Suggestion overrides the Stroop effect in highly hypnotizable individuals

Amir Raz *, Miguel Moreno-Íniguez, Laura Martin, Hongtu Zhu

Columbia University College of Physicians and Surgeons and New York State Psychiatric Institute, MRI Unit in the Department of Psychiatry, Division of Child and Adolescent Psychiatry, 1051 Riverside Drive, Unit 74, New York, NY 10032, USA

Received 11 July 2005
Available online 9 June 2006

Abstract

Cognitive scientists distinguish between automatic and controlled mental processes. Automatic processes are either innately involuntary or become automatized through extensive practice. For example, reading words is a purportedly automatic process for proficient readers and the Stroop effect is consequently considered the “gold standard” of automated performance. Although the question of whether it is possible to regain control over an automatic process is mostly unasked, we provide compelling data showing that posthypnotic suggestion reduced and even removed Stroop interference in highly hypnotizable individuals. Drawing on a large sample of highly hypnotizable participants, we examined the effects of suggestion on Stroop performance both with and without a posthypnotic suggestion to perceive the input stream as meaningless symbols. We show that suggestion administered to highly hypnotizable persons significantly reduced Stroop interference and derailed a seemingly automatic process.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Automaticity; Stroop; Hypnosis; Attention; Suggestion

1. Introduction

Cognitive psychologists generally agree that mental processes come in two varieties: controlled and automatic (Shiffrin & Schneider, 1977). Some processes are thought to be innately automatic; others become automatic through practice (Spelke, Hirst, & Neisser, 1976). General accounts posit that once automatized, these processes are initiated unintentionally, effortlessly, even ballistically, and cannot be easily interrupted or prevented. While most researchers who are interested in automatic processes focus on the nature of automaticity and how a process becomes automatic, they have given little or no attention to whether one can regain control over a process that had been automatized. Despite its theoretical and clinical importance, the latter question is not only unanswered, but mostly unasked. The present study examined whether we can “unring” the proverbial bell (i.e., regain control over an ostensibly involuntary process). We report how a posthypnotic suggestion...
administered to highly hypnotizable individuals can override a process that has become automatized through practice (i.e., reading), and discuss the theoretical and applied implications of our findings.

Reading words is considered an automatic process; a proficient reader cannot withhold accessing word meaning despite explicit instructions to attend only to the ink color. The Stroop task provides evidence for the automaticity of reading: in responding to the ink color of an incongruent color word (e.g., the word “GREEN” inked in red), participants are usually slower and less accurate than in identifying the ink color of either a neutral or congruent word (e.g., “LOT” or “RED” inked in red) (Stroop, 1935). The difference between incongruent and congruent stimuli—the Stroop effect—is one of the most robust and well-studied phenomena in attentional research (MacLeod, 1991; MacLeod & MacDonald, 2000). The standard account posits that words are processed automatically to the semantic level and that the Stroop effect is the “gold standard” of automated performance (MacLeod, 1992).

A few meditative practices claim to achieve “de-automatization” (Dillbeck, 1982) with some sparse evidence of reduced Stroop interference (Alexander, Langer, Newman, Chandler, & Davies, 1989; Wenk-Sormaz, 2005). In addition, a number of studies have independently challenged the robustness of the Stroop effect showing either decrease or elimination of Stroop interference (Besner, 2001; Besner & Stolz, 1999a; Besner & Stolz, 1999b; Besner & Stolz, 1999c; Besner, Stolz, & Boutillier, 1997; Dishon-Berkovits & Algom, 2000; Melara & Algom, 2003; Pansky & Algom, 2002). Although critiqued (Neely & Kahan, 2001), interpretation of these and other findings of either reduction (Long & Prat, 2002) or removal (Kuhl & Kazén, 1999) of Stroop conflict contends that rather than being inevitable, other factors (e.g., attention, memory, and affect) may govern automatic processing. These findings suggest that a seemingly automatic process can be derailed.

Drawing on a larger sample, the present study provides a replication of our previous results and addresses the following question: Is it possible to regain control over a process that has been automatized? To answer this query, we drew on data from multiple experiments showing that a suggestion to experience Stroop words as meaningless symbols can modulate the Stroop effect in highly hypnotizable individuals—about 10–15% of the adult population who can be reliably identified as highly compliant with hypnotic suggestion using standardized scales (Comey & Kirsch, 1999; Kirsch, Capafons, Cardena-Buelna, & Amigo, 1999; Kirsch, Lynn, & Rhue, 1993; Shor & Orne, 1962; Weitzenhoffer & Hilgard, 1962). This approach (Raz, Shapiro, Fan, & Posner, 2002) was reported in a series of studies employing hypnotic suggestions (Egner et al., in press) and consequently extended to a posthypnotic suggestion—a suggestion made during hypnosis indicating that a particular experience or behavior will occur on cue following termination of the hypnotic session (Raz, 2004; Raz, Fan, & Posner, 2005; Raz et al., 2003; Raz et al., 2002)—and was recently replicated in an independent laboratory (Raz, Kirsch, Pollard, & Nitkin-Kaner, 2006). However, since this notion and findings were first reported in the clinical literature (Raz & Shapiro, 2002; Raz et al., 2002), it may have been missed by the cognitive scientists who would be most interested in it. Here, we report behavioral data from a larger sample of highly hypnotizable persons, discuss the findings in light of our recent neuroimaging data, and outline the potential implications of overriding an automatic process for cognitive neuroscience.

2. Method

2.1. Participants

Participants were 49 right-handed proficient readers of English (24 female) aged 20–35 (mean = 27) years. All participants were recruited from a pool of about 350 volunteers who had been screened for suggestibility in a hypnotic context using both the Harvard Group Scale of Hypnotic Susceptibility, Form A (HGSHS:A), (Shor & Orne, 1962) and then the Stanford Hypnotic Susceptibility Scale, Form C, (SHSS:C) sans the anosmia to ammonia challenge (Weitzenhoffer & Hilgard, 1962). All participants scored in the highly susceptible range (10–12 out of a possible 12 on the HGSHS:A; 9–11 out of a possible 11 on the SHSS:C). Data were acquired from all participants, some of whom have participated in previous studies.

After receiving an explanation of the procedures, participants provided written informed consent. Preceding the experiment, an experimenter notified the participants that the purpose of the study was to investigate the effects of suggestion on cognitive performance in highly suggestible individuals.
2.2. Materials and apparatus

We used the same Stroop materials that were used by Raz et al. (2002). Participants sat at a viewing distance of approximately 65 cm in front of a color computer monitor. Stimuli consisted of a single word written in one of four ink colors (red, blue, green, or yellow) appearing at the center of the computer screen where a black fixation cross was visible. All characters were displayed in upper case font against a white background and the stimuli subtended visual angles of 0.5° vertically, and 1.3° to 1.9° horizontally (depending on word length). Two classes of words were used: color words (RED, BLUE, GREEN, and YELLOW) and neutral words (LOT, SHIP, KNIFE, and FLOWER), the latter class being frequency- as well as length-matched to the color words.

Three experimental conditions were employed: a congruent condition consisting of a color word inked in its own color; a neutral condition consisting of a neutral word inked in any one of the four colors; and an incongruent condition consisting of a color word inked in any of the three colors other than the one to which it referred (e.g., the color word BLUE inked in green). During each trial, participants were asked to indicate the ink color in which a word was written by depressing one of four keys on a keyboard. The color-labeled response keys were V, B, N, and M for the colors red, blue, green, and yellow, respectively. Two fingers of each hand were used to press these response keys (e.g., left middle finger for V and right index finger for N). Speed and accuracy were emphasized equally.

2.3. Design and procedure

Participants were told that suggestions may be administered at certain points during the experiment and that they would be asked to play a computer game (i.e., Stroop) with the experimenter present in the room. Before the computer testing commenced, the following posthypnotic suggestion was administered to all participants after a standard hypnotic induction (Weitzenhoffer & Hilgard, 1962): “Very soon you will be playing the computer game. When I clap my hands once, meaningless symbols will appear in the middle of the screen. They will feel like characters of a foreign language that you do not know, and you will not attempt to attribute any meaning to them. This gibberish will be printed in one of four ink colors: red, blue, green or yellow. Although you will only be able to attend to the symbols’ ink color, you will look straight at the scrambled signs and crisply see all of them. Your job is to quickly and accurately depress the key that corresponds to the ink color shown. You will find that you can play this game easily and effortlessly. When I clap my hands twice, you will be able to easily read what appears on the screen.” Participants were then brought back to their common wakefulness. Thus, by the time participants were ready to perform on the task, they were dehypnotized with only a lingering posthypnotic suggestion to be triggered by a cue (i.e., the hand clap). Under the no-suggestion condition, participants were conventionally instructed to respond as quickly and as accurately to the ink color of the visual stimuli. In all cases, participants were instructed to foveate on the center fixation cross and respond by depressing a key indicating the ink color of the stimulus.

The experiment followed a within-subject design where each participant was asked to engage in the Stroop task twice, once with a posthypnotic suggestion that the stimuli would appear to be meaningless symbols and once without this suggestion. Both conditions were run outside of hypnosis. Order was counterbalanced, so that half of the participants were randomized to do the Stroop task without suggestion first and the remaining half to do it with suggestion first. Thus, the experimental design featured Suggestion (Present, Absent) and Congruency (Congruent, Neutral, Incongruent) as within-subject factors.

Participants were instructed to focus their eyes upon a fixation cross at the center of the screen. Then, a stimulus would appear on the screen replacing the crosshair. The stimulus remained on the screen for a maximum of 2 s or until participants responded. Upon a response, veridical visual feedback was provided (i.e., CORRECT or INCORRECT was centrally flashed in black ink) and the fixation cross was redisplayed at the center for a variable duration contingent upon the subject’s reaction time. At this point a new stimulus appeared on the screen again replacing the fixation cross and beginning the next trial. The inter-stimulus interval was 4 s.

A full practice session (i.e., 144 trials) preceded the first measurement for each subject. This training session was used to confirm that participants were able to understand the task, proficiently map the four ink colors to the appropriate response keys, and respond quickly and accurately. Following this training session, partici-
pants took a short break and then completed 144 experimental trials. One-third of all trials were neutral, congruent, and incongruent, respectively. Trial order was randomized throughout the experiment. Depending on the experimental order, the suggestion was then either triggered or deactivated. Following a brief rest period participants completed another set of 144 experimental trials. At the end of the experiment, the posthypnotic suggestion was cancelled.

3. Results

Table 1 shows mean reaction times (RTs) and accuracy scores as a function of suggestion and Stroop congruency conditions (congruent, neutral, or incongruent). As in previous assays, administration order was not significant and the data were accordingly collapsed. Incorrect responses were excluded from the RT analyses, as were latencies that were 3 standard deviations either above or below the mean per Stroop condition per participant. About 2% of the data were excluded due to such deviant RTs.

3.1. RT data

RTs data were analyzed using a linear mixed model. The RT analyses revealed main effects for Suggestion ($F(1,38) = 239.61, p < .0001, \text{MSe} = 9247529.62$) and Congruency ($F(2,96) = 148.92, p < .0001, \text{MSe} = 6194297.55$). In addition, the interaction Suggestion*Congruency was significant ($F(2,76) = 89.30, p < .0001, \text{MSe} = 3498704.57$). Whereas suggestion did not influence RTs for congruent trials, RTs for both neutral and incongruent trials were significantly shorter under suggestion compared to no-suggestion ($t(76) = 6.44, p < .0001$) and ($t(76) = 19.29, p < .0001$), respectively.

Stroop interference and facilitation are assessed as differences in RTs between incongruent and neutral trials (I–N) and between neutral and congruent trials (N–C), respectively; the Stroop effect is the gross difference between incongruent and congruent trials (I–C). Table 2 shows that while Stroop effect, interference, and facilitation were all significant in the absence of suggestion, both Stroop interference and facilitation were removed following suggestion. Table 3 shows that reductions in all three Stroop subtractions were significant as a function of suggestion.

3.2. Accuracy data

Because accuracy is a binary variable that is not normally distributed (i.e., it is heavily skewed toward zero), accuracy data were analyzed using both a linear mixed model (following the RT analysis) and a generalized

<table>
<thead>
<tr>
<th>Measure</th>
<th>Suggestion</th>
<th>Congruent</th>
<th>Neutral</th>
<th>Incongruent</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT (ms)</td>
<td>Absent</td>
<td>641 (179)</td>
<td>681 (193)</td>
<td>759 (269)</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td>636 (164)</td>
<td>646 (172)</td>
<td>652 (189)</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>Absent</td>
<td>96 (19)</td>
<td>94 (24)</td>
<td>93 (26)</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td>94 (24)</td>
<td>94 (24)</td>
<td>94 (24)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measure</th>
<th>Suggestion</th>
<th>Stroop Effect</th>
<th>Interference</th>
<th>Facilitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT</td>
<td>Absent</td>
<td>$F(1,209) = 477.03, p &lt; .0001$</td>
<td>$F(1,209) = 208.30, p &lt; .0001$</td>
<td>$F(1,209) = 55.86, p &lt; .0001$</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td>$F(1,209) = 7.87, p &lt; .001$</td>
<td>$F(1,209) = 1.15, p = .29$</td>
<td>$F(1,209) = 3.02, p = .08$</td>
</tr>
<tr>
<td>Error</td>
<td>Absent</td>
<td>1.89 (.43–2.50), p &lt; .0001</td>
<td>1.50 (1.20–1.87), p &lt; .0005</td>
<td>1.26 (1.04–1.52), p &lt; .05</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td>1.06 (.87–1.29) p = .58</td>
<td>1.05 (.81–1.36) p = .69</td>
<td>1.89 (1.43–2.50) p = .98</td>
</tr>
</tbody>
</table>

Significant results are shown in bold.
estimation equation (proc genmod in SAS). Results from these two analyses were comparable, albeit the latter (i.e., odds ratio) was a more sensitive measure. While Stroop effect, interference, and facilitation were significant in the absence of suggestion, following suggestion all were removed (Table 2). Comparing errors as a function of suggestion for congruent, neutral, and incongruent trials did not reveal a significant difference. However, errors for Stroop effect and interference reduced significantly under suggestion (Table 3).

4. Discussion

Consistent with previous research (Raz, 2004; Raz, Fan, et al., 2005; Raz et al., 2003; Raz & Shapiro, 2002; Raz et al., 2002), our findings indicate that a specific suggestion to construe words as meaningless symbols of an unknown foreign language substantially reduces the Stroop effect in highly hypnotizable individuals. This finding indicates that cognitive processes that had been automatized through practice can be de-automatized and brought under cognitive control.

Examination of the results reveals that in the absence of suggestion, all Stroop subtractions were significant. After suggestion, however, RT interference and facilitation were removed and the Stroop effect was reduced; faster RTs were detected for both the neutral and incongruent trials. Accuracy data show that whereas all Stroop subtractions were significant without suggestion, following administration of the suggestion these effects were removed (Table 2). Indeed, accuracy performance was leveled across the three Stroop conditions and identical to accuracy scores of the neutral condition in the absence of suggestion (Table 1). With the exception of facilitation for the accuracy data, all Stroop subtractions for both RT and accuracy changed significantly as a function of suggestion. That RTs for congruent trials did not get any faster following suggestion may indicate that participants already perform optimally for compatible stimuli and can only improve, without compromising their accuracy, by speedier processing of the neutral and incompatible Stroop words.

The Stroop effect is thought to illustrate how processing of irrelevant information takes place even when it is unfavorable to the task at hand. Because the Stroop test is a ballistic attentional phenomenon and is difficult to reduce with practice, there is widespread accord that skilled readers obligatorily process printed stimuli presented to the fovea. Our present data, in conjunction with our previous assays, confirm that within specific attentional contexts the Stroop effect can be significantly reduced (Raz, Fan, et al., 2005) and even completely eliminated (Raz et al., 2003; Raz et al., 2002) using a pristine Stroop paradigm. Whereas other studies have also reported reduction and even elimination of Stroop interference (Besner, 2001; Melara & Algom, 2003), these studies manipulated the original experimental design and deviated from the classical Stroop paradigm. The novelty in our approach is that we kept the design pure and manipulated the participants instead (Raz, 2004; Raz & Buhle, 2006).

Unlike our previous behavioral findings (Raz et al., 2003; Raz et al., 2002), the present data show that although Stroop interference and facilitation were eliminated, the Stroop effect was not completely removed. A recent functional magnetic resonance imaging (fMRI) study also reported reduction, not elimination, of the Stroop effect (Raz, Fan, et al., 2005). However, reduction rather than removal of the Stroop effect in the context of fMRI likely pertains to the specific constraints of this imaging methodology which may make the efficacy of suggestion suboptimal: fMRI requires participants to lie down and hold still for the duration of a scan in the presence of loud ambient noise (Raz, Lieber, et al., 2005). Indeed, we have recently demonstrated that reduction in the Stroop effect in an fMRI setting can turn into complete removal when the same participants are sitting upright in the quieter event-related potentials (ERPs) environment (Raz, 2004; Raz, Fan, et al., 2005).
Another instance of Stroop reduction, rather than removal, was recently reported in an independent replication and extension study where reduction of Stroop interference was accomplished regardless of whether hypnosis was formally induced and where nonhypnotic-suggestion, rather than hypnotic-, reduced the Stroop effect (Raz et al., 2006). Anecdotal case reports support the removal of the Stroop effect at the individual level (MacLeod & Sheehan, 2003; Schatzman, 1980); the same is true for unpublished informal accounts (e.g., Thalia Wheatley, personal communication). However, why suggestion removes the Stroop effect in some studies (Raz et al., 2003; Raz et al., 2002), reduces it in others (Raz, Fan, et al., 2005; Raz et al., 2006), and fails to produce a reliable group effect despite the presence of individual effects (Amanda Barnier, Max Coltheart, and Derek Besner, personal communication) remains to be reconciled. Differences between the procedures of our studies and those of other researchers include methodological procedures, the setting and physical environment, participant population, suggestibility scale used to screen participants, prior familiarity with the Stroop task, authority and presence of the experimenter, the specific induction used, and audio-taped versus live administration of the suggestion. These parameters likely account for at least some of the outcome variability. In addition, the results reported here further propose that these differences may represent random variations in the size of a decisive effect.

Interpretation of recent neuroimaging data proposes that our experimental suggestion worked by influencing occipital regions and then the anterior cingulate cortex (ACC) (Raz, 2004; Raz, Fan, et al., 2005): the former may be related to processing visual words, but controversy persists regarding the visual word form area (Price & Devlin, 2003); the latter is involved in conflict monitoring and resolution (Bush, Luu, & Posner, 2000). Nonetheless, the combined imaging findings propose that rather than selective abrogation of orthographic processing, the entire visual input stream was dampened down.

If suggestion can override what cognitive scientists typically consider an automatic process (i.e., reading in proficient readers), the label “automaticity” may require further refinement. Furthermore, the influence of suggestion may have important therapeutic implications in the deautomatization of other ingrained behaviors in the context of certain psychopathologies (e.g., impulse control disorders) and in elucidating the neural substrates of the placebo response (Wager et al., 2004). For example, interpretation of our recent pilot data from children diagnosed with Tourette’s syndrome indicates that hypnotic suggestion can transiently ameliorate tic symptoms (Raz & Norman, 2004). The notion of testing an individual while changing her attentional efficiency with suggestion, instead of changing the experimental task, accords with recent reports describing the effects of attentional training and meditation (Carter et al., 2005; Rueda, Rothbart, McCandliss, Saccamanno, & Posner, 2005). Furthermore, hypnotizability may have a genetic component (Raz, 2005; Raz, Fan, & Posner, in press; Raz, Fossella, McGuiness, Zephrani, & Posner, 2004). Thus, hypnotic virtuosos may well be excellent candidates to unravel the effect of volitional control within this newly formed field of “atypical attention” (Raz & Buhle, 2006).

We conclude that a suggestion to construe words as meaningless scribbles reduced the Stroop effect in highly hypnotizable individuals. In conjunction with previous studies of this effect, including those which manipulated the experimental design regardless of participant suggestibility, the current study challenges the dominant view that word recognition is obligatory for proficient readers. Furthermore, this effect does not seem to depend on the formal induction of hypnosis in highly suggestible individuals (Raz et al., 2006). Effective suggestion likely operates through a top-down mechanism that modifies the processing of input words (Raz, 2004). We have begun to identify the neural correlates subserving the influence of suggestion and recently reported a combined ERP and fMRI study that highlighted the role of altered activations in early visual processing and ACC (Raz, Fan, et al., 2005). Further studying the brain mechanisms associated with suggestion will likely elucidate its top-down impact and may shed more light on such factors as anticipation and placebo. Here, we show that it is possible to “un-ring the bell” and regain control over an automatic process. Such de-automatization may be fundamental for clinical interventions (e.g., cognitive therapy) as well as for furthering our understanding of automatic processes, volitional control, and their impact on consciousness.

References


