Consciousness and Cognition 75 (2019) 102809

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

Involuntary mental rotation and visuospatial imagery from external control

Donish Cushing[a], Adam Gazzaley[b,c], Ezequiel Morsella[a,b,*]

[a] Department of Psychology, San Francisco State University, United States
[b] Department of Neurology, University of California, San Francisco, United States
[c] Departments of Psychiatry and Physiology, University of California, San Francisco, United States

ARTICLE INFO

Keywords:
Reflexive imagery task
Mental rotation
Visuospatial imagery
Unconscious processing
Mental imagery

ABSTRACT

The Reflexive Imagery Task (RIT) was developed to investigate the entry into consciousness of involuntary imagery. Subjects are presented with objects and instructed to not think of the names of the objects. Involuntary subvocalizations arise on many trials. RIT effects reveal the capacities of involuntary processing. These cognitions do not require symbol manipulation. Can mental rotation and visuospatial imagery, too, arise in this involuntary manner? In the mental rotation task, subjects were first taught to mentally rotate two-dimensional objects. Subjects were then instructed to not mentally rotate objects. In the chess task, subjects were taught how to move in their minds objects in specified ways, much as one could imagine how chess pieces move on a chessboard. Subjects were then instructed to not have such visuospatial imagery. For both tasks, involuntary imagery occurred on a substantial proportion of trials, revealing that symbol manipulation can be influenced involuntarily through external control.

1. Introduction

Early in the morning, the eyes open and one immediately experiences percepts and urges—the sight of sunlight entering the room, the smell of breakfast, and the urge to have a drink of water. To the observer, these *conscious contents* simply "just happen" (Morsella, Godwin, Jantz, Krieger, & Gazzaley, 2016a). This everyday scenario illustrates that most of the contents composing the *conscious field*[1] arise effortlessly, passively, and involuntarily (Morsella et al., 2016a). Perception research (e.g., Allen, Krisst, Montemayor, & Morsella, 2016; Firestone & Scholl, 2016) reveals that entry into consciousness of this nature ("involuntary entry," for short) is influenced by many variables.[2] As illustrated in

* Corresponding author at: Department of Psychology, San Francisco State University, 1600 Holloway Avenue, EP 301, San Francisco, CA 94132-4168, United States.
E-mail address: morsella@sfsu.edu (E. Morsella).

[1] Each thing one is conscious of is referred to as a "conscious content" (e.g., a yellow afterimage or nausea). The *conscious field* is composed of all the conscious contents activated at one time.

[2] The variables influencing entry of a particular stimulus into consciousness include the salience, novelty, motion, or incentive/emotional quality of the stimulus (Gazzaley & D'Esposito, 2007; Goodhew, 2017). The mechanisms underlying involuntary entry seem to vary across modalities. For instance, a "pop-out" effect (Treisman & Gelade, 1980) may influence entry in vision, but it is less likely to do so in olfaction.

https://doi.org/10.1016/j.concog.2019.102809
Received 21 April 2019; Received in revised form 23 August 2019; Accepted 26 August 2019
Available online 12 September 2019
1053-8100/

our example, urges, too, can enter consciousness in this way (Loewenstein, 1996; Morsella, Gray, Krieger, & Bargh, 2009; Morsella, Wilson et al., 2009).[3]

Other forms of involuntary entry stem from a combination of sets[4] and external stimuli. Take, for example, the situation noted long ago by Ach (1905/1951): If one has activated the set to add before hearing "one and one," then one will experience the conscious content "two." If, however, the set that had been activated was to subtract instead of add, then one would have experienced "zero" instead of "two." This form of involuntary entry has been referred to as *set-based entry* (Bhangal, Merrick, Cho, & Morsella, 2018), which has been distinguished from the involuntary entry mentioned above concerning percepts (e.g., a loud sound) and visceral urges (e.g., hunger).

1.1. The Reflexive Imagery Task

The Reflexive Imagery Task (RIT; Allen, Wilkins, Gazzaley, & Morsella, 2013; see review in Bhangal, Cho, Geisler, & Morsella, 2016) was developed to investigate the nature of involuntary entry from a combination of external stimuli and the activation of sets. The task stems from "subjective" variants of the Eriksen flanker task (e.g., Morsella, Gray et al., 2009; Morsella, Wilson et al., 2009; see discussion in Desender, van Opstal, & van den Bussche, 2014, and in Questienne, Atas, Burle, & Gevers, 2018), in which distractors activate involuntary urges and other conscious contents.[5] Other aspects of the RIT are based on theories (Morsella et al., 2016a; see Discussion) and on experiments by Ach (1905/1951), Stroop (1935), Uznadze (1966), Wegner (1989), and Gollwitzer (1999).

In the RIT, subjects are instructed to not perform a mental operation (e.g., to count or name an object) on to-be-presented stimuli. For example, before being presented with three circles, subjects might be instructed to not count the number of objects presented on the screen (Bhangal et al., 2018). On most trials, despite the intentions of the subject, the undesired mental operations still arise, yielding "three." The RIT uses negative instructions only to diminish artifacts stemming from demand characteristics and strategic processing on the part of the subject. However, without such negative instructions, RIT effects still arise at comparable rates.[6]

RIT effects of a complex nature have been observed. For example, in Merrick, Farnia, Jantz, Gazzaley, and Morsella (2015), subjects were presented with drawings of objects and instructed to (a) not think of the name of the object, and (b) not count the number of letters composing the object name. RIT effects arose for both mental operations on ~30% of the trials. In another study (Cho, Zarolia, Gazzaley, & Morsella, 2016), subjects first learned to transform words according to a rule resembling that of the childhood game of Pig Latin. After training, subjects were presented with words and instructed to not transform the words according to the newly-learned rule. Involuntary transformations arose on ~40% of the trials. It is worth noting that this involuntary effect requires, not only memory retrieval, but symbol manipulation, a process associated with frontal cortex (Miller & Cummings, 2007).

Might tasks that are more complex than those in Merrick et al. (2015) and in Cho et al. (2016) reveal the boundary conditions of the RIT effect? Such boundary conditions would shed light on the limits of involuntary processes and thereby illuminate the contributions of conscious processing.

1.2. The validity of subjects' self-reports

Demand characteristics and inaccurate memories of ephemeral conscious contents (Block, 2007) could yield inaccurate self-reports in an RIT (see discussion in Morsella, Wilson et al., 2009). However, evidence suggests that subjects' self-reports are accurate. First, in an RIT (Cushing, Gazzaley, & Morsella, 2017) in which subjects reported the occurrence of the basic RIT effect and also had to press a button if the involuntary subvocalization they experienced rhymed with a word held in mind, performance (>80% mean accuracy across trials) corroborated that subjects did experience involuntary subvocalizations. This is because detecting a rhyme requires retrieval of the phonological form of a word.

Second, in Bhangal et al. (2018), subjects were presented with an array of visual objects and instructed to not count the number of objects. When the number of objects was small (2–5 objects), the involuntary counting was very accurate (~90% mean accuracy), suggesting that the counting did in fact occur. Third, in RITs involving lexical retrieval, involuntary subvocalizations are influenced by the word frequency of the name of the object: High-frequency words are more likely to yield an RIT effect than low-frequency words (Bhangal, Merrick, & Morsella, 2015). Such a frequency effect would be unlikely to stem from demand characteristics. Regarding the possibility of artifacts resulting from strategic processing, on many trials of the RIT, the effect arises too quickly to be

[3] Investigations on action control have illuminated that involuntary entry of urges can arise from bodily needs (Loewenstein, 1996) and from the activation of conflicting action plans (Desender et al., 2014; Lewin, 1935; Morsella, Gray et al., 2009; Morsella, Wilson et al., 2009; Questienne et al., 2018). Moreover, metacognitions (e.g., action-related urges) can enter consciousness insuppressibly as a function of set and the presentation of external stimuli (Garcia, Bhangal, Velasquez, Geisler, & Morsella, 2016).

[4] Sets, such as mindsets or task sets, are dispositions to behave or think in certain ways.

[5] The flanker task precedes research on *ironic processing*,[8] which is associated with failures of self-regulation (e.g., in nicotine addiction; Wegner, 1989). (The ironic effect was noted long ago by Dostoevsky (1863/2008).) It is important to note that research on the RIT and on ironic processing stem from different theoretical backgrounds and are concerned with different phenomena and with answering different questions. For example, unlike ironic processing, the RIT was designed to investigate, not failures in self-control, but rather the nature of involuntary entry from, specifically, the combination of external stimuli and activated sets.

[6] It is important to note that RIT effects have arisen in RITs that lack any kind of negative instruction to not perform some kind of mental operation (e.g., see the Baseline Condition in Allen et al., 2013). To take one example, in Allen et al. (2016), subjects were instructed to hold in mind one way of perceiving an ambiguous object (e.g., Necker cube). Despite the lack of negative instructions, involuntary "perceptual reversals," involving entry into consciousness, occurred on ~80% of the trials.

caused by such processing (Allen et al., 2013; Cho, Godwin, Geisler, & Morsella, 2014).[7] In addition, neuroimaging data from studies (not involving the RIT) in which subjects self-report about the occurrence of involuntary thoughts or about subvocalizations corroborate that subjects do not confabulate about their reported mental events (Mason et al., 2007; McVay & Kane, 2010; Mitchell et al., 2007; Pasley et al., 2012; Wyland, Kelley, Macrae, Gordon, & Heatherton, 2003).

1.3. The involuntary nature of the RIT effect

The view that the RIT effect is involuntary stems in part from theoretical explanations of the effect. According to Wegner (1994), undesired, "ironic" effects[8] such as the RIT effect arise from an interaction between two distinct mechanisms. One mechanism is an *operating* process. This process is associated with the conscious intention to maintain a particular mental state. The process tends to be effortful, capacity-limited, and consciously mediated (Wegner, 1994). The mechanism actively scans mental contents (e.g., thoughts, sensations) that can help maintain the desired mental state (e.g., to be calm). The other mechanism is an 'ironic' *monitoring* process that automatically scans activated mental contents to detect contents signaling the failure to establish the desired mental state. When such a content is detected, that content then enters the conscious field. Of import, in Wegner (1994), the monitoring process is usually unconscious and autonomous. In other accounts of the RIT effect (Ach, 1905/1951; Bhangal et al., 2016), the effect is the consequence of sets having been activated by the verbal instructions provided to the subject. From this standpoint, merely hearing the word "add" in the instruction "Do not add the following numbers" increases the activation level of the set to add, yielding "four" in response to the stimuli "2 and 2." This *set activation* account is consistent with the tenets of *parallel distributed processing* (Rumelhart, McClelland, & the PDP Research Group, 1986), in which stimuli can activate, in an interactive network, "units" representing sets or rules. These units can in turn then automatically influence the activation thresholds of other, related units (e.g., representing "FOUR") in the network. The RIT effect would also be deemed to be involuntary from the standpoint of a "crossmodal" account of mental representation. From this standpoint, the mental representations of objects are inherently multimodal (Ernst & Bülthoff, 2004; Lacey & Lawson, 2013; Spence & Deroy, 2013), in such a way that the activation of one sensory attribute of an object (e.g., the image CAT) will trigger automatically the activation of the other sensory features composing the multimodal representation, including those associated with linguistic label "cat." Of most importance, in all theorizing regarding the mechanisms underlying the RIT effect, the nature of the effect is deemed to be involuntary.

1.4. Extension of the RIT effect to the realm of high-level, symbol manipulation

One early concern was that the RIT effect was not noteworthy because stimulus-elicited memory retrieval is often automatic (Schacter & Tulving, 1994). However, such an interpretation cannot account for the RIT effects found in Cho et al. (2016), in which the involuntary verbal imagery required symbol manipulation, which is more than just memory retrieval. However, one criticism of Cho et al. (2016), in which subjects experienced involuntary word transformations, is that the involuntary effect observed (symbol manipulation involving syntax) is not noteworthy because syntax is largely unconscious. This led to the hypothesis that such symbol manipulation should fail to arise for a modality other than language (e.g., vision), a hypothesis whose falsification has important implications for theories regarding the limitations of involuntary, unconscious processes.

Perhaps the true boundary conditions of the RIT effect lie in mental operations requiring other forms of symbol manipulation, such as visuospatial imagery and mental rotation. Addressing this possibility is informative because symbol manipulation is associated with both executive function and cognitive control (Enger, 2017). With this in mind, one could conclude that perhaps *RIT effects cannot arise for operations that require symbol manipulations that are visuospatial in nature.* Identifying such a boundary condition would be informative, for the boundary conditions of the RIT effect reveal some of the limits of involuntary processing and thereby shed light on the contributions of conscious processing.

1.5. The present approach

In the mental rotation task, subjects were first taught to mentally rotate (30°, 60°, or 90°) two-dimensional nonsense objects. After training, subjects were instructed to not mentally rotate in these ways a different set of objects. In the chess task, subjects were taught how to move in their minds (i.e., visuospatial imagery) objects in specified ways, much as one could imagine how, in the game of chess, a given piece can navigate the chessboard. Each object was associated with a unique pattern of potential movement on a chessboard-like grid. After training, subjects were instructed to not think of where each object could move on the grid. The order of presentation of the two tasks was fully counterbalanced across subjects.

Our aim was to assess whether high-level phenomena such as symbol manipulation can be influenced involuntarily and systematically through external control. We sought to obtain substantive evidence that, under controlled laboratory conditions designed to minimize artifacts and measurement error, these effects on symbol manipulation can occur involuntarily and at a reliable rate. This

[7] Additional behavioral data that corroborate subjects' reports about the RIT effect are the following. The effect arises when there is cognitive load, a condition in which it is difficult for subjects to carry out strategic processing (Cho et al., 2014). In addition, RIT effects are systematically more likely for some sensory modalities (e.g., vision or phonological) than for others (e.g., olfaction; Dou, Li, Geisler, & Morsella, 2018). Such a pattern of results is unlikely to arise from strategic processing or demand characteristics.

[8] Ironic effects arise when one thinks about a certain thing, such as a memory or some form of mental imagery, while attempting to not think about that thing. For reviews of ironic processing and thought suppression, see Rassin (2005) and Wegner (1989).

would provide evidence that high-level, symbol manipulation can occur involuntarily.

We should add that the RIT is the kind of paradigm that, because it builds incrementally on robust phenomena and previous research, has been encouraged by leading researchers in the field (e.g., Fiedler, 2017; Nosek, Spies, & Motyl, 2012). The reliable, component processes of the present variant of the RIT are of interest in disparate subfields of the study of mind and brain, including consciousness, cognitive control, and imagery.

2. Method

2.1. Subjects

Four San Francisco State University students ($M_{\text{Age}} = 20$, $SD_{\text{Age}} = 1.63$, females = 3) volunteered to participate, and thirty more ($M_{\text{Age}} = 22.89$, $SD_{\text{Age}} = 7.42$, females = 21) participated for course credit. The involvement of human subjects in our study was approved by the Institutional Review Board at San Francisco State University. The sample size ($n > 10$) was based on the effect size (Cohen's d [on raw proportions] = 1.72; Cohen's h [on raw proportions] = 1.44; Cohen's d [on arcsine transformations of the proportion data] = 1.38), SD (0.25), and other aspects of a previous project (Cho et al., 2016) that, similar to the present project, was designed to illuminate the boundary conditions of the RIT effect. To determine the sample size, we used the program *G*Power 3* (Faul, Erdfelder, Lang, & Buchner, 2007). The input parameters were: Cohen's $d = 1.72$, one sample t-test, tails = one, power = 0.95, and $\alpha = 0.05$. The output parameters were: noncentrality parameter = 4.21, critical $t = 2.02$, and actual power = 0.97.

2.2. Stimuli and apparatus

Stimulus objects were presented on an Apple iMac computer monitor with a 50.8 cm screen. Subjects were seated approximately 48 cm away. Stimulus presentation was controlled by PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). All instructions, prompts, and questions were presented in black 36 pt. Helvetica font on a white background. Stimuli for both tasks were black images of two-dimensional nonsense objects. Five of these objects were used successfully in previous research (Merrick, Cooper, Jantz, & Morsella, 2013).

In the chess task, four black-and-white nonsense objects were used (Fig. 1). None of these objects connoted a direction based on their shape. Each object moved in only one manner. Two of the objects moved two spaces (e.g., one space to the right, and one space up), and two of the objects moved three spaces (e.g., two spaces to the right, and one space down; Fig. 1). The objects that moved three spaces did not simply move in opposite directions. When presented in the center of the screen during the initial instructions, the objects had a subtended visual angle of $4.77^{\circ} \times 4.77^{\circ}$ (4 cm × 4 cm). When appearing within the 6-cell by 5-cell grid during the instructions, as when subsequently appearing in the practice trials and the critical trials, these objects all had a subtended visual angle of $3.28^{\circ} \times 3.28^{\circ}$ (2.75 cm × 2.75 cm). The grid they appeared in had a subtended visual angle of $21.82^{\circ} \times 18.34^{\circ}$ (18.5 cm × 15.5 cm). The grid was positioned in the center of the screen. Each object appeared in one of six starting locations in the grid. These starting locations included cells that were, horizontally, two spaces left of center, one space left of center, or one space right of center and, vertically, on center, or one space below center. When presented during the learning trials, the starting locations of these objects were never in these six cells. The center of the grid was marked by a fixation cross (+) that was white with a thick black border. The fixation cross was made this way so it would be visible while on the grid. The fixation cross had a subtended visual angle of $1.19^{\circ} \times 1.19^{\circ}$ (1 cm × 1 cm).

The objects used in the mental rotation task were all presented in the center of the screen with a subtended visual angle less than $5.37^{\circ} \times 5.37^{\circ}$ (4.5 cm × 4.5 cm) (Appendix A). The three objects used in the learning trials, the four objects used in the practice trials, and the four objects used in the critical trials were all distinct.

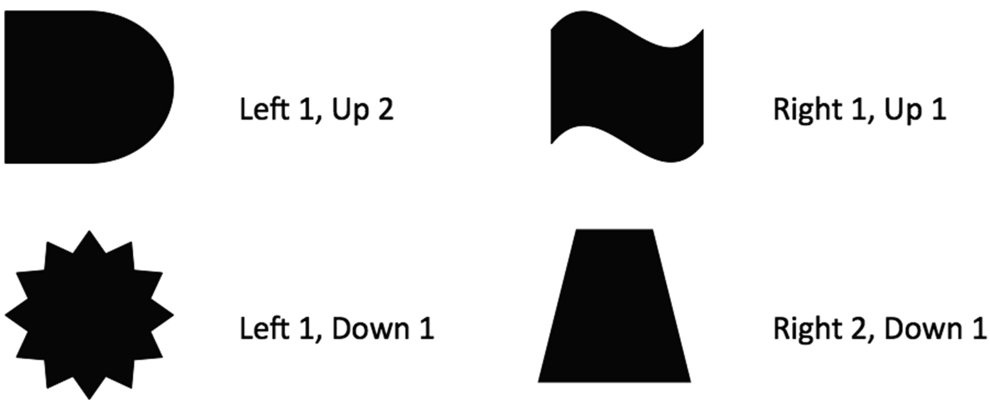


Fig. 1. Stimuli from chess task. Not drawn to scale.

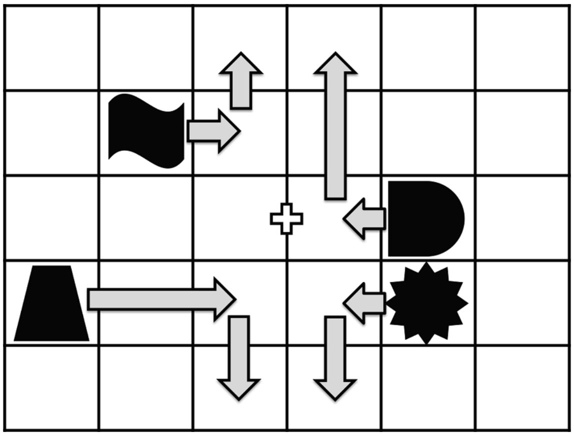


Fig. 2. Illustration of how each object moves across the grid. Not drawn to scale.

2.3. Procedures

Each subject was run individually. The order in which subjects completed the chess task and the mental rotation task was fully counterbalanced.

In the chess task, subjects were instructed that they would learn how four distinct objects move on a grid. Each object would move only one way, and would move the same way regardless of where it started on the grid (Figs. 1 and 2). After subjects indicated that they understood these instructions, they were invited to press the return key to advance to further instructions.

These further instructions followed the following format: First, subjects were shown the image of an object with instructions for how that object moves (e.g., “This object moves one space to the left, and two spaces up”; Fig. 1). Second, subjects were shown an example of how the object moves across the grid (Fig. 2). Third, the object was presented in a starting location on the grid, and subjects were instructed to indicate, on a paper version of the grid, what the starting and ending locations of the object were, as well as draw arrows describing the path it took across the grid. Subjects pressed the return key to advance through each stage of each learning trial. This sequence occurred twice for each object. Each object had the same movement pattern for each subject. This minimized experimenter error when providing feedback about the correct ending location of the object.

Subjects were instructed that, in the trials that followed, they would be presented with a series of objects on a grid. They would look at the fixation cross until the object was presented, and then look at the object. They would press the spacebar when they thought of the ending location of the object. They were told as well that they would indicate, by using another grid, the ending location of the object. With these instructions in mind, subjects were given two minutes to use their paper grids to memorize how each object moves. When the two minutes expired, subjects handed their paper grids back to the researcher. Next, subjects were reminded (a) that they were to press the spacebar as soon as they thought of the ending location of the object, (b) to indicate what that ending location was; and (c) the image would be presented for a set amount of time regardless of whether the spacebar was pressed. Subjects were additionally instructed to keep their finger or their thumb rested on the spacebar at all times. Subjects pressed the return key to advance to the practice trials. (Aside from the instructions to think of the ending location of the object, and to use the paper grids to memorize the way each object moves, all instructions for the practice trials were repeated for the critical trials.)

In the practice trials (n = 48), subjects were presented with the instruction, “Think of the ending location of the object.” This instruction served as a “ready” prompt. To commence each trial, subjects pressed the return key. The grid and a fixation cross were simultaneously presented for 500 ms before an object additionally appeared on the grid. The object, grid, and fixation cross remained on the screen for 6 s. During this time, response times (RTs) for spacebar presses were recorded. Next, a second grid appeared on the screen with a number-letter combination in each of its cells (e.g., 1a, 2a, 4c, etc.). Subjects typed the number and letter of the cell corresponding to the ending location of the object. If the ending location of the object did not occur to them, they typed, “I did not think of it.” Following this indication, they pressed the return key to complete the trial. During these trials, the experimenter was present to provide feedback regarding the accuracy of subjects’ responses when indicating the ending location of the object. On trials in which subjects produced an incorrect answer, or did not know the answer, the experimenter reiterated the instructions that were associated with the object in the learning trials (e.g., “That object moves one space to the left, and one space down”). On trials in which subjects produced the correct answer, the experimenter said, “That is correct.” If, after 10 trials, subjects were unable to produce the ending locations for a majority of the objects, they were given an additional 60 s to use their paper grids to memorize how each object moves. Every object appeared in each of the six starting locations twice. The order in which each object appeared in each starting location was random.

For the subsequent critical trials (n = 24), subjects were instructed to *not* think of the ending location of the object (see sample trial in Fig. 3). However, if they happened to think of the ending location of the object, then they pressed the spacebar. Subjects were asked to use the same standard as before for judging whether or not they thought of the ending location of the object. The prompt appearing at

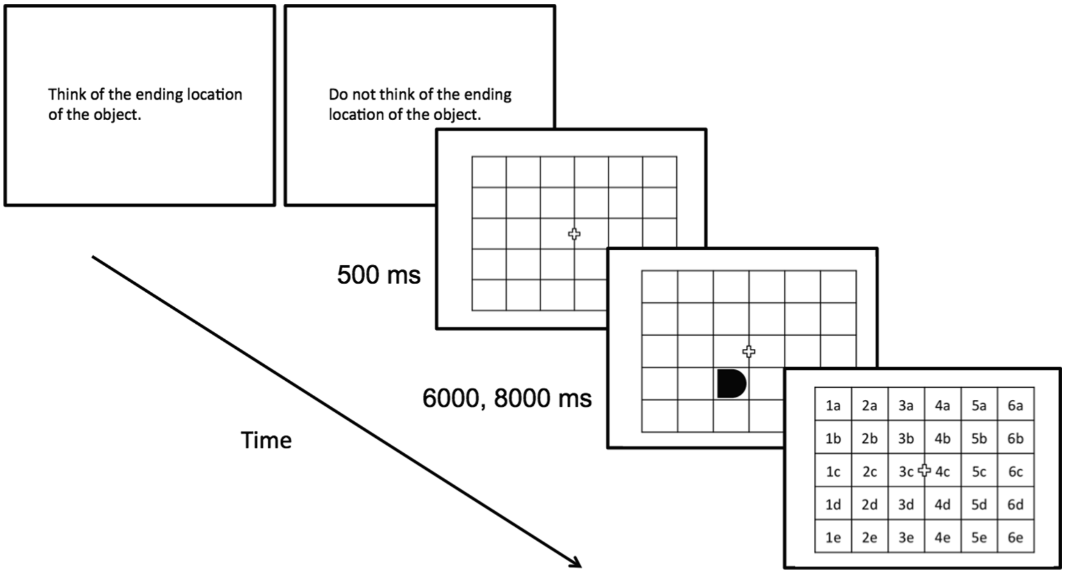


Fig. 3. Schematic depiction of two typical trials in the chess task. Not drawn to scale.

the beginning of each critical trial was "Do not think of the ending location of the object." The same objects were presented in the same starting locations as in the practice trials. The fixation cross was also presented for the same amount of time. The object, the grid, and the fixation cross remained on the screen for 8 s. During this time, spacebar RTs were recorded. Subjects used the second grid to indicate the ending location of the object in the same way as in the practice trials. Subjects pressed the return key to conclude each trial. During the critical trials, the experimenter remained nearby in case subjects had any questions. Again, the order in which each object appeared in each starting location was random. Every object appeared in each of the six starting locations only once.

In the mental rotation task, subjects were instructed that their task was to learn how to mentally rotate objects three different distances: "a short distance", "a medium distance", and "a long distance." All rotations would occur in the clockwise direction. After subjects indicated that they understood these instructions, they were invited to press the return key.

Subjects were taught the different mental rotation distances in the following way. First, they were shown an object at a starting orientation. Next, they were shown that same object rotated "a short distance" (30°), "a medium distance" (60°), and "a long distance" (90°). The rotated object was presented with instructions specifying the distance rotated (e.g., "This is what the object looks like when rotated 'a short distance'"). This sequence occurred three times, each time with a different object. Subjects were never explicitly told the degree measurement that related to each one of the rotation distances. This made it easier for subjects to understand the instructions. Objects used during these instructions were never used again during the practice or critical trials. By using rotations of 90° or less, no situation arose wherein it would be easier or more convenient for subjects to rotate the object counterclockwise.

Subjects were instructed that, in the trials that would follow, (a) they would be presented with a series of objects that they would mentally rotate one of the three distances, (b) they would press the spacebar when they mentally rotated the object, (c) they were to keep their finger or their thumb rested on the spacebar at all times, (d) the object would appear on the screen for a fixed amount of time whether or not they pressed the spacebar, and (e) they would indicate, on a second screen, if the amount they rotated the object in their mind matched an image of the object rotated by typing "y" (for yes) or "n" (for no). Subjects pressed the return key to advance to the practice trials. (Aside from the instructions to mentally rotate the object, all instructions for the practice trials were repeated for the critical trials.)

At the beginning of each practice trial (n = 48), subjects were presented with one of the following instructions, which served as a ready prompt: (1) "Think of the object rotated a short distance," (2) "Think of the object rotated a medium distance," and (3) "Think of the object rotated a long distance." Subjects pressed the return key to commence each trial. A fixation cross was presented for 500 ms. Next, one object that was distinct from those in the learning trials was presented for 6 s. During this time, RTs for spacebar presses were recorded. Next, a second image of the object was shown that was rotated by the amount that was specified by the prompt at the beginning of the trial. Subjects typed "y" or "n" to indicate if the distance they rotated the object in their mind matched this second image, and then pressed the return key to conclude the trial. During the practice trial block, there were four objects each presented at two different starting orientations (0° and 180°). Each ready prompt was paired with each object at each starting orientation twice. The order of the presentations of these prompt-object pairs was random. Although the second image of the object was always rotated the amount that was specified by the ready prompt, subjects were never explicitly told that this was the case.

For the subsequent critical trials (n = 24; Fig. 4), subjects were presented with one of the following instructions, which served as a ready prompt: (1) "Do not think of the object rotated a short distance", (2) "Do not think of the object rotated a medium distance", (3) "Do not think of the object rotated a long distance." If, however, subjects happened to mentally rotate the object, then they pressed

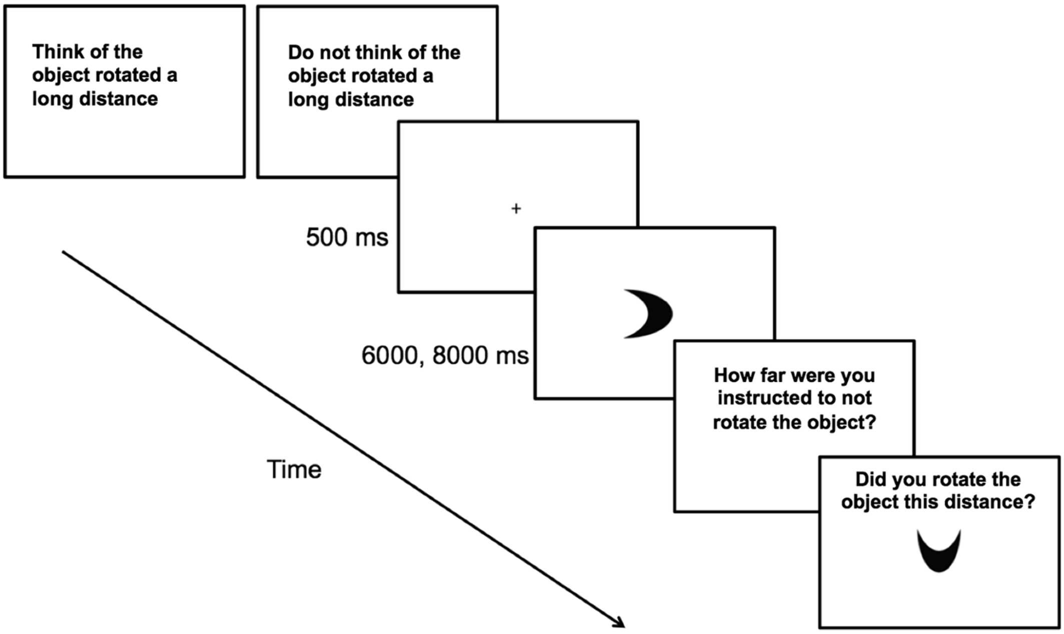


Fig. 4. Schematic depiction of two typical trials in the mental rotation task. Not drawn to scale.

the spacebar. A fixation cross was presented for 500 ms. Next, one object that was unique to the critical trials was presented for 8 s. During this time, RTs for spacebar presses were recorded. Next, in order to eliminate the potential strategy of forgetting, or not paying attention to the prompted rotation distance, subjects were instructed to type the distance specified by the prompt at the beginning of the trial. For example, if the prompt was, "Do not think of the object rotated a short distance", subjects had to type, "A short distance." If they could not remember the prompted distance, they typed, "I do not remember." Subjects pressed the return key to continue. Subjects again indicated if the amount they rotated the object in their mind matched a picture of the object rotated by typing "y" or "n" and then pressed the return key to conclude the trial. During the critical trial block, there were four objects each presented at two different starting orientations (0° and 180°). Each ready prompt was paired with each object at each starting orientation once. The order of the presentations of these prompt-object pairs was random. Subjects were never shown an object rotated the instructed distance before being instructed to mentally rotate the object that distance.

Once subjects completed the experiment, they completed a paper-and-pencil funneled debriefing questionnaire (following the procedures of Bargh & Chartrand, 2000). The questionnaire was designed to identify any subjects whose data should be excluded from analysis. The funneled debriefing included general questions to assess whether subjects (a) were aware of the purpose of the study, (b) had any strategies for completing the tasks, (c) had anything interfere with their performance on the tasks, and (d) were capable of learning, in the practice trials, what was necessary to perform meaningfully in the critical blocks.

3. Results

3.1. Chess task

Subjects indicated by button press that involuntary imagery, that is, the RIT effect, occurred on a substantive proportion of the trials ($M = 0.62$, $SD = 0.30$, $SE = 0.05$, Range = 0–1), a proportion that was significantly different from zero, t (31) = 11.90, $p < .0001$. The mean RT for these presses was 2,978.49 ms ($SD = 838.45$, $SE = 150.59$, Range = 612.71–4,389.50 ms). When considering only the trials in which the imagery was accurate (reflecting the correct final position of the piece), the rate was still substantive ($M = 0.54$, $SD = 0.29$, $SE = 0.05$, Range = 0–0.96) and significantly different from zero, t (31) = 10.58, $p < .0001$ (Cohen's d [on raw proportions] = 1.86; Cohen's h [on raw proportions] = 1.66). A comparable effect size is obtained with arcsine transformations of the proportion data (Cohen's $d = 1.46$).

The mean RT (Fig. 5) for these accurate presses was 2,956.61 ms ($SD = 941.43$, $SE = 171.88$, Range = 612.83–5,145.75 ms). In response to the funneled debriefing question, "*Did you have a strategy for NOT thinking of the ending location of the object on the grid? What was it?*", 27 out of 32 subjects indicated clearly that they attempted some kind of strategy to thwart the RIT effect.

The Chess RIT yielded a large effect size (e.g., Cohen's h [on raw proportions] = 1.66). This magnitude of an effect is comparable to that of projects (e.g., Cho et al., 2016; Yankulova, Bui, & Morsella, 2018) designed to illuminate the boundary conditions of the RIT effect. In these previous projects, the effects were large (Cohen's $d > 1.60$), but not as large as that found in the basic RIT (Cohen's $d = 3.58$ [on raw proportions]; Cohen's $h = 2.38$ [on raw proportions]; Cohen's $d = 2.33$ [on arcsine transformations of the proportion data]), a project that was not designed to test the limitations of the RIT effect.

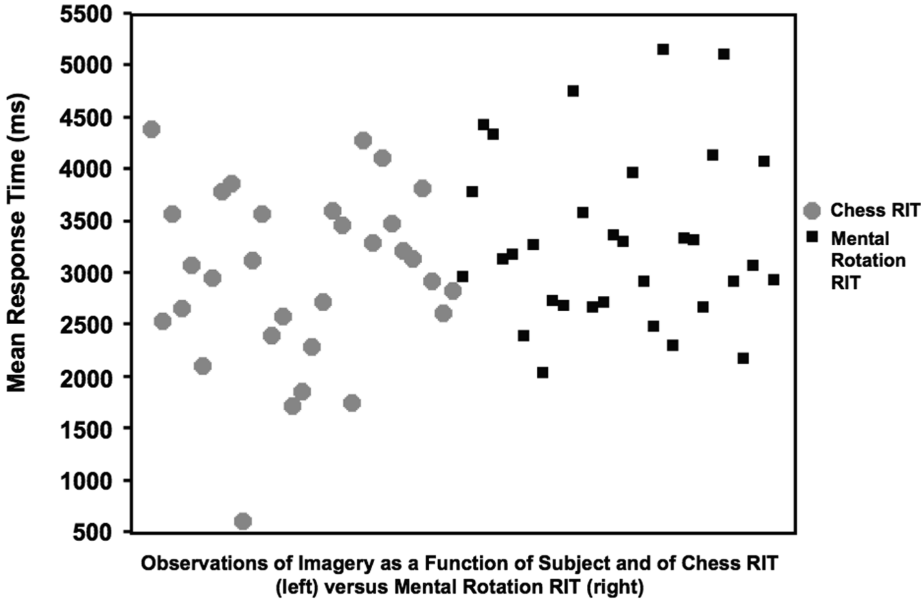


Fig. 5. Mean response time (ms) of involuntary imagery as a function of task and subject.

3.2. Mental rotation task

For this task, the RIT effect (involuntary mental rotation of the symbol) occurred on a substantive proportion of the trials (M = 0.51, SD = 0.31, SE = 0.05, Range = 0–1), a proportion that was significantly different from zero, t (32) = 9.54, p < .0001. The mean RT for these effects was 3,315.05 ms (SD = 823.54, SE = 145.58, Range = 2,036.81–5,154.67 ms). Mean accuracy was 0.45 (SD = 0.28, SE = 0.05, Range = 0–1). When analyzing only the trials in which the rotation was accurate, the mean proportion was M = 0.38 (SD = 0.25, SE = 0.04, Range = 0–0.83) and, too, significantly different from zero, t (32) = 8.68, p < .0001 (Cohen's d [on raw proportions] = 1.52; Cohen's h [on raw proportions] = 1.32). A comparable effect size is found with arcsine transformations of the proportion data (Cohen's d = 1.27).

The mean RT (Fig. 5) for these effects was 3,321.16 ms (SD = 892.99, SE = 163.04, Range = 1,962.00–5,447.33 ms). In response to the funneled debriefing question, "*Did you have a strategy for NOT mentally rotating the object? What was it?*", 27 out of 33 indicated clearly that they attempted some kind of strategy to thwart the RIT effect.

The Mental Rotation RIT yielded a large effect size (e.g., Cohen's h [on raw proportions] = 1.32). As with the effect in the Chess RIT, the magnitude of the present effect is comparable to that of projects (e.g., Cho et al., 2016: Yankulova et al., 2018) designed to illuminate the boundary conditions of the RIT effect. In these previous projects, the effects were large (Cohen's d > 1.60), but not as large as that found in the basic RIT (Cohen's d = 3.58 [on raw proportions]; Cohen's h = 2.38 [on raw proportions]; Cohen's d = 2.33 [on arcsine transformations of the proportion data]), a project that was not designed to test the limitations of the RIT effect.

3.3. Correlational analysis

A subject's mean rate (across trials) of the RIT effect in the Chess RIT correlated with the rate found in the Rotation RIT. This was the case when considering only accurate trials (r = 0.47, Fisher's r to z, p = .007) or all trials (r = 0.40, Fisher's r to z, p = .027). In short, a subject who tended to have a high proportion of involuntary entry across the trials from one task also tended to have a high proportion of involuntary entry across the trials of the other task. By-subject mean RTs from the two tasks did not correlate significantly, rs < 0.25, Fisher's r to z, ps > 0.20.

4. Discussion

The RIT has revealed how higher-order cognitions can be elicited involuntarily through the combination of experimental manipulations: the activation of sets and the presentation of external stimuli. The present data reveal that high-level processes involving symbol manipulation, too, can be controlled in this involuntary manner. In the mental rotation task, subjects were first taught to mentally rotate (30°, 60°, or 90°) two-dimensional nonsense objects. After training, subjects were instructed to not mentally rotate in these ways a different set of objects. In the chess task, subjects were taught how to move in their minds (i.e., visuospatial imagery) objects in specified ways, much as one could imagine how, in the game of chess, a given piece can navigate the chessboard. Each object was associated with a unique pattern of potential movement on a chessboard-like grid. After training, subjects were instructed to not think of where each object could move on the grid. Systematic, involuntary imagery occurred on a substantial proportion of trials for the mental rotation task (M = 0.51) and chess task (M = 0.62).

In many *response interference* paradigms (e.g., the Stroop task (Stroop, 1935), the antisaccade task (Hallett, 1978), the flanker task (Eriksen & Eriksen, 1974; Eriksen & Schultz, 1979), and the stop signal paradigm (Lappin & Eriksen, 1966)), the subject must refrain from expressing a "prepotent," dominant response to a stimulus. As in our project, in many paradigms (e.g., the flanker task and stop signal paradigm), this prepotent response is learned in the laboratory during a training session that precedes the critical test trials in which, under certain conditions, the subject must suppress the learned prepotent response.

With the present data, one could conclude that the boundary conditions of the RIT effect do not lie in mental operations requiring symbol manipulation.[9] Casual observation and some research (e.g., Cho, Zarolia, Velasquez, & Morsella, 2015) reveals that RIT effects are less likely to arise in situations involving the generation of incentive states (e.g., thirst) and emotional states (Bhangal et al., 2016). For example, it seems that, when presented with the instruction, *"Do not make yourself feel ecstatic,"* one might succeed in complying with the instruction more easily than when presented in an RIT with the instruction, *"Do not think of the name of the following object."* The effect will also not arise for basic processes associated with autonomic function (e.g., an RIT task in which the instruction is to not dilate one's pupils; Bhangal et al., 2016).

Cushing, Ghafur, and Morsella (2017) propose that this difference between these two RIT conditions reflects more than just a matter of the complexity underlying the mental operations. Rather, it has been proposed that RIT effects can arise only for activities of the *corticospinal tract* (Morsella et al., 2016a, 2016b), the tract in the nervous system that historically has been associated with voluntary action. Effects are proposed to not arise in this manner for activities of the autonomic nervous system, which is difficult to control voluntarily. Consider that *method* actors spend a great deal of effort to 'put themselves' into a certain state (e.g., to make themselves sad in order to portray a sad personage; Morsella, Larson, & Bargh, 2010).

One limitation of the present experiment is that subjects' self-reports could be inaccurate because of response bias, demand characteristics, or incorrect introspections (e.g., memory distortions; Block, 2007). However, the accuracy rates in both tasks suggest that subjects were in fact correctly reporting about their mental experiences. Future versions of the two RITs could be combined with neuroimaging technologies to provide additional, neural-based measures (e.g., activation of the neural correlates of mental rotation and visuospatial imagery) that could corroborate still further the self-reports made by subjects.

External stimuli often activate conscious contents (e.g., percepts and urges) in a direct, involuntary manner. Overt behavior, however, is often not influenced in this direct way. (Such a direct influence on overt behavior would not be adaptive, as is evident in some neurological conditions [e.g., anarchic hand syndrome; utilization behavior].) Hence, in most scenarios, inclinations triggered by external stimuli are *behaviorally suppressible*, but they are often not *mentally suppressible* (Bargh & Morsella, 2008; Morsella, 2005). According to Passive Frame Theory (Morsella et al., 2016a), this arrangement in which behavior is more suppressible than is the activation of conscious contents is evolutionarily adaptive and necessary for instrumental behavior. From this standpoint, the kind of symbol manipulation observed in the present project, though presumably more sophisticated and elaborate an act than, say, touching an object, is actually less suppressible, and more likely to arise involuntarily, than a motor act. This contrast might shed light on why habitual behaviors are more suppressible than some habitual mental processes.

In both the Chess RIT and the Mental Rotation RIT, the subject experienced the consequences of mental transformation that arose involuntarily. This finding is consistent with the views of several theorists (e.g., Brentano, 1874; Helmholtz, 1856/1925; Lashley, 1956; Miller, 1962) who propose that mental operations, whether voluntary or involuntary, are usually not introspectable and that, phenomenologically, one experiences only the products of mental operations rather than the operations themselves. Many experiences in everyday life corroborate this view (Nisbett & Wilson, 1977). In short, most entry into consciousness appears to occur involuntarily. The RIT was designed to investigate the nature and limitations of such entry. The present data suggest that the boundary conditions of the RIT effect, and, more generally, involuntary entry, do not lie in mental operations requiring a high-level executive processes such as symbol manipulation. The present findings and the theoretical views which they support have implication for various subfields of psychology, including visual imagery, perception, and cognitive control.

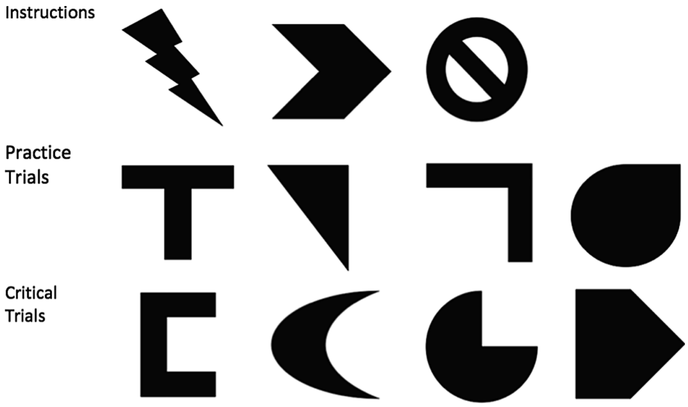
Acknowledgment

Lara Krisst conducted the pilot study in which subliminal stimuli (orthographs) were presented as the stimuli in the Reflexive Imagery Task.

Appendix A

Objects in the 0° starting orientation (mental rotation task). Not drawn to scale.

[9] Regarding boundary conditions, the RIT effect will obviously not arise for subliminal stimuli: In a pilot study ($n = 8$, trials = 8; see Acknowledgment), no RIT effects were observed with orthographs rendered subliminal through masking. In addition, the RIT effect will not arise for overt action. In Allen et al. (2013), the subjects were clearly capable of, when instructed, not overtly uttering the name of objects. This discrepancy between the control of overt naming versus covert naming is consistent with the conclusion that, though one can suppress the overt expression of behavioral inclinations, one cannot so easily suppress these inclinations mentally (Bargh & Morsella, 2008).



Appendix B. Supplementary material

Supplementary data to this article can be found online at https://osf.io/zr74s/?view_only=9bc9379f16cc46bca27b6b50323a3f95 and https://osf.io/f6wak.

References

Ach, N. (1905/1951). Determining tendencies: Awareness. In D. Rapaport (Ed.). *Organization and pathology of thought* (pp. 15–38). New York: Columbia University Press (Original work published in 1905.).

Allen, A. K., Krisst, L., Montemayor, C., & Morsella, E. (2016). Entry of involuntary conscious contents from ambiguous images. *Psychology of Consciousness: Theory, Research, and Practice, 3*, 326–337.

Allen, A. K., Wilkins, K., Gazzaley, A., & Morsella, E. (2013). Conscious thoughts from reflex-like processes: A new experimental paradigm for consciousness research. *Consciousness and Cognition, 22*, 1318–1331.

Bargh, J. A., & Chartrand, T. L. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis, & C. M. Judd (Eds.). *Handbook of research methods in social and personality psychology* (pp. 253–285). Cambridge, England: Cambridge University Press.

Bargh, J. A., & Morsella, E. (2008). The unconscious mind. *Perspectives on Psychological Science, 3*, 73–79.

Bhangal, S., Cho, H., Geisler, M. W., & Morsella, E. (2016). The prospective nature of voluntary action: Insights from the reflexive imagery task. *Review of General Psychology, 20*, 101–117.

Bhangal, S., Merrick, C., Cho, H., & Morsella, E. (2018). Involuntary entry into consciousness from the activation of sets: Object counting and color naming. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2018.01017.

Bhangal, S., Merrick, C., & Morsella, E. (2015). Ironic effects as reflexive responses: Evidence from word frequency effects on involuntary subvocalizations. *Acta Psychologica, 159*, 33–40.

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences, 30*, 481–548.

Brentano, F. (1874). *Psychology from an empirical standpoint.* Oxford, UK: Oxford University Press.

Cho, H., Godwin, C. A., Geisler, M. W., & Morsella, E. (2014). Internally generated conscious contents: Interactions between sustained mental imagery and involuntary subvocalizations. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.01445.

Cho, H., Zarolia, P., Velasquez, A., & Morsella, E. (2015). Cognitive-versus emotion-based involuntary cognitions: An informative contrast for the reflexive imagery task. *Poster presented at the Annual Convention of the Association for Psychological Science, New York*.

Cho, H., Zarolia, P., Gazzaley, A., & Morsella, E. (2016). Involuntary symbol manipulation (Pig Latin) from external control: Implications for thought suppression. *Acta Psychologica, 166*, 37–41.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments, & Computers, 25*, 257–271.

Cushing, D., Gazzaley, A., & Morsella, E. (2017). Externally controlled involuntary cognitions and their relations with other representations in consciousness. *Consciousness and Cognition, 55*, 1–10.

Cushing, D., Ghafur, R. D., & Morsella, E. (2017). The interdependence between conscious and unconscious processes. In Z. Radman (Ed.). *Before consciousness: In search of the fundamentals of mind* (pp. 50–82). UK: Imprint Academic.

Desender, K., van Opstal, F. V., & van den Bussche, E. (2014). Feeling the conflict: The crucial role of conflict experience in adaptation. *Psychological Science, 25*, 675–683.

Dostoevsky, F. (1863/2008). *Winter notes on summer impressions.* London: One-world Classics.

Dou, W., Li, Y., Geisler, M. W., & Morsella, E. (2018). Involuntary polymodal imagery involving olfaction, audition, touch, taste, and vision. *Consciousness and Cognition, 62*, 9–20.

Enger, T. (2017). *The Wiley handbook of cognitive control.* West Sussex, UK: John Wiley and Sons.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics, 16*, 143–149.

Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception and Psychophysics, 25*, 249–263.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences, 8*, 162–169.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

Fiedler, K. (2017). What constitutes strong psychological science? The (neglected) role of diagnosticity and a priori theorizing. *Perspectives on Psychological Science, 12*, 46–61.

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for 'top-down' effects. *Behavioral and Brain Sciences, 39*, 1–77 [Target Article].

Garcia, A. C., Bhangal, S., Velasquez, A. G., Geisler, M. W., & Morsella, E. (2016). Metacognition of working memory performance: Trial-by-trial subjective effects from a new paradigm. *Frontiers in Psychology, 7*, 927. https://doi.org/10.3389/fpsyg.2016.00927.

Gazzaley, A., & D'Esposito, M. (2007). Unifying prefrontal cortex function: Executive control, neural networks and top-down modulation. In B. Miller, & J. Cummings (Eds.). *The human frontal lobes: Functions and disorders* (pp. 187–206). New York: Guilford Press.
Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist, 54*, 493–503.
Goodhew, S. C. (2017). What have we learned from two decades of object substitution masking? Time to update: Object individuation prevails over substitution. *Journal of Experimental Psychology: Human Perception & Performance, 43*(6), 1249–1262.
Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research, 18*, 1279–1296.
Helmholtz, H.v. (1856/1925). Treatise of physiological optics: Concerning the perceptions in general. In T. Shipley (Ed.). *Classics in psychology* (pp. 79–127). New York: Philosophy Library.
Lacey, S., & Lawson, R. (2013). *Multisensory imagery.* New York, NY: Springer.
Lappin, J. S., & Eriksen, C. W. (1966). Use of a delayed signal to stop a visual reaction time response. *Journal of Experimental Psychology, 72*, 805–811.
Lashley, K. S. (1956). Cerebral organization and behavior. *Proceedings of the Association for Research in Nervouse and Mental Diseases, 36*, 1–18.
Lewin, K. (1935). *A dynamic theory of personality.* New York: McGraw-Hill.
Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes, 65*, 272–292.
Mason, M. F., Norton, M. I., van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science, 315*, 393–1345.
McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological Bulletin, 136*, 188–207.
Merrick, C., Cooper, S., Jantz, T. K., & Morsella, E. (2013). Working memory based action control: Increased refreshing rate in the face of interference. Poster presented at the Annual California Cognitive Science Conference, Berkeley.
Merrick, C., Farnia, M., Jantz, T. K., Gazzaley, A., & Morsella, E. (2015). External control of the stream of consciousness: Stimulus-based effects on involuntary thought sequences. *Consciousness and Cognition, 33*, 217–225.
Miller, B. L., & Cummings, J. L. (2007). *The human frontal lobes: Functions and disorders* (second ed.). New York: Guilford Press.
Miller, G. A. (1962). *Psychology: The science of mental life.* New York: Adams, Bannister, & Cox.
Mitchell, J. P., Heatherton, T. F., Kelley, W. M., Wyland, C. L., Wegner, D. M., & Macrae, C. N. (2007). Separating sustained from transient aspects of cognitive control during thought suppression. *Psychological Science, 18*, 292–297.
Morsella, E. (2005). The function of phenomenal states: Supramodular interaction theory. *Psychological Review, 112*, 1000–1021.
Morsella, E., Godwin, C. A., Jantz, T. J., Krieger, S. C., & Gazzaley, A. (2016a). Homing in on consciousness in the nervous system: An action-based synthesis. *Behavioral and Brain Sciences, 39*, 1–17 [Target Article].
Morsella, E., Godwin, C. A., Jantz, T. K., Krieger, S. C., & Gazzaley, A. (2016b). Passive frame theory: A new synthesis. *Behavioral and Brain Sciences, 39*, 44–70.
Morsella, E., Gray, J. R., Krieger, S. C., & Bargh, J. A. (2009). The essence of conscious conflict: Subjective effects of sustaining incompatible intentions. *Emotion, 9*, 717–728.
Morsella, E., Larson, L. R. L., & Bargh, J. A. (2010). Indirect cognitive control, working-memory-related movements, and sources of automatisms. In E. Morsella (Ed.). *Expressing oneself / expressing one's self: Communication, cognition, language, and identity* (pp. 61–90). London: Psychology Press.
Morsella, E., Wilson, L. E., Berger, C. C., Honhongva, M., Gazzaley, A., & Bargh, J. A. (2009). Subjective aspects of cognitive control at different stages of processing. *Attention, Perception, & Psychophysics, 71*, 1807–1824.
Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231–259.
Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia II: Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science, 7*, 615–631.
Pasley, B. N., David, S. V., Mesgarani, N., Flinker, A., Shamma, S. A., Crone, N. E., ... Chang, E. F. (2012). Reconstructing speech from human auditory cortex. *PLoS Biology, 10*(1), e1001251.
Questienne, L., Atas, A., Burle, B., & Gevers, W. (2018). Objectifying the subjective: Building blocks of metacognitive experiences in conflict tasks. *Journal of Experimental Psychology: General, 147*, 125–131.
Spence, C., & Deroy, O. (2013). Crossmodal mental imagery. In S. Lacey, & R. Lawson (Eds.). *Multisensory imagery* (pp. 157–183). New York, NY: Springer.
Rassin, E. (2005). *Thought suppression.* Amsterdam, The Netherlands: Elsevier.
Rumelhart, D. E., McClelland, J. L., & the PDP Research Group (1986). Parallel distributed processing: Explorations in the microstructure of cognition, Vols. 1 and 2. Cambridge, MA: Massachusetts Institute of Technology.
Schacter, D. L., & Tulving, E. (1994). *Memory systems.* Cambridge, MA: The MIT Press.
Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.
Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136.
Uznadze, D. (1966). *The psychology of set.* New York: Consultants Bureau.
Wegner, D. M. (1989). *White bears and other unwanted thoughts.* New York: Viking/Penguin.
Wegner, D. M. (1994). Ironic processes of thought control. *Psychological Review, 101*, 34–52.
Wyland, C. L., Kelley, W. M., Macrae, C. N., Gordon, H. L., & Heatherton, T. F. (2003). Neural correlates of thought suppression. *Neuropsychologia, 41*, 1863–1867.
Yankulova, J. K., Bui, N-C. T., & Morsella, E. (2018). Involuntary insight, judgments, and syntactic processing and their experimentally-controlled determinants. Poster presented at the Annual Convention of the Association for Psychological Science, San Francisco.