Consciousness and Cognition 55 (2017) 1–10

Contents lists available at ScienceDirect

Consciousness and Cognition

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

Externally controlled involuntary cognitions and their relations with other representations in consciousness

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:
Consciousness
Voluntary action
Passive frame theory
Reflexive imagery task

ABSTRACT

Percepts and action-related urges often enter consciousness insuppressibly. The Reflexive Imagery Task (RIT) was developed to investigate how high-level cognitions (*e.g.*, subvocalizations), too, can enter consciousness in this manner. Limitations of the paradigm include (*a*) that no data have confirmed subjects' introspections about the involuntary subvocalizations, and (*b*) that, in everyday life, adaptive responses to involuntary cognitions often depend on the nature of the other contents in consciousness. To address *a* and *b*, we developed an RIT in which subjects were presented with visual objects and instructed to not think of the object names. If a subvocalization did arise, however, subjects responded motorically only if the subvocalization rhymed with a word held in memory and if there was a visual "go" cue. Subjects successfully (on 0.83 of the trials) emitted this complex, "multi-determined" response, which provides evidence for the occurrence of the involuntary subvocalizations and illuminates the function of consciousness.

1. Introduction

It is a fact of everyday experience that percepts and urges often enter one's consciousness involuntarily: The eyes open and a visual world, replete with objects and colors, is perceived instantaneously, effortlessly, and regardless of one's desires. Later in the day, one might experience, say, the desire to have a cup of coffee. To the thinker, such *conscious contents*[1] "just happen"—involuntarily and, most often, insuppressibly.

Recent theorizing (see review in Morsella, Godwin, Jantz, Krieger, & Gazzaley, 2016a) and experimental paradigms reveal that high-level cognitions, too, can arise in this involuntary and insuppressible manner. For example, in the Reflexive Imagery Task (RIT; see review in Bhangal, Cho, Geisler, & Morsella, 2016), high-level conscious thoughts are triggered involuntarily, through external control stemming from the actions of the experimenter. The RIT is based on the experimental approaches of Ach (1905/1951), Wegner (1989), and Gollwitzer (1999). It was developed to investigate experimentally the unconscious generation of high-level, involuntary conscious contents and to complement the theoretical standpoint that action inclinations experienced in consciousness

* 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] In this treatise, we are speaking about the most basic form of consciousness, such as the awareness of the color blue, of an afterimage, or of the urge to cough. From this standpoint, a "conscious content" is something that one is aware of (Merker, 2007). For example, it might be a color, an urge, or an earworm, which is a song that one "cannot get out of one's mind." The "conscious field" is all that one is aware of at one moment in time, which is the combination of all activated conscious contents (Freeman, 2004; Köhler, 1947; Searle, 2000).

http://dx.doi.org/10.1016/j.concog.2017.07.003
Received 5 October 2016; Received in revised form 3 July 2017; Accepted 9 July 2017
Available online 20 July 2017

are often mentally insuppressible (Bargh & Morsella, 2008).

In the initial and most basic version of the task (Allen, Wilkins, Gazzaley, & Morsella, 2013), subjects were instructed to not subvocalize (*i.e.*, say in their head but not aloud) the names of objects (*e.g.*, line drawings from Snodgrass & Vanderwart, 1980). (*Subvocalization* occurs when one 'talks in one's head' or names an object in such a manner; this mental representation has been referred to as the 'pre-articulatory output'; Slevc & Ferreira, 2006.) In Allen et al. (2013), subjects were presented before each trial with the instruction, "Don't Think of the Name of the Object" before an object was presented for 4 s, during which time subjects indicated by button press if they happened to have subvocalized the name of the object involuntarily. On the majority of the trials (0.86 in Allen et al., 2013; 0.87 in Cho, Godwin, Geisler, & Morsella, 2014; and 0.73 in Merrick, Farnia, Jantz, Gazzaley, & Morsella, 2015), subjects failed to suppress such subvocalizations.

In a more complex version of the task (Merrick et al., 2015), subjects were instructed to (*a*) not subvocalize the name of the visual object, and (*b*) not subvocalize the number of letters in the object name. On a significant proportion of the trials (0.30 [$SE = 0.04$]), subjects reported experiencing both kinds of imagery. Each kind of imagery arose from distinct, high-level processes (*i.e.*, that of object naming versus letter counting). In another complex version of the task, RIT effects arose even though the effect involved a word-manipulation task similar to Pig Latin (*e.g.*, "CAR" becomes "AR-CAY"). In this variant of the RIT (Cho, Zarolia, Gazzaley, & Morsella, 2016), subjects were instructed to not transform stimulus words in this way, but involuntary transformations still arose on more than a proportion of 0.40 of the trials. This effect is striking because the involuntary transformation requires complex symbol manipulation, which is known to be associated with high-level processes in frontal cortex (Miller & Cummings, 2007).

To illustrate the basic version of the RIT effect, momentarily, we will present to you an object enclosed within parentheses. Your task is to *not* subvocalize the name of the object. Here is the stimulus (▲). When presented with this instruction (which induces a certain *action set*) and then presented with this stimulus, most people cannot suppress the conscious experience of the phonological form of the word "triangle." This RIT effect requires the sophisticated process of object naming, in which only one of tens of thousands of phonological representations is selected for production in response to a visual stimulus (*e.g.*, CAT yields /k/, /œ/, and /t/; Levelt, 1989). After the presentation of the stimulus, the RIT effect arises after a few moments ($M = 1451.27$ ms [$SD = 611.42$] in Allen et al., 2013; $M = 2323.91$ ms [$SD = 1183.01$] in Cho et al., 2014; $M = 1745.97$ ms [$SD = 620.86$] in Merrick et al., 2015).

It is important to appreciate that the RIT provides an additional method to contrast the capacities of conscious and unconscious processes. This method does not involve subliminal stimuli, which are problematic: Subliminal stimuli are not only unconscious, but they are also of very weak strength, unlike the supraliminal representations that unconscious mechanisms often process (Bargh & Morsella, 2008). Thus, the RIT involves what can be regarded as the Helmholtzian-Freudian unconscious, which operates over supraliminal stimuli (see related account in Nisbett & Wilson, 1977). With this in mind, it is important to consider the observation that, in a pilot study ($n = 8$, trials $= 8$; see Acknowledgment), no RIT effects were observed when the stimuli (orthographs) were rendered subliminal through visual masking. Hence, it seems that the automatic and unconscious processes giving rise to the RIT effect require that the stimuli be supraliminal.

1.1. Evidence that the RIT effect resembles a reflex

Empirical evidence and theory (including Wegner's (1994) model of *ironic processing*; see discussion of relationship between Wegner's (1994) model and the RIT in Bhangal, Cho, et al., 2016) suggest that, for subjects, the effect "just happens" and is not an artifact of high-level strategic processes on their part. Corroborating this conclusion, in one version of the RIT, subjects reported on the majority of trials that the involuntary subvocalization felt "immediate" (Bhangal, Merrick, & Morsella, 2015). There is other evidence supporting the view that the effect is not an artifact of strategic processes. First, the nature of the subvocalizations is influenced systematically by factors such as word frequency (Bhangal et al., 2015). Such an artifact of experimental demand would require subjects to know how word frequency should influence responses in an experiment. Second, on some trials, the effect arises too quickly to be caused by strategic processing (Allen et al., 2013; Cho et al., 2014). Third, the effect still arises under conditions of cognitive load, in which it is difficult for subjects to implement strategic processing (Cho et al., 2014). Fourth, the RIT effect habituates (*i.e.*, is less likely to arise) after repeated presentation of the same stimulus object, which suggests that the RIT effect is activated in a reflex-like manner and involuntarily (Bhangal, Allen, Geisler, & Morsella, 2016).

1.2. Two limitations of the RIT

The RIT effect reveals that high-level conscious contents can enter the conscious field in an insuppressible, non-trivial, and reflex-like manner. Although the evidence accumulated thus far suggests that the effect is both robust and reliable, some important questions remain concerning the validity of the effect. For instance, one major criticism of the RIT is that it relies on the technique of self-report. This technique has well-known limitations. For example, self-reports can be inaccurate as a result of subjects basing their reports on a strategy of how to comport oneself during an experiment (see discussion in Morsella et al., 2009). In addition, inaccurate memories of fleeting conscious contents can lead to incorrect self-reports (Block, 2007). Evidence from neuroimaging studies in which subjects must report about the occurrence of involuntary conscious contents suggests that subjects are not confabulating about the occurrence of these mental events (Mason et al., 2007; McVay & Kane, 2010; Mitchell et al., 2007; Pasley et al., 2012; Wyland, Kelley, Macrae, Gordon, & Heatherton, 2003). However, this neuroimaging-based evidence stems from tasks that are different in significant ways from the RIT. For instance, the dependent measures in these paradigms do not involve the involuntary naming (subvocally) of stimulus objects. In addition, the neural correlates of subvocalizing remain controversial (Hickok, 2009; Schomers, Kirilina, Weigand, Bajbouj, & Pulvermüller, 2014; see discussion in Bhangal, Cho, et al., 2016). Hence, at this stage of understanding, neural evidence alone cannot be used as an unequivocal index of the occurrence of involuntary subvocalizing in the RIT.

Another limitation pertains to ecological validity. In everyday life, involuntary contents, such as those of the RIT and urges, occupy the conscious field along with other contents activated by, for example, the stimulus scene or memory. An adaptive response to an involuntary content must often be based in part on the nature of the other contents composing the conscious field at that time. (Several theoretical frameworks propose that this bringing together of diverse sources of information to yield adaptive action is an important role of conscious processing [*e.g.*, Baars, 1988; Dehaene, 2014; Norman & Shallice, 1980].)[2] It is usually within such a multifaceted context—one having many "moving parts"—that one must respond adaptively to an involuntarily-generated content. The RIT fails to capture such *inter-representational dynamics* (Morsella, Zarolia, & Gazzaley, 2012) underlying adaptive responses to involuntary cognitions. Can the RIT effect arise in such a dynamic and complex (and more ecologically-valid) context, and, in such a context, can subjects respond adaptively to the involuntary effect? To address this question and also provide much needed evidence that corroborates subjects' self-reports about the occurrence of the RIT effect, we developed a new variant of the RIT.

1.3. The current approach

In the present variant of the RIT, subjects indicated by button press the basic RIT effect but, in addition, they had to press another button if and only if (*a*) the subvocalization rhymed with a word held in mind ("stir"), and (*b*) there was a solid border around the image (as opposed to a dotted border). We developed *a* to address our first limitation regarding the accuracy of subjects' self-reports. The rhyme task provides evidence that subjects experience involuntary subvocalizations of the object names, for detecting a rhyme requires the retrieval of either the whole object name or, at minimum, the coda of the object name. We developed *b* to address the second limitation mentioned above. To this end, in our computer-based task, the subject must perform a response (a button press) toward a visual stimulus (a line drawing of an everyday object) only if several conditions are met. First, the subject must experience an involuntary high-level cognition (the involuntary subvocalization of the object name, which is the RIT effect). Second, the subject must (*a*) detect that the involuntary subvocalization happens to rhyme with a word held in memory and (*b*) perceive a certain visual stimulus (a black frame surrounding the visual object). Thus, in our task, the correct response to the involuntary cognition depends in part on the nature of the other conscious contents composing the field at that time. The accurate response with the second button (henceforth "the multi-determined response") requires a host of polymodal processing (*e.g.*, subvocalization, prospective memory, and visual processing). Many laboratory tasks demand a multi-determined response, but few tasks have as the target of action a high-level, involuntary cognition (*i.e.*, the RIT effect) that arises in consciousness through external control.

Can the RIT effect survive in such a complex scenario? Would the effect arise on a large proportion (> 0.70) of the trials, as was found in previous studies (*e.g.*, 0.86 in Allen et al., 2013; 0.87 in Cho et al., 2014; and 0.73 in Merrick et al., 2015)? If so, can subjects perform the task successfully and, when appropriate, emit successfully the multi-determined response? For example, can they emit this response accurately on a substantive proportion of the trials (~0.80; Buccafusco, 2001)? (We should add that, before the commencement of data collection, several scientists informed one of us [EM] that subjects would find it challenging to emit the multi-determined response and would yield only low rates of success [< 0.30 accuracy across trials].)

If subjects can perform this task, then that would be very interesting, for the task requires that the following six component processes transpire successfully. First, the phonological form of the word "stir" (or, minimally, the coda of the word) must somehow be held in memory from the beginning of the trial until the occurrence of the multi-determined response. Second, the visual object must be perceived. Third, on trials in which an RIT effect arises, the visual object must be processed by the semantic system. Fourth, on trials having a correct multi-determined response, the entire phonological representation of the object name must be activated (Levelt, 1989). Fifth, on these critical trials, the subject must perceive whether the frame surrounding the object is solid or dotted. Sixth, for all button presses, there must be accurate perception-to-action response mappings.

The observation of successful performance on a substantive proportion of the trials (that is, ~0.80; Buccafusco, 2001) implies that the experimental paradigm can be employed to further mine experimentally each of the six component tasks as well as the complex interactions among them. Our scientific approach was influenced by Tolman's (1948) experimental strategy for corroborating the existence of the cognitive map. In his classic experiments, if the rat was capable of solving a particular maze under certain controlled conditions (*e.g.*, the rat could not solve the maze by relying on external cues), then it must be the case that several component processes, including the mental representation of a cognitive map, must be at play. Similarly, in our approach, we reasoned that, if subjects could perform this task, then the six component processes must be at play. The conscious field seems necessary for at least one of the processes—the RIT effect, for the effect does not arise for subliminal stimuli (see Acknowledgment). Future variants of our task could be used to further investigate each of the component processes.

[2] For example, in one account (Morsella & Bargh, 2011), when the conscious field is operating normally, it yields "integrated behavior," a kind of behavior which is influenced by multiple types of information and inclinations. This behavior occurs, for example, when one is underwater and has the urge to inhale but suppresses this response, or when one is carrying a hot dish of food and experiences both the urge to drop the dish and to refrain from dropping it (Morsella, 2005). It also occurs when, while driving, one is faced with a yellow traffic light and must decide whether to begin to speed up or to brake, which often depends on the context (Morsella, Godwin, Jantz, Krieger, & Gazzaley, 2016b). Without consciousness, behavior (including sophisticated motor programming) can arise but the behavior will not be integrated. For example, in neurological conditions in which actions are decoupled from consciousness and arise involuntarily, sophisticated actions such as manipulating tools or removing clothing can arise but the actions are not influenced by all the kinds of information by which they should be influenced (Morsella & Bargh, 2011). These actions are, in a sense, "un-integrated," thereby appearing insensitive to context.

1.4. The research approach is not the typical, hypothesis-testing experiment

We developed a multifaceted and theoretically-informed task that could yield a reliable multi-determined response, one which could be investigated further in future projects. Unlike in standard "hypothesis-testing" experiments, our goal was not simply to assess whether new data falsify or fail to falsify some hypothesis. Instead, we set out to determine whether subjects can perform this multifaceted task, one which is suitable for further, in-depth investigation. Thus, the research project, involving a "proof of concept," or, more accurately, a "proof of mechanism," is not the standard "hypothesis-testing" experiment. We should add that the present paradigm (*a*) builds systematically on previous research, (*b*) includes stimuli (*e.g.*, object names) about which much is known (Miller, 1996), and (*c*) yields a robust and reliable effect (the RIT effect) that is predicted by several frameworks (*e.g.*, by Gollwitzer, 1999; Miller, 1959; Morsella, 2005; Wegner, 1989). In summary, it is the kind of *cumulative, theory-driven* research that leaders in the field of experimental psychology (*e.g.*, Nosek, Spies, & Motyl, 2012) have strongly encouraged (see recent discussion in Fiedler, 2017). The task also provides a way of examining the mechanisms underlying entry into consciousness, one of the greatest enigmas in science (Crick & Koch, 2003). The many features of our task would be of interest to researchers in the fields of cognitive control, systems neuroscience, action control, prospective memory, mental imagery, and psychopathology, in which involuntary cognitions can be debilitating.

To conclude, we predicted that our new task would (*a*) provide strong, much needed evidence for subjects' self-reports about the RIT effect, (*b*) reveal that the RIT effect can arise at high rates even in a complex scenario that features the kinds of inter-representational dynamics which often occur outside of the laboratory, and (*c*) subjects will be able to carry out the multi-determined response with high rates of accuracy (a proportion of trials *circa* 0.80).

2. Method

2.1. Subjects

San Francisco State University students ($n = 33$; $M_{age} = 21.27$, $SD_{age} = 4.41$, females = 22) participated for course credit. The Institutional Review Board at San Francisco State University approved the involvement of human subjects in our project. All subjects were proficient at speaking, reading, and understanding English. There was not one case in which a subject's primary language was one which reads from right to left.

2.2. Stimuli and apparatus

Stimulus objects were presented on an Apple iMac computer monitor with a 508 mm screen with a seated viewing distance of approximately 480 mm. The stimulus for the initial practice trial was a black-and-white line drawing of a bird. The stimulus (35 × 30 mm, with a subtended visual angle of 4.18° × 3.58°) was presented in the center of the screen. This image was taken from the Snodgrass Image List (Snodgrass & Vanderwart, 1980). Images from the Snodgrass list have been used successfully in previous research employing the RIT. Stimulus presentation was controlled by PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). All questions and instructions were written in black 36-point Helvetica font on a white background.

The stimuli for the subsequent practice trials were ten black-and-white line drawings of well-known objects. These images were similar to those from the Snodgrass Image List. The size of each of these ten images was equal to or less than 55 × 40 mm. The subtended visual angle of each image was equal to or less than 6.56° × 4.77° (Appendix). The stimuli for the critical trials were 50 black-and-white line drawings of well-known objects taken from the Snodgrass Image List. All stimuli were displayed in the center of the screen. The size of each image was equal to or less than 80 × 70 mm. The subtended visual angle of each image was equal to or less than 9.53° × 8.34° (Appendix). Object names for the stimuli in subsequent practice trials had between two and four syllables and between five and eleven letters. These stimuli were presented in random order. With the exception of one stimulus (screwdriver), these stimuli were never presented again during the experimental session. The object names for the six stimuli in the rhyme condition had between five and eleven letters, and between two and four syllables. The object names for the four stimuli in the non-rhyme condition had between five and nine letters and between two and three syllables.

Object names (see Appendix for list) for the standard trials had between four and fourteen letters, and between one and five syllables. Object names ($n = 25$) in the rhyme condition rhymed with the word "stir." Objects in the non-rhyme condition ($n = 25$) did not rhyme with the word "stir." Regarding word-length and syllable-length, the names associated with the rhyme condition were comparable to those in the non-rhyme condition, $ps > 0.70$. Francis-Kucera (Francis & Kucera, 1982) word frequency counts (from Snodgrass & Vanderwart, 1980) did not differ significantly between the words in the rhyme condition and those in the non-rhyme condition, $p > 0.40$. Surrounding the image of each object was a square solid border or a square dotted border (Fig. 1). Each border was 380 × 255 mm, with a subtended visual angle of 43.19° × 29.75°. The borders were 3 mm thick. Each image was presented twice, once with each kind of border, totaling 100 critical trials per subject.

2.3. Procedures

Each subject was run individually. Before the start of the first trial, subjects were instructed to keep their eyes focused on the center of the screen. Subjects were instructed to not think of the name of the object that, on each trial, would be presented to them. If, however, they happened to think of the name of the object, then they had to press the "z" key. This button was covered with a pink

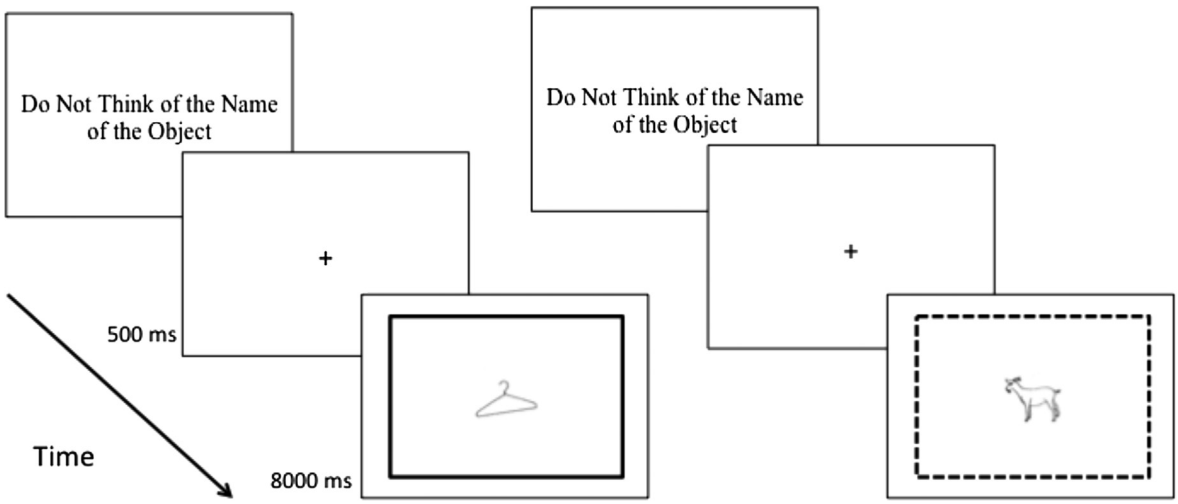

Fig. 1. Schematic depiction of a critical trial (left) and of a trial that should not yield the multi-determined response (right). Not drawn to scale.

paper overlay to minimize confusion. Instructions presented to the subject referred to the "z" key as the "pink" key. Subjects were instructed to keep their left index finger rested on the pink key at all times throughout the trials and were informed that the image of an object would be shown for a fixed amount of time, regardless of whether or not the pink key was pressed.

After subjects indicated that they understood the instructions, they were invited to press the space bar to advance to the first trial, which was the initial practice trial. The sentence *"Do not think of the name of the object"* was presented in the center of the screen. This instruction served as a "ready" prompt. Subjects pressed the space bar when ready. Following this button press, a fixation cross (+) was presented for 500 ms. Next, the line drawing image of a bird was presented for 8 s. The ready prompt, the fixation cross, and the image were all presented in the center of the screen (Fig. 1). Surrounding the bird was a solid border. Importantly, subjects were not told in advance that there would be any border surrounding the image. After the image was presented for 8 s, subjects were asked if they noticed a border surrounding the image. They answered this question by pressing either "y" or "n" on the keyboard and then pressing the space bar. The proportion of subjects who responded in the affirmative to this question was 0.63.

For the subsequent practice trials ($n = 20$), subjects were instructed that, if they thought of the name of the object, they should press the pink key, just as before. However, subjects were also informed that, if they did happen to think of the name of the object and the object happened to rhyme with the word "stir," and if there was a solid border around the image, they should also press a second key (a green key). (The "/" key was covered with a green paper overlay.) The solid border appeared on half of the trials. For the remaining half of the trials, a dotted border was presented. Each object appeared once with each kind of border. In these and in subsequent trials, whether the object rhymed with "stir" or not, and whether the border was solid or dotted, was random.

Subjects kept their right index fingers rested on the green key at all times throughout the experiment. Response times (RTs) for each key press were recorded. (We selected the "z" and "/" keys because these keys are equidistant from the space bar on the keyboard.) Subjects pressed the space bar when they finished reading the instructions and were ready to begin the practice trials. To ensure that subjects understood the task, the experimenter stood nearby each subject during the practice trials and pointed out whenever errors were made. Importantly, during the practice trials, subjects learned to emit the critical, multi-determined response on six of the twenty trials.

After the conclusion of the practice session, subjects performed the 100 trials. To commence each trial, subjects pressed the space bar. As in previous trials, following this button press, a fixation cross (+) was presented for 500 ms. Next, the image of an object was presented for 8 s. Subjects were aware that the image would be presented for this amount of time regardless of whether or not they pressed any keys. As in the practice trials, the ready prompt, the fixation cross, and the image were all presented in the center of the screen. After the image of an object was presented, subjects pressed the pink key if they happened to think of the name of the object. If that object name rhymed with "stir," and if the border around the image was a solid border, they had to also press the green key. To ensure that subjects would not forget the instructions regarding when to press the green key, the instructions were presented on the computer screen four times during the block of 100 trials, thereby appearing every 25 trials. Trials in which RTs for a button press were less than or equal to 200 ms were excluded from analysis. There was only one such trial.

At the conclusion of the trials, subjects completed a paper-and-pencil, funneled debriefing questionnaire (following the procedures of Bargh & Chartrand, 2000). The questionnaire was designed to identify subjects whose data should not be included in the statistical analyses. The questions assessed whether (*a*) subjects were aware of the purpose of the study, (*b*) subjects had any strategies for completing the task, (*c*) anything interfered with subjects' performance on the task, (*d*) subjects were able to remember the rhyme target during the task, and (*e*) subjects had difficulty with the English language. The information gathered from the funneled debriefing revealed that it was unnecessary to exclude any trial data from our analyses. On the basis of other considerations, the data from six subjects were excluded from data analysis because either the subjects did not follow directions or because the computer malfunctioned.

3. Results

The mean proportion of RIT effects (involuntary subvocalizations) across all trials was 0.87 (SD = 0.14, SE = 0.03), which replicates previous findings (Allen et al., 2013; Bhangal, Cho, et al., 2016; Bhangal et al., 2015; Cho et al., 2014, 2016; Merrick et al., 2015). This proportion was significantly different from zero, $t\,(26) = 32.49$, $p < 0.0001$. The same significant result was found with arcsine transformations of the proportion data, $t\,(26) = 30.77$, $p < 0.0001$. (Arcsine transformations are often used to statistically normalize data that are in the form of proportions.) Of the 27 subjects, 14 had an RIT effect on over 0.90 of the trials; 8 had an RIT effect on 0.80–0.90 of the trials; and the proportions for the remaining 5 subjects were 0.74, 0.71, 0.64, 0.62, and 0.43. The mean latency of the RIT effect (*i.e.*, the time between the presentation of the visual stimulus and the subject's first press [*i.e.*, the pink key]) was 2225.14 ms (SD = 695.14, SE = 133.78).

For the critical trials (*i.e.*, the trials in which the word rhymed with the target, and the frame around the stimulus was a solid line), the RIT effect (proportion) was 0.92 (SD = 0.13, SE = 0.03), with a mean latency of 2041.73 ms (SD = 609.82, SE = 119.60). The mean accuracy across these critical trials was 0.83 (SD = 0.14, SE = 0.03), a level of accuracy which is significantly different from zero, $t\,(26) = 30.92$, $p < 0.0001$. The same result was found with arcsine transformations of the proportion data, $t\,(26) = 28.99$, $p < 0.0001$. Of the 27 subjects, 11 had an accuracy rate of over 0.90; 6 had a proportion of 0.80–0.90; 6 had rates from 0.70 to 0.79; and the rates for the remaining 4 subjects were 0.68, 0.63, 0.52, and 0.52. The mean latency of these correct responses (involving the green key) was 2584.77 ms (SD = 661.12, SE = 127.23).

On the non-critical trials, in which the three conditions were not met and subjects should not have executed the second button press, the mean rate of incorrect button presses was 0.03 (SD = 0.03, SE = 0.01). On this small proportion of trials, the mean proportion of RIT effects across trials was 0.85 (SD = 0.15, SE = 0.03), with a mean latency of 2283.68 ms (SD = 755.17, SE = 145.33).

4. Discussion

The data revealed (*a*) strong evidence for subjects' self-reports concerning the RIT effect, (*b*) that, despite the complexity and many "moving parts" involved in our new task, the RIT effect still arose (> 0.80 of the trials; replicating Allen et al., 2013; Bhangal, Cho, et al., 2016; Bhangal et al., 2015; Cho et al., 2014, 2016; Merrick et al., 2015), and (*c*) that subjects were able to emit successfully (0.83 accuracy across trials) the critical, multi-determined response. The observation noted in *b* corroborates that the RIT effect is both a robust and reliable phenomenon (see review of findings in Bhangal, Cho, et al., 2016). Regarding *c*, it should be noted that, in many established laboratory tasks (*e.g.*, *response interference paradigms*), there is the elicitation of a multi-determined response that affects the conscious field in some way (*e.g.*, as in the case of action-related urges; see review in Morsella, Berger, & Krieger, 2011). However, in contrast to most of these tasks, our variant of the RIT had as the object of intended action a high-level, involuntary cognition (*i.e.*, the RIT effect) that was elicited reliably through external control. Concerning the evidence for subjects' self-reports about the RIT effect, our rhyme task requires the retrieval of either the whole object name or, at a minimum, the coda of the object name. As discussed above, this behavioral (and much needed) evidence for the occurrence of the RIT effect is more compelling than any data that are based on neuroimaging technologies.

To thwart the RIT effect, subjects may be using various strategies, including that of self-distraction (Wegner, Schneider, Carter, & White, 1987). Referring to an experimental finding by Hertel and Calcaterra (2005), Bulevich, Roediger, Balota, and Butler (2006) state that "suppression instructions to not think of an unwanted response may succeed if subjects are given the strategy (or themselves hit upon the strategy) of always thinking of some other item when they are trying to suppress an unwanted response" (p. 1575). In the original RIT study (Allen et al., 2013), the funneled debriefing data revealed that 8 subjects attempted to suppress the RIT effect by subvocalizing other names or by thinking about something else (*e.g.*, subvocalizing a melody). Some subjects reported that they attempted to succeed at the task by focusing on parts of the object and not focusing on the object as a whole.

The answers to the questions from the funneled debriefing questionnaire from the present study reveal that eleven subjects reported that, to thwart the RIT effect, they did not employ any kind of strategy. (One must be conservative when interpreting the results from this funneled debriefing questionnaire because the questionnaire was designed only to identify subjects whose data should be excluded from analysis; Bargh & Chartrand, 2000.) One additional subject reported that he did not employ any strategy until the end of the experiment, when he began trying to distract himself by intentionally thinking of other words. (For this subject, the RIT effect occurred on every single trial.) One subject did not provide an answer as to whether she employed any strategy during the experiment. Another subject attempted to maintain a "blank" state of mind. The rest of the subjects attempted to employ some kind of self-distraction strategy, including covert singing (2 subjects; by-subject proportion of RIT effects across trials = 0.92 and 0.64), the covert repeating of words (4 subjects; by-subject proportion of RIT effects across trials = 0.87, 0.87, 0.82, 0.89), or a strategy in which they attended to parts of the object but not to the whole object (4 subjects; by-subject proportion of RIT effects across trials = 0.99, 0.81, 0.43, 0.74). Of the subjects who reported that they attempted to use such strategies, two concluded that these strategies were not effective.

Consistent with this conclusion are the findings from Cho et al. (2014). In that study, subjects attempted to thwart the RIT effect by occupying the "verbal buffer" (Cho et al., 2014). In one condition, participants were instructed to reiteratively subvocalize a speech sound ("da, da, da") throughout the trial. This internally generated content was self-generated and intentional. Involuntary subvocalizations of object names still arose on over 80% of the trials. One could hypothesize that subvocalizations occurred because of the pauses between the intended speech sounds, but this is inconsistent with the observation that comparable results arose even when participants subvocalized a continuous, unbroken hum ("daaa….") throughout the trial. Perhaps what is needed to thwart the subvocalization effect in the RIT are verbal representations that are more complex in nature than what was employed by subjects in Cho et al. (2016). For example, perhaps the intentional repetition of a string of real words could diminish substantially the likelihood of an RIT effect. This is a topic worthy of future investigation.

4.1. The RIT effect

In the present task, the RIT effect appears to rely on several mechanisms (Allen et al., 2013). First, it seems that the relevant action set must be activated. This activation stems from the instructions provided by the experimenter (Allen et al., 2013). It is unlikely that, without such an activation of set, subjects would experience the phonological representations of the names of the objects that happen to be perceived visually. This action set is somehow held in memory, at least from the beginning of the trial until the onset of the visual object. The final process occurs when the onset of the visual object begins the stages of processing that, somehow, lead to the consciousness of the involuntary imagery (*e.g.*, subvocalization of the object name). The RIT effect corroborates what can be observed in everyday life: That conscious contents, including high-level, sophisticated contents, often “just happen.”

One important difference between the involuntary subvocalization that constitutes the RIT effect and the kinds of effects found in most experiments concerning ironic processing is that, in the latter, subjects are presented with a verbal description (*e.g.*, verbal stimuli such as “Do not think of white bears”) and then the subjects experience involuntary perceptuo-semantic imagery, but in the RIT, the stimuli are visual and the involuntary imagery is phonological in nature. One could state that the RIT involves the opposite direction of activation of that found in the classic studies on ironic processing (see review of the classic research in Wegner, 1989).

In the RIT, subjects may experience the urge to name the object aloud (see Allen et al., 2013), but this action plan is not expressed overtly because of the influence of a separate and competing action set—that of not uttering the object name (see discussion in Bhangal, Cho, et al., 2016). This kind of action control often fails to arise when actions are decoupled from consciousness, as occurs in some neurological disorders (Morsella & Bargh, 2011). In such cases, actions appear impulsive and not influenced by all the kinds of information by which the actions should be influenced. These maladaptive actions are often sophisticated but “un-integrated” (Morsella & Bargh, 2011).

Regarding integrated actions, it has been proposed that, for the generation of these kinds of actions, conscious contents must reside in the same decision space and must share the same format (which appears to be perceptual-like; see discussion in Cabanac, 1996; Gray, 1995; Koch, Massimini, Boly, & Tononi, 2016; Morsella et al., 2016a). For example, in the model proposed by Cabanac (1996), all conscious contents, whether for high-level or low-level processes, must reside in the same decision space and have the same, *sensation-based* format. Cabanac (1996) adds that “consciousness evolved from sensation” (p. 34), which, from the standpoint of Cabanac, implies that all kinds of conscious content will retain the properties and characteristics (*e.g.*, quality, duration, and intensity) that are found in sensations. Moreover, Cabanac (1996) proposes that affective responses such as pleasure serve an important role in action selection: “Sensory pleasure serves as the common currency for the trade-offs that take place in the mind to achieve the ranking of priorities and ensure that the most urgent motivation has access first to the behavioral final common path” (p. 38).

Consistent with these views, both the visual stimulus and unintentional subvocalization in our experiment are perceptual-like contents. For example, the latter is based, not on the articulatory code, but on a phonological code that is based on audition (Levelt, 1989), which is perceptual in nature (Pasley et al., 2012). Thus, the eliciting stimulus and the subvocalization are in the same format, which happens to be perceptual-like.

4.2. The multi-determined response

Despite the complexity of our task, subjects were able to perform the multi-determined response. This response required integrations among disparate tokens and modalities, including involuntary conscious contents (the subvocalization), an item held in prospective memory (the phonological form of the rhyme target), and a visual go/no-go cue (the solid or dotted border). These disparate, multimodal conscious contents somehow yielded a single “integrated” action—the multi-determined response involving the second (green) button. The task captures the six aforementioned component processes and also the kind of processing that is polymodal, context-sensitive, and associated with the conscious field. Our data are consistent with frameworks proposing that the conscious field brings together diverse sources of information to yield adaptive action (Baars, 1988; Dehaene, 2014; Norman & Shallice, 1980).

In the current task, the generation of high-level conscious contents (*e.g.*, subvocalizations) was involuntary. In one theory (Passive Frame Theory; Morsella et al., 2016a), the mechanisms generating the conscious contents (*i.e.*, the “content generators”) are themselves unconscious, and so are the mechanisms responding to those contents (which are mechanisms that are distinct from the systems that generate conscious contents). In short, in a form of “unidirectional communication,” conscious contents (*e.g.,* a red apple and an urge) are “sampled” only by action systems, which are themselves unconscious. Hence, one conscious content does not, in a sense, “know” of the nature of other conscious contents nor of the nature of ongoing behavior. It has been argued that this form of built-in ignorance on the part of the system is actually adaptive (see Baumeister, Vohs, DeWall, & Zhang, 2007; Firestone & Scholl, 2016).

To function adaptively, such an architecture—in which conditions of action-relevance are likely to be represented in subsystems distinct from those generating conscious contents—must be capable of solving certain challenges. For example, in one circumstance, it might be the case that the adaptive response to conscious content *X* depends on the spatial distance between, say, concurrently activated contents *Y* and *Z*, with *Y* and *Z* being just two out of many conscious contents composing the field. To prepare for such a scenario, the conscious field must somehow retain and represent the spatial relations amongst *all* of its many contents. It must do this at each moment in time, and for all of its contents, simply because the field and its content generators "do not know," in a sense, which subset of stimuli, or spatial relations amongst stimuli, will be action-relevant. The spatial relations amongst stimuli could thus be regarded as themselves being a kind of higher-order, *discriminative stimulus*. Because of this, adaptive action selection cannot be determined just by the relative strength of activations of the response codes activated by each of the contents in the field; the relations amongst contents, whether spatial or non-spatial, too, must influence the strength of activation of response codes and, subsequently, the process of action selection. This illuminates why the conscious field incessantly provides such a rich representation of the world, including its spatial properties, even though, at one moment in time, only a subset of the stimuli might appear to be action-relevant (see discussion in Morsella, Godwin, Jantz, Krieger, & Gazzaley, 2016b).

With this theorizing in mind, it is important to appreciate that, before the presentation of the critical trials, at least 0.63 of all the subjects in our experiment were aware of the frame surrounding the image even when the frame was not in any way action-relevant. From the point of view of Passive Frame Theory, the "content generator" producing the percept of the frame surrounding the stimulus does not, in a sense, know whether that percept is action-relevant: At one point in the study, the frame was not action-relevant; however, it certainly became action-relevant during the subsequent, critical trials. In short, regardless of how action-relevant the solid frame was, the percept was presented in the conscious field. (Also consistent with Passive Frame Theory was the observation of the "encapsulation" [Firestone & Scholl, 2016] of the generation of the high-level content [*i.e.*, the subvocalization], which occurred even though it conflicted with the subject's intentions.)

It is worth reiterating that our project is progressive in that it does not represent the standard "hypothesis-testing" approach; instead, in the spirit of Tolman (1948), it is a "proof of mechanism," one revealing that subjects can, under complex circumstances, perform a particular multi-determined behavior in response to high-level, involuntary cognitions. The approach is the kind of cumulative, theory-driven investigation that is today encouraged by leaders in the field (*e.g.*, Fiedler, 2017; Nosek et al., 2012). It is our hope that our new task will be investigated further in several ways, including with clinical populations and by use of neuroimaging technologies. The empirical finding and the theories that motivated the present project have implications for the study of multi-determined voluntary actions and neurological disorders involving such actions. Apart from these considerations, our behavioral paradigm and associated theorizing have implications for investigations regarding cognitive control, action control, prospective memory, mental imagery, psychopathology, and the entry of contents into the conscious field, which remains one of the greatest mysteries in science.

Acknowledgment

This research was supported by the Center for Human Culture and Behavior at San Francisco State University, United States. Lara Krisst conducted the pilot study in which subliminal stimuli (orthographs) were presented as the stimuli in the Reflexive Imagery Task.

Appendix

	Stimulus	Visual angle
First Trial	Bird	4.18 × 3.58 (35 × 30 mm)
Practice Trials (Rhyme Words)	Boxer	2.98 × 3.58 (25 × 30 mm)
	Cheerleader	4.18 × 4.77 (35 × 40 mm)
	Feather	6.56 × 4.77 (55 × 40 mm)
	Lawnmower	4.18 × 4.18 (35 × 35 mm)
	Pacifier	3.58 × 4.77 (30 × 40 mm)
	Screwdriver	4.77 × 3.58 (40 × 30 mm)
Practice Trials (Non-Rhymes)	Butterfly	4.18 × 4.18 (35 × 35 mm)
	Elephant	4.18 × 3.58 (35 × 30 mm)
	Piano	4.18 × 3.58 (35 × 30 mm)
	Snowman	3.58 × 4.18 (30 × 35 mm)
Critical Trials (Rhyme Words)	Alligator	3.58 × 2.39 (30 × 20 mm)
	Anchor	2.98 × 2.98 (25 × 25 mm)
	Doctor	2.98 × 5.96 (25 × 50 mm)
	Dresser	3.58 × 2.98 (30 × 25 mm)
	Finger	2.39 × 2.39 (20 × 20 mm)
	Fire	6.56 × 4.18 (55 × 35 mm)
	Flower	2.98 × 2.98 (25 × 25 mm)
	Grasshopper	3.58 × 2.98 (30 × 25 mm)
	Hammer	3.58 × 3.58 (30 × 30 mm)
	Hanger	3.58 × 2.98 (30 × 25 mm)

	Helicopter	4.77 × 3.58 (40 × 30 mm)
	Ladder	4.18 × 2.98 (35 × 25 mm)
	Lobster	3.58 × 3.58 (30 × 30 mm)
	Paper	5.37 × 5.96 (45 × 50 mm)
	Pepper	2.98 × 2.98 (25 × 25 mm)
	Pitcher	2.39 × 2.98 (20 × 25 mm)
	Refrigerator	2.98 × 3.58 (25 × 30 mm)
	Rooster	3.58 × 3.58 (30 × 30 mm)
	Ruler	3.58 × 2.98 (30 × 25 mm)
	Saltshaker	2.98 × 2.98 (25 × 25 mm)
	Screwdriver	3.58 × 3.58 (30 × 30 mm)
	Spider	3.58 × 2.98 (30 × 25 mm)
	Sweater	2.98 × 2.98 (25 × 25 mm)
	Tiger	3.58 × 3.58 (30 × 30 mm)
	Tower	1.79 × 5.96 (15 × 50 mm)
Non-Critical Trials (Non-Rhymes)	Accordion	2.98 × 2.39 (25 × 20 mm)
	Balloon	2.98 × 2.98 (25 × 25 mm)
	Barrel	3.58 × 3.58 (30 × 30 mm)
	Baseball Bat	3.58 × 3.58 (30 × 30 mm)
	Beetle	2.98 × 2.98 (25 × 25 mm)
	Camel	3.58 × 2.98 (30 × 25 mm)
	Candle	2.39 × 3.58 (20 × 30 mm)
	Cannon	4.77 × 2.98 (40 × 25 mm)
	Carrot	3.58 × 3.58 (30 × 30 mm)
	Chisel	9.53 × 8.34 (80 × 70 mm)
	Cigar	3.58 × 3.58 (30 × 30 mm)
	Donkey	3.58 × 3.58 (30 × 30 mm)
	Glasses	2.98 × 2.98 (25 × 25 mm)
	Goat	3.58 × 3.58 (30 × 30 mm)
	Guitar	3.58 × 2.98 (30 × 25 mm)
	Mitten	2.98 × 2.98 (25 × 25 mm)
	Pencil	3.58 × 2.98 (30 × 25 mm)
	Pliers	5.96 × 5.96 (50 × 50 mm)
	Raccoon	2.98 × 2.39 (25 × 20 mm)
	Roller Skate	3.58 × 2.98 (30 × 25 mm)
	Tennis Racket	4.18 × 3.58 (35 × 30 mm)
	Watering Can	3.58 × 3.58 (30 × 30 mm)
	Watermelon	3.58 × 2.98 (30 × 25 mm)
	Whistle	3.58 × 3.58 (30 × 30 mm)
	Zebra	3.58 × 2.39 (30 × 20 mm)

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., 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.

Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge, England: Cambridge University Press.

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.

Baumeister, R. F., Vohs, K. D., DeWall, N., & Zhang, L. (2007). How emotion shapes behavior: Feedback, anticipation, and reflection, rather than direct causation. *Personality and Social Psychology Review*, *11*, 167–203.

Bhangal, S., Allen, A. K., Geisler, M. W., & Morsella, E. (2016). Conscious contents as reflexive processes: Evidence from the habituation of high-level cognitions. *Consciousness and Cognition*, *41*, 177–188.

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., & 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.

Buccafusco, J. J. (2001). *Methods of behavior analysis in neuroscience*. Boca Raton, FL: CRC Press.

Bulevich, J. B., Roediger, H. L., Balota, D. A., & Butler, A. C. (2006). Failures to find suppression of episodic memories in the think/no-think paradigm. *Memory & Cognition*, *34*, 1569–1577.

Cabanac, M. (1996). On the origin of consciousness, a postulate and its corollary. *Neuroscience and Biobehavioral Reviews*, *20*, 33–40.
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*. http://dx.doi.org/10.3389/fpsyg.2014.01445.
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.
Crick, F., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, *6*, 1–8.
Dehaene, S. (2014). *Consciousness and the brain: Deciphering how the brain codes our thoughts*. New York: Viking.
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 [Target Article]*, *39*, 1–77.
Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton-Mifflin.
Freeman, W. J. (2004). William James on consciousness, revisited. *Chaos and Complexity Letters*, *1*, 17–42.
Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*, 493–503.
Gray, J. A. (1995). The contents of consciousness: A neuropsychological conjecture. *Behavioral and Brain Sciences*, *18*, 659–676.
Hertel, P. T., & Calcaterra, G. (2005). Intentional forgetting benefits from thought substitution. *Psychonomic Bulletin & Review*, *12*, 484–489.
Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, *21*, 1229–1243.
Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: Progress and problems. *Nature Reviews Neuroscience*, *17*, 307–321.
Köhler, W. (1947). *Gestalt psychology: An introduction to new concepts in modern psychology*. New York: Liveright Publishing Corporation.
Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.
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*, 345–393.
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.
Merker, B. (2007). Consciousness without a cerebral cortex: A challenge for neuroscience and medicine. *Behavioral and Brain Sciences*, *30*, 63–134.
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, N. E. (1959). Liberalization of basic S-R concepts: Extensions to conflict behavior, motivation, and social learning. In S. Koch (Vol. Ed.), *Psychology: A study of a science*. *Vol. 2*, (pp. 196–292). New York: McGraw-Hill.
Miller, G. A. (1996). *The science of words (Scientific American Library)*. New York: W. H. Freeman & Co.
Miller, B. L., & Cummings, J. L. (2007). *The human frontal lobes: Functions and disorders* (2d ed.). New York: Guilford Press.
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., & Bargh, J. A. (2011). Unconscious action tendencies: Sources of 'un-integrated' action. In J. T. Cacioppo, & J. Decety (Eds.), *The Oxford handbook of social neuroscience* (pp. 335–347). New York: Oxford University Press.
Morsella, E., Berger, C. C., & Krieger, S. C. (2011). Cognitive and neural components of the phenomenology of agency. *Neurocase*, *17*, 209–230.
Morsella, E., Godwin, C. A., Jantz, T. J., Krieger, S. C., & Gazzaley, A. (2016b). Passive frame theory: A new synthesis. *Behavioral and Brain Sciences* [Authors' Reply to Commentaries], 39, 44–70.
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 [Target Article]*, *39*, 1–17.
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.
Morsella, E., Zarolia, P., & Gazzaley, A. (2012). Cognitive conflict and consciousness. In B. Gawronski, & F. Strack (Eds.), *Cognitive consistency: A unifying concept in social psychology*(pp. 19–46). New York: Guilford Press.
Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231–259.
Norman, D. A., & Shallice, T. (1980). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (pp. 1–18). New York: Plenum Press.
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.
Schomers, M. R., Kirilina, E., Weigand, A., Bajbouj, M., & Pulvermüller, F. (2014). Causal influence of articulatory motor cortex on comprehending single spoken words: TMS evidence. *Cerebral Cortex*, *1–9*. http://dx.doi.org/10.1093/cercor/bhu274.
Searle, J. R. (2000). Consciousness. *Annual Review of Neurosciences*, *23*, 557–578.
Slevc, L. R., & Ferreira, V. S. (2006). Halting in single word production: A test of the perceptual loop theory of speech monitoring. *Journal of Memory and Language*, *54*, 515–540.
Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174–215.
Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*, 189–208.
Wegner, D. M. (1989). *White bears and other unwanted thoughts*. New York: Viking/Penguin.
Wegner, D. M. (1994). Ironic processes of thought control. *Psychogical Review*, *101*, 34–52.
Wegner, D. M., Schneider, D. J., Carter, S. R., & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology*, *53*, 5–13.
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.