The impact of hypnotic suggestibility in clinical care settings

Guy H. Montgomery[1], **Julie B. Schnur**[1], and **Daniel David**[2]

[1]Integrative Behavioral Medicine Program, Department of Oncological Sciences, Mount Sinai School of Medicine, New York, New York, United States of America

[2]Department of Clinical Psychology and Psychotherapy, Babes-Bolyai University, Cluj-Napoca, Romania

Abstract

Hypnotic suggestibility has been described as a powerful predictor of outcomes associated with hypnotic interventions. However, there have been no systematic approaches to quantifying this effect across the literature. The present meta-analysis evaluates the magnitude of the effect of hypnotic suggestibility on hypnotic outcomes in clinical settings. PsycINFO and PubMed were searched from their inception through July 2009. Thirty-four effects from ten studies and 283 participants are reported. Results revealed a statistically significant overall effect size in the small to medium range ($r = 0.24$; 95% Confidence Interval = −0.28 to 0.75), indicating that greater hypnotic suggestibility led to greater effects of hypnosis interventions. Hypnotic suggestibility accounted for 6% of the variance in outcomes. Smaller sample size studies, use of the SHCS, and pediatric samples tended to result in larger effect sizes. Results question the usefulness of assessing hypnotic suggestibility in clinical contexts.

Hypnotic suggestibility, also referred to as hypnotizability and hypnotic susceptibility, refers to the degree to which a participant responds to hypnotic suggestions. Hypnotic suggestibility is typically recognized as a stable individual difference characteristic, as demonstrated by robust test-retest correlations for periods of up to 25 years (Piccione, Hilgard, & Zimbardo, 1989). Classic data from Hilgard (1965) support the position that hypnotic suggestibility roughly conforms to a bell-shaped distribution in the population, with fewer people falling at the extremes and most people in the middle. Cross-cultural studies have replicated these findings in Australian (Sheehan & McConkey, 1979), Canadian (Laurence & Perry, 1982), Danish (Zachariae, Sommerlund, & Molay, 1996), Finnish (Kallio & Ihamuotila, 1999), German (Bongartz, 1985), Italian (DePascalis, Russo, & Marucci, 2000), Romanian (David, Montgomery, & Holdevici, 2003), and Spanish (Lamas, delValle-Inclan, Blanco, & Diaz, 1989) samples.

The construct of hypnotic suggestibility has been conceptualized as a moderator. That is, degree or level of hypnotic suggestibility is proposed to be associated with degree or level of responsiveness to hypnotic suggestions. The higher that an individual scores on a measure of hypnotic suggestibility, the more responsive one would predict that person to be to hypnotic suggestions.

Support for hypnotic suggestibility as a predictor of hypnotic outcomes is strong in the experimental literature. For example, studies have demonstrated that participants scoring higher in hypnotic suggestibility tend to have greater experimental pain relief (Milling, Shores, Coursen, Menario, & Farris, 2007) and color pattern recognition following hypnosis

Corresponding Author: Dr. Guy H. Montgomery, Department of Oncological Sciences, Box 1130, Mount Sinai School of Medicine, 1425 Madison Avenue, New York, NY 10029-6574. Phone: 212-659-5561; Fax: 212-849-2566; guy.montgomery@mssm.edu.

(Kosslyn, Thompson, Costantini-Ferrando, Alpert, & et al., 2000). These studies have led some to conclude that hypnotic suggestibility plays a central role as a predictor of hypnotic responsiveness across a variety of settings (Barabasz & Perez, 2007).

However, despite the experimental evidence, the utility of assessing levels of hypnotic suggestibility in clinical settings, where time is often short and patient burden often high, remains in question for at least three reasons: 1) Assessments of hypnotic suggestibility can often be lengthy (Spanos, Radtke, Hodgins, Stam, & Bertrand, 1983; Weitzenhoffer & Hilgard, 1962) and can take longer to administer than brief efficacious hypnotic interventions themselves (Montgomery et al., 2007); 2) Meta-analyses have supported the position that the vast majority of patients undergoing medical procedures benefit from hypnotic interventions (Montgomery, David, Winkel, Silverstein, & Bovbjerg, 2002; Schnur, Kafer, Marcus, & Montgomery, 2008); and, 3) The percentage of patients who benefit from hypnotic interventions in clinical settings far exceeds the percentage of individuals scoring in the high range of hypnotic suggestibility scales (Hilgard, 1965; Montgomery et al., 2002).

The primary goal of this meta-analysis is to assess the influence of hypnotic suggestibility on hypnosis outcomes in the context of clinical settings. Should the influence of hypnotic suggestibility on responses to hypnotic interventions be large (based on Cohen's criteria (1992)), the argument to screen patients for levels of hypnotic suggestibility would be strongly supported. Clinicians could benefit from assessing patients' levels of hypnotic suggestibility in order to best determine treatment course. However, should the effect fall into the small or medium range, the case for screening patients prior to hypnotic interventions would be less well-supported. Additional patient burden, staff time, and costs associated with assessing patients' levels of hypnotic suggestibility might not be justified. A secondary goal of this meta-analysis is to identify potential factors which might affect the strength or direction of the relationship between hypnotic suggestibility and hypnotic outcomes in clinical settings. In particular, we will examine the influence of sample size, the scale used to assess hypnotic suggestibility, the nature of the outcome, as well as the impact of pediatric versus adult samples.

Methods

Search strategy

Two electronic databases, PsycInfo and PubMed were searched from their respective inceptions through the end of July 2009.

For PsycInfo, the major search terms were hypnosis and ("hypnotic suggestibility" or "hypnotizability" or "hypnotic susceptibility") [mp=title, abstract, heading word, table of contents, key concepts]. The limitations placed on the search were: peer reviewed journal, human, English language, abstracts, quantitative study or treatment outcome/randomized clinical trial. This approach yielded a total of 95 abstracts.

For PubMed, the search terms were: ("hypnosis" [MeSH Terms] OR "hypnosis" [All Fields]) AND (hypnotizability [All Fields] OR "hypnotic suggestibility" [All Fields] OR "hypnotic susceptibility" [All Fields]) AND (has abstract [text] AND "humans" [MeSH Terms] AND Randomized Controlled Trial [ptyp] AND English [lang]). This approach yielded a total of 63 abstracts.

Selection Strategy

The abstracts of all articles identified by electronic searches (158 in total) were carefully screened by the authors in the study to determine if the abstracts meet the following

inclusion criteria: (a) published in a peer reviewed journal; (b) full abstract available online; (c) randomized; (d) written in English; (e) included at least one control condition; (f) hypnosis was listed as at least one of the intervention conditions; (g) hypnosis was being used in a clinical population (defined as a medical, dental, or mental health treatment setting); (h) some empirical, validated measure of hypnotic suggestibility; (i) the study had sufficient data to calculate an effect size; and (j) the study was not a duplicate (i.e., if an article was cited in both PubMed and PsycInfo, it was only used once).

Subsequent to abstract review, 18 manuscripts were obtained and read in full. A standardized form assessing the inclusion criteria was completed for each paper. The co-authors then discussed the 18 manuscripts. Based on consensus review by all authors, 10 of the 18 papers were included in the meta-analysis. The reasons for exclusion were as follows: 1 paper was a case study, in 1 study a portion of the sample was included in another study already included in the meta-analysis (overlapping data), in 3 studies, participants were not randomly assigned to study group, 1 study included an ineligible population, 1 study had a lack of information regarding hypnosis, and 1 study did not have sufficient data to calculate an effect size (see Study Flow Chart, Figure 1).

Data abstraction and study characteristics

For each of the 10 papers which were accepted for inclusion into the meta-analysis, relevant data were abstracted using a standardized worksheet. Each paper was abstracted independently by two of the co-authors. Any discrepancies were discussed among all of the authors with reference to the original manuscript until consensus was reached. Specific data collected included: 1) results describing the correlation between hypnotic suggestibility and outcomes within hypnosis treatment groups (e.g., r, R^2, test statistics); 2) the type of clinical setting (i.e., medical/dental vs. mental health); 3) study design; 4) sample size; 5) hypnotic suggestibility scale used; 6) population characteristics (i.e., pediatric vs. adult).

Quantitative data synthesis

For every study, we calculated an effect size (r) for each association between the predictor of interest (i.e., hypnotic suggestibility) and a hypnosis intervention outcome in hypnosis intervention groups. If a given study had more than one such association, then an r was computed for each individual association.

The aggregate effect sizes reported below were calculated using the meta-analysis program by Schwarzer (2008) (using the "*Effect Sizes r*" utility). The program produced effect sizes weighted by sample size using a random effects model. The random effects model does not assume that the set of effect sizes is homogeneous, and we expected the set of effects to display heterogeneity. Additionally, it has been recommended to use the random effects model, as this model is more generalizable to the broader population of studies (Rosenthal & DiMatteo, 2001).

The results yielded by the program also address two important concerns: the homogeneity of the set of effects, and the "file-drawer" problem. Homogeneity of the set of effects is critical, as it indicates the trustworthiness of the overall effect size generated by the meta-analysis (Schwarzer, 2008). To examine homogeneity of the effect sizes, the percent of variance attributable to sampling error versus the percent of variance attributable to systematic factors was examined (Q statistic). The Q statistic indicates whether the variability present in the group of effect sizes is significantly greater than chance, and thus suggestive of the presence of potential moderators. Moderator analyses following a finding of heterogeneity of effect sizes (a significant Q value), were conducted by analyses of

variance in SAS 9.2 (SAS Institute Inc, 2009). To examine publication bias, otherwise known as the "file-drawer problem," we used Orwin's (1983) method.

Results

Included studies were published between 1991 and July 2009. Sample sizes in the hypnosis groups ranged from 15 to 77.

The ten studies meeting the inclusion criteria yielded 34 effect sizes. Effect sizes were based on 283 participants within the studies. Table 1 presents the chief characteristics and effect sizes for study.

The aggregated effect sizes demonstrated that hypnotic suggestibility was significantly associated with hypnotic intervention outcomes in clinical settings ($z = 7.56$; $p < .001$). The mean effect ($r = 0.24$, 95% Confidence Interval = −0.28 to 0.75) using a random effects model was in the positive direction (greater hypnotic suggestibility was associated with greater effectiveness of hypnosis interventions), and in the small to medium range (Cohen, 1992). Across studies, the effects of hypnotic suggestibility accounted for 6% of the variance in outcomes.

Effect sizes used for this analysis were all drawn from published, peer-reviewed journals, which may raise concerns about publication bias. To address this possible concern, the number of studies with an effect size of zero required to decrease the overall mean effect size to small (based on Orwin, 1983) was calculated. Results indicated that 47 studies with effect sizes of zero would be needed to reduce the small to medium effect size of 0.24 found here to a small effect size of $r = 0.10$.

Results of homogeneity tests indicated that the sample of effect sizes was heterogeneous ($Q = 142.03$, $df = 33$, $p < 0.001$). These results suggest that moderators are likely to be present.

First, we explored whether sample size was correlated with effect size. Results revealed a significant, negative correlation ($r = −0.67$, $p < .001$). That is, larger studies tended to produce smaller associations between hypnotic suggestibility and outcomes of hypnotic interventions. These results were consistent with the funnel plot (see Figure 2).

Second, we explored whether the particular scale used to assess hypnotic suggestibility affected the association. Effect sizes were divided into those that used the Hypnotic Induction Profile (HIP) (Spiegel, 1977; Stern, Spiegel, & Nee, 1979) and those that used the Stanford Hypnotic Clinical Scale (SHCS) (Morgan & Hilgard, 1975). Mean effect size for those effects using the HIP ($r = 0.12$; $SD = 0.26$; based on eight effects) was significantly lower than for those studies which used the SHCS ($r = 0.47$; $SD = 0.28$; based on 26 effects) [$F(1, 32) = 9.97$; $p < .004$].

Third, we examined whether the nature of the outcome variable was related to effect sizes. Outcomes were divided into five categories, pain, anxiety, patient distress, behavioral outcomes and depression. Descriptive statistics are presented in Table 2. ANOVA output indicates that the moderating influence of outcome category on the association between hypnotic suggestibility and the effectiveness of the hypnotic intervention was not significant [$F(3, 25) = 2.84$; $p < .08$]. Depression as an outcome was not included in these analyses as there was only a single effect ($r = −0.15$).

Fourth, we considered the influence of age on the results. In three of the studies (Liossi & Hatira, 1999; Liossi & Hatira, 2003; Liossi, White, & Hatira, 2006), the samples consisted of pediatric patients. With these studies removed from the data set, hypnotic suggestibility

remained significantly associated with hypnotic intervention outcomes in clinical settings with adults ($z = 3.27$; $p < .001$). However, the effect shrunk to the small range ($r = 0.12$, 95% Confidence Interval = −0.19 to 0.42). The average weighted effect size for the pediatric samples was also significant ($z = 10.82$; $p < .001$), but was in the large range ($r = 0.67$, 95% Confidence Interval = 0.43 to 0.91). Further analyses revealed that effect sizes were significantly larger in pediatric samples than in adult samples [$F(1, 32) = 35.42$; $p < .001$].

Fifth, three of the studies were conducted with mental health patients (Moene, Spinhoven, Hoogduin, & Van Dyck, 2003; Moene, Spinhoven, Hoogduin, & Van Dyck, 2002; Van Dyck & Spinhoven, 1997), and the remainder were with medical patients. There were no differences in effect sizes based on this factor [$F(1, 32) = 0.73$; $p < .40$].

Discussion

The present meta-analysis represents a review of the studies to date which have included assessments of the moderating influence of hypnotic suggestibility on the effects of hypnotic interventions in clinical (medical, dental, or mental health) settings. The results revealed a small to medium overall weighted effect size ($r = 0.24$) (Cohen, 1992), and raise the question of the overall utility of assessing hypnotic suggestibility in clinical settings.

The utility of any screening instrument in health care settings lies in its ability to predict clinical outcomes or to identify patient needs. With regard to the application of hypnosis interventions, it has been argued that hypnotic suggestibility is a core construct in the clinic as well as in the laboratory, and is a powerful predictor of outcome. The present data and practical considerations temper these assumptions. First, in aggregate, hypnotic suggestibility only accounted for 6% of the variance in responsiveness to hypnosis interventions in clinical settings. Though statistically significant, this effect is relatively small and suggests little clinical importance. It seems that the assessment of hypnotic suggestibility prior to offering a hypnosis intervention will predict which clinical patients will do slightly better, and not much more. Second, several assessment approaches to hypnotic suggestibility are often lengthy (e.g., SHCS (Morgan et al., 1975), CURSS (Spanos et al., 1983)), even taking more time to administer than hypnosis itself (e.g., (Montgomery et al., 2007). Briefer measures, like the HIP (Spiegel, 1977) could be used, but the present data suggest that the predictive utility of the briefer HIP in this context is significantly less than that of other lengthier measures, negating the time savings advantage. Perhaps, a better briefer measure of hypnotic suggestibility is needed for clinical settings. Third, with regard to the patient's experience, assessment of hypnotic suggestibility could be counterproductive. The vast majority of hypnotic suggestibility measures include very difficult items (e.g., amnesia, hallucinations), ensuring that most patients will have failure experiences. Such failures are likely to negatively affect expectancies for positive clinical outcomes, which according to theory (Kirsch, 1999), meta-analysis (Sohl, Schnur, & Montgomery, 2009) and empirical studies (Montgomery et al., 2010; Montgomery, Schnur, Erblich, Diefenbach, & Bovbjerg, 2010) contribute to hypnosis intervention effects. In essence, patients may respond better to hypnosis interventions, and receive greater clinical benefit, if they do *not* go through the hypnotic suggestibility assessment process. In our clinical experience, breast cancer patients who reported having benefited from our hypnosis interventions are often puzzled and disappointed when they do not "score well" on hypnotic suggestibility measures. It can take a fair amount of time to reassure them that their "hypnotic suggestibility" score does not in any way diminish their prior response to hypnosis, nor does it mean they will not benefit in the future. Fourth, clinical effect sizes for hypnosis interventions in clinical settings have consistently been in the medium to large range (Montgomery et al., 2002; Schnur et al., 2008). Given that the majority of patients (up to 89%) (Montgomery et al., 2002) benefit from hypnosis interventions in clinical settings,

the most efficient and least burdensome clinical approach may be to simply administer the intervention.

Examination of potential moderators of the effects of hypnotic suggestibility in clinical settings yielded several interesting findings. First, effects sizes were significantly and negatively correlated with sample sizes. Larger effect sizes tended to be found in smaller samples, and smaller effect sizes tended to be reported in larger samples. These results raise the question of whether effect sizes describing the influence of hypnotic suggestibility were biased by (statistically speaking) small sample size reports. For example, the strongest correlations ($r = .81$) were reported in sample sizes of only 20 participants (Liossi et al., 2003).

Another potential explanation of the results regarding the inverse correlation of sample size and effect size is that the studies by Liossi and colleagues, which have both small sample sizes and large effect sizes (Liossi et al., 1999; Liossi et al., 2003; Liossi et al., 2006), involved pediatric patients. Removing these studies from the sample resulted in a smaller overall effect size. It is possible therefore that the effects of hypnotic suggestibility in clinical populations are stronger in pediatric populations than among adults. This area deserves further study.

Effect sizes did not significantly differ based on outcome category (see Table 2). However, the results ($p < .08$) are somewhat provocative, and there may be a tendency to have smaller effects with behavioral outcomes. As they stand, the results are consistent with other reports indicating that effects of hypnosis interventions are equivalent across clinical outcome categories in medical settings (Montgomery et al., 2002). However, differences in the effects of hypnotic suggestibility may indeed vary across outcomes if enough statistical power were available. This is an additional area which deserves greater scientific attention.

An important limitation of the present paper is that the studies with the largest effect sizes were both small sample size studies and involved pediatric patients. It is therefore impossible to determine whether effects are due to sampling error, or due to the difference in the developmental stage of the population. If one were to design a new study testing associations between hypnotic suggestibility and hypnosis outcomes, the present results (an $r = 0.24$) suggest that a sample size of 132 participants would be required to maintain power above the traditional .80 with a two-tailed alpha of .05 (Gorman, Primavera, & Allison, 1995).

A second limitation of the analysis is that given the length of time that hypnosis has been used in clinical settings (over a century (Gravitz, 1988)), and the length of time the SCHS and HIP have existed, there are relatively few papers which included assessments of hypnotic suggestibility as a potential moderator. This lack of studies could reflect the limited real-world practicality of the assessments overall, or simply a clinical focus on improving outcomes.

In conclusion, the present data support the position that there is a small but significant association between hypnotic suggestibility and hypnosis intervention outcomes in clinical settings. However, given that hypnotic suggestibility only accounted for 6% of the variance in outcomes, from the clinical perspective, there appears to be an imbalance between the value of the information gained and the burden of administering hypnotic suggestibility assessments in these settings. It would be of interest to determine if these results are consistent across a variety of clinical populations, ages, and outcomes. Based on the literature, it appears that clinical success or failure of a hypnotic intervention does not hinge on level of hypnotic suggestibility.

Acknowledgments

References

Barabasz A, Perez N. Salient findings: hypnotizability as core construct and the clinical utility of hypnosis. International Journal of Clinical and Experimental Hypnosis. 2007; 55:372–379. [PubMed: 17558724]

Bongartz W. German norms for the Harvard Group Scale of Hypnotic Susceptibility, Form A. International Journal of Clinical and Experimental Hypnosis. 1985; 33:131–140. [PubMed: 4018922]

Cohen JA. A power primer. Psychological Bulletin. 1992; 112:155–159. [PubMed: 19565683]

David D, Montgomery G, Holdevici I. Romanian norms for the Harvard Group Scale of Hypnotic Susceptibility, Form A. International Journal of Clinical and Experimental Hypnosis. 2003; 51:66–76. [PubMed: 12825919]

DePascalis V, Russo P, Marucci FS. Italian norms for the Harvard Group Scale of Hypnotic Susceptibility, Form A. International Journal of Clinical and Experimental Hypnosis. 2000; 48:44–55. [PubMed: 10641432]

Gorman BS, Primavera LH, Allison DB. Powpal - A Program for Estimating Effect Sizes, Statistical Power, and Sample Sizes. Educational and Psychological Measurement. 1995; 55:773–776.

Gravitz MA. Early uses of hypnosis as surgical anesthesia. American Journal of Clinical Hypnosis. 1988; 30:201–208. [PubMed: 3278584]

Hilgard, ER. Hypnotic Susceptibility. New York: Brace and Wared; 1965.

Kallio SI, Ihamuotila MJ. Finnish norms of the Harvard Group Scale of Hypnotic Susceptibility, Form A. International Journal of Clinical and Experimental Hypnosis. 1999; 47:227–235. [PubMed: 10616256]

Kirsch, I. How expectancies shape experience. 1. Washington, DC: American Psychological Association; 1999.

Kosslyn SM, Thompson WL, Costantini-Ferrando MF, Alpert NM, et al. Hypnotic visual illusion alters color processing in the brain. American Journal of Psychiatry. 2000; 157:1279–1284. [PubMed: 10910791]

Lamas JR, delValle-Inclan F, Blanco MJ, Diaz AA. Spanish norms for the Harvard Group Scale of Hypnotic Susceptibility, Form A. International Journal of Clinical and Experimental Hypnosis. 1989; 37:264–273. [PubMed: 2753576]

Lang EV, Joyce JS, Spiegel D, Hamilton D, Lee KK. Self-hypnotic relaxation during interventional radiological procedures: effects on pain perception and intravenous drug use. International Journal of Clinical and Experimental Hypnosis. 1996; 44:106–119. [PubMed: 8871338]

Laurence JR, Perry C. Montreal norms for the Harvard Group Scale of Hypnotic Susceptibility, Form A. International Journal of Clinical and Experimental Hypnosis. 1982; 30:167–176. [PubMed: 7085138]

Liossi C, Hatira P. Clinical hypnosis versus cognitive behavioral training for pain management with pediatric cancer patients undergoing bone marrow aspirations. International Journal of Clinical and Experimental Hypnosis. 1999; 47:104–116. [PubMed: 10208073]

Liossi C, Hatira P. Clinical hypnosis in the alleviation of procedure-related pain in pediatric oncology patients. International Journal of Clinical and Experimental Hypnosis. 2003; 51:4–28. [PubMed: 12825916]

Liossi C, White P, Hatira P. Randomized clinical trial of local anesthetic versus a combination of local anesthetic with self-hypnosis in the management of pediatric procedure-related pain. Health Psychology. 2006; 25:307–315. [PubMed: 16719602]

Lutgendorf SK, Lang EV, Berbaum KS, Russell D, Berbaum ML, Logan H, et al. Effects of age on responsiveness to adjunct hypnotic analgesia during invasive medical procedures. Psychosomatic Medicine. 2007; 69:191–199. [PubMed: 17289823]

Milling LS, Shores JS, Coursen EL, Menario DJ, Farris CD. Response expectancies, treatment credibility, and hypnotic suggestibility: mediator and moderator effects in hypnotic and cognitive-behavioral pain interventions. Annals of Behavioral Medicine. 2007; 33:167–178. [PubMed: 17447869]

Moene FC, Spinhoven P, Hoogduin KA, van DR. A randomized controlled clinical trial of a hypnosis-based treatment for patients with conversion disorder, motor type. International Journal of Clinical and Experimental Hypnosis. 2003; 51:29–50. [PubMed: 12825917]

Moene FC, Spinhoven P, Hoogduin KA, van DR. A randomised controlled clinical trial on the additional effect of hypnosis in a comprehensive treatment programme for in-patients with conversion disorder of the motor type. Psychotherapy and Psychosomatics. 2002; 71:66–76. [PubMed: 11844942]

Montgomery GH, Bovbjerg DH, Schnur JB, David D, Goldfarb A, Weltz CR, et al. A randomized clinical trial of a brief hypnosis intervention to control side effects in breast surgery patients. Journal of the National Cancer Institute. 2007; 99:1304–1312. [PubMed: 17728216]

Montgomery GH, David D, Winkel G, Silverstein JH, Bovbjerg DH. The effectiveness of adjunctive hypnosis with surgical patients: A meta-analysis. Anesthesia and Analgesia. 2002; 94:1639–1645. [PubMed: 12032044]

Montgomery GH, Hallquist MN, Schnur JB, David D, Silverstein JH, Bovbjerg DH. Mediators of a brief hypnosis intervention to control side effects in breast surgery patients: Response expectancies and emotional distress. Journal of Consulting and Clinical Psychology. 2010; 78:80–88. [PubMed: 20099953]

Montgomery GH, Schnur JB, Erblich J, Diefenbach MA, Bovbjerg DH. Presurgery psychological factors predict pain, nausea, and fatigue one week after breast cancer surgery. Journal of Pain and Symptom Management. 2010; 39:1043–1052. [PubMed: 20538186]

Morgan, AH.; Hilgard, JR. The Stanford Hypnotic Clinical Scale. In: Hilgard, ER.; Hilgard, JR., editors. Hypnosis in the Relief of Pain. Los Altos, CA: William Kaufman; 1975.

Orwin RG. A fail safe N for effect size in meta-analysis. Journal for Educational Statistics. 1983; 8:157–159.

Piccione C, Hilgard ER, Zimbardo PG. On the degree of stability of measured hypnotizability over a 25-year period. Journal of Personality and Social Psychology. 1989; 56:289–295. [PubMed: 2926631]

Rosenthal R, DiMatteo MR. Meta-analysis: Recent developments in quantitative methods for literature reviews. Annual Review of Psychology. 2001; 52:59–82.

SAS Institute Inc. SAS (r) Proprietary Software 9.2 [Computer software]. Cary, N.C: SAS Institute Inc; 2009.

Schnur JB, Kafer I, Marcus C, Montgomery GH. Hypnosis to manage distress related to medical procedures: A meta-analysis. Contemporary Hypnosis. 2008; 25:114–128. [PubMed: 19746190]

Schwarzer, R. Statistics Software for Meta-Analysis [Computer software]. 2008.

Sheehan PW, McConkey KM. Australian norms for the Harvard Group Scale of Hypnotic Susceptibility form A. International Journal of Clinical and Experimental Hypnosis. 1979; 27:294–304. [PubMed: 541138]

Sohl S, Schnur J, Montgomery G. A Meta-Analysis of the Relationship between Response Expectancies and Cancer-Treatment-Related Side Effects. Journal of Pain and Symptom Management. 2009; 38:775–784. [PubMed: 19775863]

Spanos NP, Radtke HL, Hodgins DC, Stam HJ, Bertrand LD. The Carleton University responsiveness to suggestion scale: normative data and psychometric properties. Psychological Reports. 1983; 53:523–535. [PubMed: 6647700]

Spiegel H. The Hypnotic Induction Profile HIP: A review of its development. Annals of The New York Academy of Sciences. 1977; 296:129–142. [PubMed: 279239]

Stern DB, Spiegel H, Nee JCM. The hypnotic induction profile: Normative observation, reliability and validity. The American Journal of Clinical Hypnosis. 1979; 21:109–133. [PubMed: 747161]

ter Kuile MM, Spinhoven P, Linsen ACG, Zitman FG, Van Dyck R, Rooijmans HGM. Autogenic training and cognitive self-hypnosis for the treatment of recurrent headaches in three different subject groups. Pain. 1994; 58:331–339. [PubMed: 7838582]

Van Dyck R, Zitman FG, Linssen ACG, Spinhoven P. Autogenic training and future oriented hypnotic imagery in the treatment of tension headache: outcome and process. The International Journal of Clinical and Experimental Hypnosis. 1991; 39:6–23. [PubMed: 1705921]

Van, Dyck R.; Spinhoven, P. Does preference for type of treatment matter? A study of exposure in vivo with or without hypnosis in the treatment of panic disorder with agoraphobia. Behavior Modification. 1997; 21:172–186. [PubMed: 9086865]

Weitzenhoffer, AM.; Hilgard, ER. Stanford Hypnotic Susceptibility Scale - Form C. Palo Alto, CA: Consulting Psychologists Press, Inc; 1962.

Zachariae R, Sommerlund B, Molay F. Danish norms of the Harvard Group Scale of Hypnotic Susceptibility, Form A. International Journal of Clinical and Experimental Hypnosis. 1996; 44:140–152. [PubMed: 8871340]
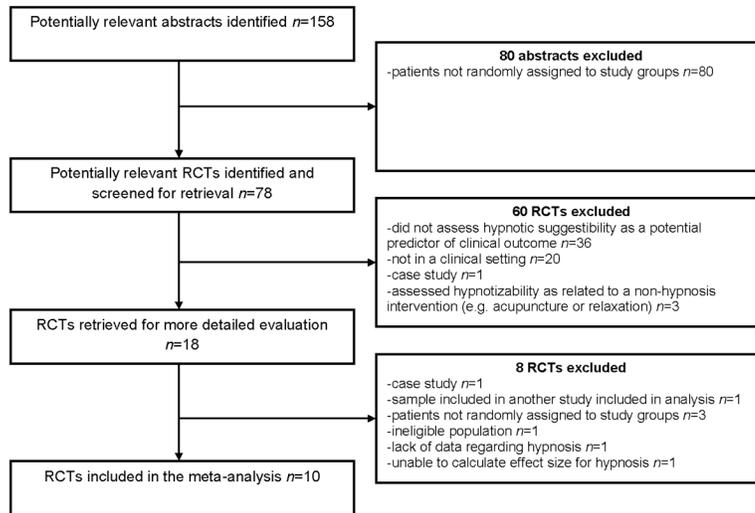
```
┌─────────────────────────────────────┐
│ Potentially relevant abstracts      │
│ identified n=158                    │
└─────────────────────────────────────┘
              │          ┌──────────────────────────────────────────┐
              │          │ 80 abstracts excluded                    │
              ├─────────▶│ -patients not randomly assigned to study │
              │          │  groups n=80                             │
              ▼          └──────────────────────────────────────────┘
┌─────────────────────────────────────┐
│ Potentially relevant RCTs identified │
│ and screened for retrieval n=78     │
└─────────────────────────────────────┘
              │          ┌──────────────────────────────────────────┐
              │          │ 60 RCTs excluded                         │
              │          │ -did not assess hypnotic suggestibility  │
              │          │  as a potential predictor of clinical    │
              ├─────────▶│  outcome n=36                            │
              │          │ -not in a clinical setting n=20          │
              │          │ -case study n=1                          │
              │          │ -assessed hypnotizability as related to  │
              │          │  a non-hypnosis intervention (e.g.       │
              ▼          │  acupuncture or relaxation) n=3          │
┌─────────────────────────────────────┐└──────────────────────────────────────────┘
│ RCTs retrieved for more detailed    │
│ evaluation n=18                     │
└─────────────────────────────────────┘
              │          ┌──────────────────────────────────────────┐
              │          │ 8 RCTs excluded                          │
              │          │ -case study n=1                          │
              │          │ -sample included in another study        │
              │          │  included in analysis n=1                │
              ├─────────▶│ -patients not randomly assigned to study │
              │          │  groups n=3                              │
              │          │ -ineligible population n=1               │
              │          │ -lack of data regarding hypnosis n=1     │
              ▼          │ -unable to calculate effect size for     │
┌─────────────────────────────────────┐│  hypnosis n=1                            │
│ RCTs included in the meta-analysis  │└──────────────────────────────────────────┘
│ n=10                                │
└─────────────────────────────────────┘
```
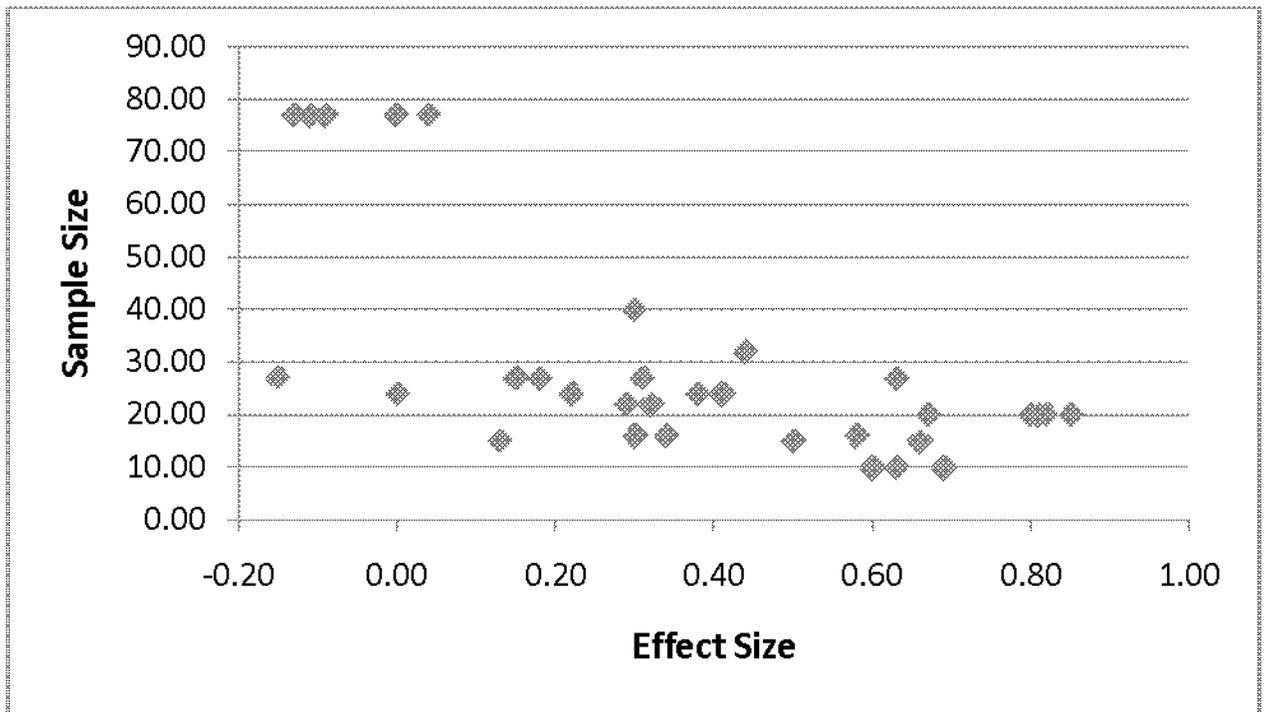
Figure 1.
Study flow chart.

Figure 2.
Effect size by study sample size (funnel plot).

Table 1

Study characteristics and mean effect sizes.

Study	Clinical Care Context	Design	Hypnosis Group *n*	Total *n*	Hypnotic Suggestibility Scale	Child vs. Adult	*r* (mean)[†]
Liossi et al. (2006)	Pediatric procedure, related pain	EMLA [*] vs. EMLA + hypnosis vs. EMLA + attention	15	45	SHCS	Child	0.49
Liossi & Hatira (2003)	Pediatric oncology	Direct hypnosis + standard vs. Indirect hypnosis + standard vs. Attention control + standard vs. Standard	20 20	80	SHCS	Child	0.76 0.82
Van Dyck & Spinhoven (1997)	Panic disorder with agoraphobia	Exposure vs. Exposure + hypnosis	32	64	SHCS	Adult	0.44
Lang et al. (1996)	Radiological Procedures	Self-hypnosis relaxation vs. Control	16	30	HIP	Adult	0.41
Van Dyck et al. (1991)	Headache	Autogenic training vs. Future oriented hypnotic imagery	27	55	SHCS	Adult	0.22
Moene et al. (2003)	Conversion disorder, motor type	Hypnosis vs. Waiting list	20	44	SHCS	Adult	0.30
ter Kuile et al. (1994)	Headache	Cognitive self-hypnosis vs. Autogenic training vs. Waiting list	40	157	SHCS	Adult	0.30
Lutgendorf et al. (2007)	Invasive medical procedures	Hypnosis vs. Attention vs. Standard	77	241	HIP	Adult	−0.05
Liossi & Hatira (1999)	Bone marrow aspirations	Hypnosis vs. CBT vs. No intervention	10	30	SHCS	Children	0.64
Moene et al. (2002)	Conversion disorder, motor type	Group therapy vs. Group therapy + hypnosis	24	45	SHCS	Adults	0.25

[*]
EMLA: eutectic mixture of local anesthetics;

[†]
r (mean): mean correlation between hypnotic suggestibility and outcomes in hypnosis groups

Table 2

Mean weighted effect size by outcome type.

Outcome Category	# of effects	Mean weighted *r*	95% Confidence Interval
Pain	12	0.38	–0.01 to 0.75
Anxiety	8	0.31	–0.33 to 0.96
Distress	3	0.57	0.12 to 1.01
Behavioral	10	0.07	–0.21 to 0.35