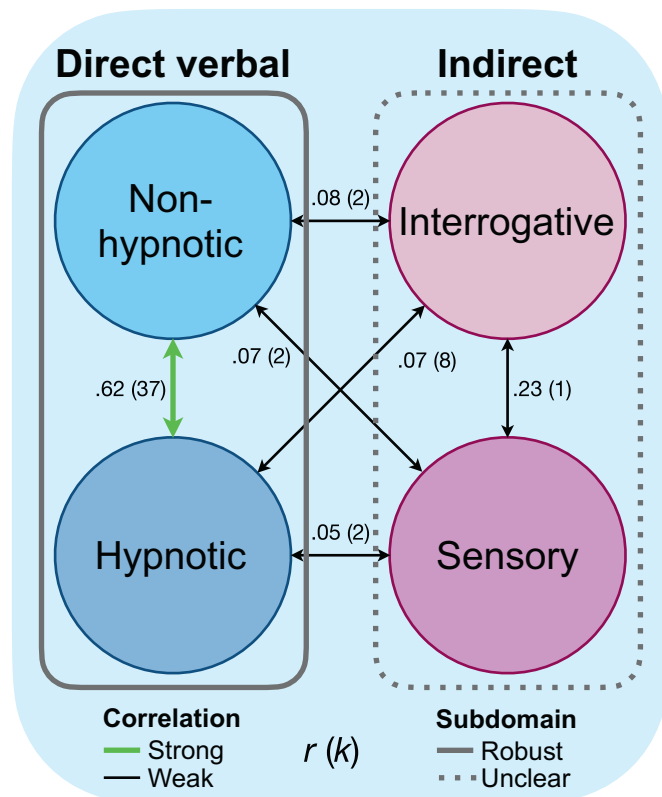
value of the term suggestibility in psychological research.

The moderation analyses present novel information regarding contextual features that shape suggestibility assessment. Scales administered in a hypnotic context (e.g., reference to hypnosis in relation to one of the scales) yielded larger inter-scale correlations than for those that were not; in the former case, this may be attributable to the scales being administered in close temporal proximity (e.g., the same day), reflecting a context effect (Council, 1993). This context artefact may also explain why the administration of the same scale yielded larger positive correlations within the total sample and the direct-direct suggestibility subgroup, as many of the studies that administered the same scale did so in a hypnotic context. The percentage of suggestion content overlap was a very weak moderator and could either reflect the presence of higher correlations among subcomponents of suggestibility (e.g., motor vs. cognitive suggestions; Barnier et al., 2022; Woody et al., 2005) or an additional artefact of context administration effects. Insofar as a large proportion of studies in our sample administered the same scale, this may have artificially decreased the correlation difference across levels of suggestion content overlap. Conversely, the absence of a hypnotic induction yielded larger effect sizes within the direct-direct subgroup, potentially because studies that employed an induction often had longer temporal delays. Individual differences in (hypnotic) suggestibility might additionally moderate scale-type associations when an induction is included. These results highlight the salience of temporal proximity in assessing scale-type associations and suggest the association among scales is probably inflated when scales are administered in the same

context, use the same scale, or include two scales with overlapping suggestion content (Council, 1993). On the other hand, studies that clearly reported participant recruitment methods demonstrated larger correlation differences, signifying that larger correlations are unlikely to be an artefact of poor methodology. Cumulatively, these results indicate the need for greater attention to contextual moderators in the assessment of suggestibility.

#### 4.1. Limitations and future directions

Despite the clear results of this meta-analysis, there are multiple limitations within the original studies that warrant consideration when interpreting our results. First, our analyses of direct-indirect and indirect-indirect suggestibility correlations were limited by the small number of available studies. In turn, these results should be treated with caution and future work should continue to evaluate these associations. Second, many of the analysed correlations were drawn from the same studies and thus were not independent. We were able to address this limitation in our analysis of all suggestibility correlations as well as the direct-direct suggestibility correlation by analysing only the average correlations from studies. These results yielded comparable results to the main analyses and thus the central results are not artefactual of this limitation. However, we were unable to perform a similar analysis of average correlations in the analyses of direct-indirect and indirect-indirect correlations due to the small number of unique studies (see Supplementary Results). However, this limitation is unlikely to detract



**Fig. 4.** Schematic depiction of the demarcation between direct verbal and indirect suggestibility domains and the magnitude of associations between scales based on available evidence.

Notes. Double-headed lines denote correlations between scales. A green line indicates a strong association whereas a black line indicates a weak association. A solid grey line indicates a robust subdomain whereas a dashed grey line indicates an unclear subdomain.

from our central findings. For example, as can be seen in Fig. 2, all correlations in the indirect-indirect suggestibility subgroup were significant, whereas all correlations in the direct-indirect subgroup were non-significant, which aligns with the results of our analyses. Next, only one study included a measure to correct suggestibility scores for compliance (Wieder & Terhune, 2019). Although often unrecognized, responses on direct verbal suggestibility scales are partly contaminated by compliant responses, in which participants perform the suggested effects in a completely voluntary manner without any genuine changes in subjective experience (Bowers, 1981; Bowers et al., 1988). This issue is even more pronounced for interrogative suggestibility scales, for which responses are mostly indistinguishable from compliant responding (Gudjonsson, 2013). Future studies should incorporate compliance corrections into suggestibility measurement (e.g., Brown et al., 2008) in order to better dissociate genuine phenomenological changes in response to suggestions from simple compliance effects.

This meta-analysis similarly highlights gaps in current knowledge due to the focus of primary research studies in this literature. Alongside broader trends in the research literature, the studies in our sample focused exclusively on verbal suggestions, leaving a significant lacuna in our understanding of the interplay between verbal and non-verbal suggestions (Halligan & Oakley, 2014). In particular, many suggestions in clinical and experimental contexts are delivered through informed consent documentation, but it remains poorly understood whether direct verbal suggestibility accounts for variability in response to such textual suggestions. In addition, sensory suggestibility has received relatively little empirical attention, and thus, its associations with direct verbal suggestibility and interrogative suggestibility remain

uncertain (Fig. 4; Polczyk & Pasek, 2006; Polczyk, 2016). A similar limitation holds for self-report retrospective suggestibility scales (e.g., Kotov et al., 2004), which were not considered in this meta-analysis due to a lack of published studies examining the association between such scales with other suggestibility scales. These scales deviate from classic work-sample suggestibility scales in that they ask participants to report their responsiveness to different suggestive environmental influences (e.g., advertisements). Owing to their ease of measurement relative to work-sample scales, their use has proliferated in recent years (Corsi & Colloca, 2017; Szigeti et al., 2024). Unpublished research suggests a moderate association between retrospective and direct verbal suggestibility (Kotov et al., 2004), but further psychometric evaluations of these scales are necessary. Finally, as we did find evidence for potential publication bias, future studies should make greater efforts to publish negative, null, or mixed results to ensure the literature is representative (Nosek et al., 2022).

## 5. Conclusions

This meta-analysis corroborates proposals that direct and interrogative suggestibility scales measure discrete suggestion effects. This measurement dissociation underscores the importance of the moderating influence of different forms of suggestibility in applications of verbal suggestion in clinical, experimental, and forensic contexts. Future empirical and theoretical work can strengthen conceptual clarity and psychometric rigor by accurately specifying the form of suggestibility being measured or hypothesised.

## CRedit authorship contribution statement

**Madeline V. Stein:** Writing – review & editing, Writing – original draft, Project administration, Formal analysis, Data curation, Conceptualization. **Afik Faerman:** Writing – review & editing, Data curation, Conceptualization. **Trevor Thompson:** Writing – review & editing, Supervision, Formal analysis, Conceptualization. **Irving Kirsch:** Writing – review & editing, Writing – original draft, Data curation, Conceptualization. **Steven J. Lynn:** Writing – review & editing, Writing – original draft, Data curation, Conceptualization. **Devin B. Terhune:** Writing – review & editing, Writing – original draft, Supervision, Formal analysis, Data curation, Conceptualization.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## Declaration of competing interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Acknowledgments

This paper is dedicated to the memory of Steven J. Lynn, who passed away during the completion of the manuscript.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.paid.2025.113181>.

## Data availability

The data are freely available in the included articles

## References

- Acunzo, D. J., & Terhune, D. B. (2021). A critical review of standardized measures of hypnotic suggestibility. *The International Journal of Clinical and Experimental Hypnosis*, 69(1), 50–71. <https://doi.org/10.1080/00207144.2021.1833209>
- Barber, T. X. (1976). *Pitfalls in human research: Ten pivotal points*. Elsevier Science.
- Barnier, A. J., Terhune, D. B., Polito, V., & Woody, E. Z. (2022). A componential approach to individual differences in hypnotizability. *Psychology of Consciousness: Theory, Research and Practice*, 9(2), 130.
- Bell, V., Oakley, D. A., Halligan, P. W., & Deeley, Q. (2011). Dissociation in hysteria and hypnosis: Evidence from cognitive neuroscience. *Journal of Neurology, Neurosurgery, and Psychiatry*, 82(3), 332–339. <https://doi.org/10.1136/jnnp.2009.199158>
- Bowers, K. S. (1981). Do the stanford scales tap the “classic suggestion effect”? *International Journal of Clinical and Experimental Hypnosis*, 29(1), 42–53. <https://doi.org/10.1080/00207148108409142>
- Bowers, P., Laurence, J.-R., & Hart, D. (1988). The experience of hypnotic suggestions. *International Journal of Clinical and Experimental Hypnosis*, 36(4), 336–349. <https://doi.org/10.1080/00207148808410523>
- Braffman, W., & Kirsch, I. (1999). Imaginative suggestibility and hypnotizability: An empirical analysis. *Journal of Personality and Social Psychology*, 77(3), 578–587. <https://doi.org/10.1037/0022-3514.77.3.578>
- Brooke, B. S., Schwartz, T. A., & Pawlik, T. M. (2021). Moose reporting guidelines for meta-analyses of observational studies. *JAMA Surgery*, 156(8), 787–788. <https://doi.org/10.1001/jamasurg.2021.0522>
- Brown, R. J., Schrag, A., Krishnamoorthy, E., & Trimble, M. R. (2008). Are patients with somatization disorder highly suggestible? *Acta Psychiatrica Scandinavica*, 117(3), 232–235. <https://doi.org/10.1111/j.1600-0447.2007.01127.x>
- Corsi, N., & Colloca, L. (2017). Placebo and nocebo effects: The advantage of measuring expectations and psychological factors. *Frontiers in Psychology*, 8, 308. <https://doi.org/10.3389/fpsyg.2017.00308>
- Council, J. R. (1993). Context effects in personality research. *Current Directions in Psychological Science*, 2(2), 31–34. <https://doi.org/10.1111/1467-8721.ep10770636>
- Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455–463. <https://doi.org/10.1111/j.0006-341x.2000.00455.x>
- Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109), 629–634. <https://doi.org/10.1136/bmj.315.7109.629>
- Enck, P., & Zipfel, S. (2019). Placebo effects in psychotherapy: A framework. *Frontiers in Psychiatry*, 10, 456. <https://doi.org/10.3389/fpsyg.2019.00456>
- Eysenck, H. J., & Furneaux, W. D. (1945). Primary and secondary suggestibility; an experimental and statistical study. *Journal of Experimental Psychology*, 35, 485–503. <https://doi.org/10.1037/h0054976>
- Fassler, O., Lynn, S. J., & Knox, J. (2008). Is hypnotic suggestibility a stable trait? *Consciousness and Cognition*, 17(1), 240–253. <https://doi.org/10.1016/j.concog.2007.05.004>
- Frischholz, E. J., Lipman, L. S., Braun, B. G., & Sachs, R. G. (1992). Psychopathology, hypnotizability, and dissociation. *The American Journal of Psychiatry*, 149(11), 1521–1525. <https://doi.org/10.1176/ajp.149.11.1521>
- Geiger, E., Peter, B., Prade, T., & Piesbergen, C. (2014). Intelligence and hypnotizability: Is there a connection? *The International Journal of Clinical and Experimental Hypnosis*, 62(3), 310–329. <https://doi.org/10.1080/00207144.2014.901083>
- Gheorghiu, V. A., Hodapp, V., & Ludwig, C. M. (1975). Attempt to construct a scale for the measurement of the effect of suggestion on perception. *Educational and Psychological Measurement*, 35(2), 341–352. <https://doi.org/10.1177/001316447503500212>
- Gudjonsson, G. H. (1984). A new scale of interrogative suggestibility. *Personality and Individual Differences*, 5(3), 303–314. [https://doi.org/10.1016/0191-8869\(84\)90069-2](https://doi.org/10.1016/0191-8869(84)90069-2)
- Gudjonsson, G. H. (1987). Historical background to suggestibility: How interrogative suggestibility differs from other types of suggestibility. *Personality and Individual Differences*, 8(3), 347–355. [https://doi.org/10.1016/0191-8869\(87\)90035-3](https://doi.org/10.1016/0191-8869(87)90035-3)
- Gudjonsson, G. H. (1988). The relationship of intelligence and memory to interrogative suggestibility: The importance of range effects. *The British Journal of Clinical Psychology*, 27(2), 185–187. <https://doi.org/10.1111/j.2044-8260.1988.tb00772.x>
- Gudjonsson, G. H. (2013). Interrogative suggestibility and compliance. In *Suggestibility in legal contexts: Psychological research and forensic implications* (pp. 45–61). Wiley Blackwell.
- Halligan, P. W., & Oakley, D. A. (2014). Hypnosis and beyond: Exploring the broader domain of suggestion. *Psychology of Consciousness: Theory, Research and Practice*, 1(2), 105–122. <https://doi.org/10.1037/cns0000019>
- Higgins, J., Morgan, R., Rooney, A., Taylor, K., Thayer, K., Silva, R., Lemeris, C., Akl, A., Arroyave, W., & Bateson, T. (2022). Risk of bias in non-randomized studies of exposure (robins-e): Launch version. Risk Of Bias.
- Higgins, J. P. T., Higgins, J. P. T., Green, S., & Cochrane, C. (2008). Cochrane handbook for systematic reviews of interventions. In *Cochrane book series*. John Wiley & Sons Ltd. Cochrane Collaboration.
- Hilgard, E. R., & Tart, C. T. (1966). Responsiveness to suggestions following waking and imagination instructions and following induction of hypnosis. *Journal of Abnormal Psychology*, 71(3), 196–208. <https://doi.org/10.1037/h0023323>
- Holman, L., Head, M. L., Lanfear, R., & Jennions, M. D. (2015). Evidence of experimental bias in the life sciences: Why we need blind data recording. *PLoS Biology*, 13(7). <https://doi.org/10.1371/journal.pbio.1002190>
- Huber, A., Lui, F., & Porro, C. A. (2013). Hypnotic susceptibility modulates brain activity related to experimental placebo analgesia. *Pain*, 154(9), 1509–1518. <https://doi.org/10.1016/j.pain.2013.03.031>
- Kirsch, I. (1990). *Changing expectations: A key to effective psychotherapy*. Thomson Brooks/Cole Publishing Co.
- Kirsch, I. (1997). Suggestibility or hypnosis: What do our scales really measure? *The International Journal of Clinical and Experimental Hypnosis*, 45(3), 212–225. <https://doi.org/10.1080/00207149708416124>
- Kirsch, I. (1999). *Hypnosis and placebo: Response expectancy as a mediator of suggestion effects*. Anales de Psicología.
- Kotov, R. I., Bellman, S. B., & Watson, D. B. (2004). *Multidimensional iowa suggestibility scale (miss) brief manual* Unpublished manuscript.
- Laurence, J.-R., Beaulieu-Prévost, D., Chéné, T., & d. (2008). Measuring and understanding individual differences in hypnotizability. In *The oxford handbook of hypnosis: Theory, research, and practice* (pp. 225–253). Oxford University Press.
- Lund, K., Petersen, G. L., Erlandsen, M., De Pascalis, V., Vase, L., Jensen, T. S., & Finnerup, N. B. (2015). The magnitude of placebo analgesia effects depends on how they are conceptualized. *Journal of Psychosomatic Research*, 79(6), 663–668. <https://doi.org/10.1016/j.jpsychores.2015.05.002>
- Lynn, S. J., Neufeld, V., & Maré, C. (1993). Direct versus indirect suggestions: A conceptual and methodological review. *The International Journal of Clinical and Experimental Hypnosis*, 41(2), 124–152. <https://doi.org/10.1080/00207149308414543>
- Matthews, W. J., Lankton, S., & Lankton, C. (1993). An ericksonian model of hypnotherapy. In *Handbook of clinical hypnosis* (pp. 187–214). American Psychological Association. <https://doi.org/10.1037/10274-009>
- Milling, L. S., Valentine, K. E., LoStimolo, L. M., Nett, A. M., & McCarley, H. S. (2021). Hypnosis and the alleviation of clinical pain: A comprehensive meta-analysis. *The International Journal of Clinical and Experimental Hypnosis*. <https://doi.org/10.1080/00207144.2021.1920330>
- Montgomery, G. H., Schnur, J. B., & David, D. (2011). The impact of hypnotic suggestibility in clinical care settings. *International Journal of Clinical and Experimental Hypnosis*, 59(3), 294–309. <https://doi.org/10.1080/00207144.2011.570656>
- Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., ... Nuijten, M. B. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology*, 73, 719–748.
- Oakley, D. A., Walsh, E., Mehta, M. A., Halligan, P. W., & Deeley, Q. (2021). Direct verbal suggestibility: Measurement and significance. *Consciousness and Cognition*, 89(103), 036. <https://doi.org/10.1016/j.concog.2020.103036>
- Otgaar, H., Schell-Leugers, J., Howe, M., De La Fuente Vilar, A., Houben, S., & Merckelbach, H. (2021). The link between suggestibility, compliance, and false confessions: A review using experimental and field studies. *Applied Cognitive Psychology*, 35. <https://doi.org/10.1002/acp.3788>
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... McKenzie, J. E. (2021). Prisma 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372, Article n160. <https://doi.org/10.1136/bmj.n160>
- Parsons, R. D., Bergmann, S., Wiech, K., & Terhune, D. B. (2021). Direct verbal suggestibility as a predictor of placebo hypoalgesia responsiveness. *Psychosomatic Medicine*, 83(9), 1041–1049. <https://doi.org/10.1097/psy.0000000000000977>
- Peters, M. J., Moritz, S., Tekin, S., Jelacic, M., & Merckelbach, H. (2012). Susceptibility to misleading information under social pressure in schizophrenia. *Comprehensive Psychiatry*, 53(8), 1187–1193. <https://doi.org/10.1016/j.comppsych.2012.04.008>
- 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. <https://doi.org/10.1037/0022-3514.56.2.289>
- Polczyk, R. (2016). Factor structure of suggestibility revisited: New evidence for direct and indirect suggestibility. *Current Issues in Personality Psychology*, 4(2), 87–96. <https://doi.org/10.5114/cipp.2016.60249>
- Polczyk, R., & Pasek, T. (2006). Types of suggestibility: Relationships among compliance, indirect, and direct suggestibility. *The International Journal of Clinical and Experimental Hypnosis*, 54(4), 392–415. <https://doi.org/10.1080/00207140600856764>
- R Core Team. (2023). R: A language and environment for statistical computing. In (Version 4.3.1 (2023-06-16)) <https://www.R-project.org/>.
- Rasch, B., & Cordi, M. J. (2023). The influence of experience and modality of presentation (online vs. offline) on hypnotizability. *Frontiers in Psychology*, 14, 1,293,070. <https://doi.org/10.3389/fpsyg.2023.1293070>
- Rassin, E. (2022). Suggested false memories of a non-existent film: Forensically relevant individual differences in the crashing memories paradigm. *Memory*, 30(9), 1205–1211. <https://doi.org/10.1080/09658211.2022.2085750>
- Ridley, A. M., & Gudjonsson, G. H. (2013). Suggestibility and individual differences: Psychosocial and memory measures. In *Suggestibility in legal contexts: Psychological research and forensic implications* (pp. 85–106). Wiley Blackwell.
- Sapkota, R. P., Brunet, A., & Kirmayer, L. J. (2020). Characteristics of adolescents affected by mass psychogenic illness outbreaks in schools in nepal: A case-control study. *Frontiers in Psychiatry*, 11(493), 094. <https://doi.org/10.3389/fpsyg.2020.493094>
- Stein, M. V., Holt, R., Wieder, L., & Terhune, D. B. (2023). Responsiveness to direct verbal suggestions and dissociation independently predict symptoms associated with environmental factors. *Psychopathology*, 56(4), 324–328.
- Stein, M. V., Heller, M., Chapman, S., Rubin, G. J., & Terhune, D. B. (2025). Trait responsiveness to verbal suggestions predicts nocebo responding: A meta-analysis. *British Journal of Health Psychology*, 30(1), Article e12774. <https://doi.org/10.1111/bjhp.12774>
- Szigeti, B., Weiss, B., Rosas, F. E., Erritzoe, D., Nutt, D., & Carhart-Harris, R. (2024). Assessing expectancy and suggestibility in a trial of escitalopram v. Psilocybin for

- depression. *Psychological Medicine*, 1–8. <https://doi.org/10.1017/S0033291723003653>
- Tasso, A. F., Pérez, N. A., Moore, M., Griffo, R., & Nash, M. R. (2020). Hypnotic responsiveness and nonhypnotic suggestibility: Disparate, similar, or the same? *International Journal of Clinical and Experimental Hypnosis*, 68(1), 38–67. <https://doi.org/10.1080/00207144.2020.1685330>
- The jamovi project. (2023). Jamovi (version 2.3) [computer software]. Retrieved from <https://www.Jamovi.Org>.
- Thompson, T., Terhune, D. B., Oram, C., Sharangparni, J., Rouf, R., Solmi, M., ... Stubbs, B. (2019). The effectiveness of hypnosis for pain relief: A systematic review and meta-analysis of 85 controlled experimental trials. *Neuroscience and Biobehavioral Reviews*, 99, 298–310. <https://doi.org/10.1016/j.neubiorev.2019.02.013>
- Varga, K. (2011). *Beyond the words: Communication and suggestion in medical practice*. Nova Science Publishers.
- Viechtbauer, W. (2010). Conducting meta-analyses in r with the metafor package. *Journal of Statistical Software*, 36(3), 1–48. <https://doi.org/10.18637/jss.v036.i03>
- Vissia, E. M., Giesen, M. E., Chalavi, S., Nijenhuis, E. R., Draijer, N., Brand, B. L., & Reinders, A. A. (2016). Is it trauma- or fantasy-based? Comparing dissociative identity disorder, post-traumatic stress disorder, simulators, and controls. *Acta Psychiatrica Scandinavica*, 134(2), 111–128. <https://doi.org/10.1111/acps.12590>
- Wieder, L., & Terhune, D. B. (2019). Trauma and anxious attachment influence the relationship between suggestibility and dissociation: A moderated-moderation analysis. *Cognitive Neuropsychiatry*, 24(3), 191–207. <https://doi.org/10.1080/13546805.2019.1606705>
- Wieder, L., Brown, R., Thompson, T., & Terhune, D. B. (2021). Suggestibility in functional neurological disorder: A meta-analysis. *Journal of Neurology, Neurosurgery and Psychiatry*, 92(2), 150–157. <https://doi.org/10.1136/jnnp-2020-323,706>
- Wieder, L., Brown, R. J., Thompson, T., & Terhune, D. B. (2022). Hypnotic suggestibility in dissociative and related disorders: A meta-analysis. *Neuroscience and Biobehavioral Reviews*, 139(104), 751. <https://doi.org/10.1016/j.neubiorev.2022.104751>
- Wieder, L., Brown, R. J., & Terhune, D. B. (2023). Revisiting the role of verbal suggestion in dissociative psychopathology. *Acta Psychiatrica Scandinavica*. <https://doi.org/10.1111/acps.13597>
- Woody, E. Z., & Barnier, A. J. (2008). Hypnosis scales for the twenty-first century: What do we need and how should we use them?. In *The oxford handbook of hypnosis: Theory, research, and practice* (pp. 255–280). Oxford University 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. <https://doi.org/10.1037/1040-3590.17.2.200>
- Yard, S. S., DuHamel, K. N., & Galynker, I. I. (2008). Hypnotizability as a potential risk factor for posttraumatic stress: A review of quantitative studies. *Intl. Journal of Clinical and Experimental Hypnosis*, 56(3), 334–356.
- Zaragoza, M. S., & Lane, S. M. (1994). Source misattributions and the suggestibility of eyewitness memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4), 934.