Acta Psychologica 166 (2016) 37–41

Contents lists available at ScienceDirect

Acta Psychologica

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

Involuntary symbol manipulation (Pig Latin) from external control: Implications for thought suppression

Hyein Cho [a], Pareezad Zarolia [b], Adam Gazzaley [c,d], Ezequiel Morsella [a,c,*]

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

ARTICLE INFO

Article history:
Received 7 August 2015
Received in revised form 16 March 2016
Accepted 17 March 2016
Available online xxxx

Keywords:
Ironic processing
Reflexive Imagery Task
Thought suppression
Encapsulation
Cognitive control

ABSTRACT

In *ironic processing*, one is more likely to think about something (e.g., white bears) when instructed to not think about that thing. Entry into consciousness of such content may be automatic, reflecting the 'encapsulated' nature of the generation of *conscious contents*. Based on this research, the Reflexive Imagery Task (RIT) reveals that, following the activation of action sets, conscious contents can arise involuntarily and systematically in response to external stimuli. In the most basic version of this paradigm, participants are presented with visual objects and instructed to not think of the names of the objects, which is challenging. Here, we addressed one criticism of the RIT—that the effect arises only for automatic processes (e.g., forms of cued-memory retrieval) and not for more complex processes (e.g., symbol manipulation). Participants were first trained to perform a word-manipulation task similar to the game of Pig Latin (e.g., "CAR" becomes "AR-CAY"). Such a task involves complex symbol manipulations that are associated with processes in frontal cortex. After training, participants were instructed to not transform stimulus words in this way. The RIT effect still arose under these conditions. This striking finding is relevant to theories of cognitive control, psychopathology, and conscious/unconscious processing.

1. Introduction

The conscious mind, replete with thoughts, feelings, and desires, hardly looks like a reflex. Reflexes (e.g., the pupillary reflex) are predictable, insuppressible, and capable of being controlled systematically by external stimuli (e.g., light). In contrast, *conscious contents* (Merker, 2007; Seth, 2007) are often unpredictable, 'offline,' and insulated from the influence of the external world (Chomsky, 1988; Fodor, 1975, 1983; Shallice, 1972; Smallwood & Schooler, 2006; Wegner & Bargh, 1998). While attending an important lecture, for example, one might experience out of the blue conscious contents about, say, a future vacation. To one, these thoughts are unpredictable; to those around one, they are both un-inferable and uncontrollable. Such observations, and important theoretical developments (e.g. Fodor, 1975; Schacter & Addis, 2007; Smallwood & Schooler, 2006), have led to the prevalent view that the workings of the conscious mind are very different in nature from those of reflexes.

Despite these prevalent views, some theorists (e.g., Freud, 1938; James, 1890; Miller, 1959; Vygotsky, 1962; Wegner, 1989) have proposed that, under the appropriate conditions, conscious contents can arise in a manner resembling that of reflexes. Under these conditions, the generation of conscious contents is predictable and susceptible to external influence (see review in Allen, Wilkins, Gazzaley, & Morsella, 2013). It has been hypothesized that this reflects the fact that, in most circumstances, the generation of conscious contents is 'encapsulated': One has no conscious access regarding the workings of these generative mechanisms, nor can one directly influence these mechanisms (Firestone & Scholl, 2014, in press; Fodor, 1983; Merrick, Godwin, Geisler, & Morsella, 2014). This notion has led to the view that one is conscious only of the outputs of processes but not of the processes themselves (Lashley, 1956; Miller, 1962). Consistent with this theorizing is Helmholtz's (1856/1925) notion of the *unconscious inference*.[1]

Consider, for example, action-related urges. When one attempts to hold one's breath while underwater, for instance, one cannot turn off the conscious inclination to inhale (Morsella, 2005). In general, action-related urges cannot be weakened or suppressed voluntarily, even when doing so would be adaptive (Morsella, 2005; Öhman & Mineka, 2001; Olsson & Phelps, 2004). These urges enter consciousness in response to certain stimulus conditions. In many cases, the urges are less controllable than is behavior, leading to the insight that, although

* 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] Helmholtz (1856/1925) suggests that unconscious inferences generate conscious contents, not only perception (e.g., depth perception), but also in high-level processes, such as automatic word reading, which requires a complex, multi-stage process (Levelt, 1989).

http://dx.doi.org/10.1016/j.actpsy.2016.03.004
0001-6918/

inclinations can be suppressed behaviorally, they often cannot be suppressed mentally (Bargh & Morsella, 2008).

2. The Reflexive Imagery Task

Building on theorizing about encapsulation, unconscious inferences, ironic processing, and other phenomena (i.e., Ach, 1905/1951; Freud, 1938; Gollwitzer, 1999; Helmholtz, 1856/1925; James, 1890; Miller, 1959; Vygotsky, 1962; Wegner, 1989), the *Reflexive Imagery Task* (RIT; Allen et al., 2013) was developed to further investigate the unintentional nature of the generation of high-level conscious contents. In the most basic version of this task, participants are instructed to not subvocalize (i.e., say in one's head) the names of objects that are presented on a computer screen. Interestingly, participants fail to suppress such vocalizations on the majority of the trials (86% in Allen et al., 2013; 87% in Cho, Godwin, Geisler, & Morsella, 2014; and 73% in Merrick, Farnia, Jantz, Gazzaley, & Morsella, 2015). When the effect of involuntary subvocalization arises, it arises only moments (~2 s) after the visual stimulus appears ($M = 1451.27$ 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).

We will now present the reader with a demonstration of this curious effect.[2] Your task is to not subvocalize the name of an object that will be presented somewhere in the next sentence. Here is the stimulus (★). Research involving the RIT reveals that this instruction, in combination with the action that the instruction activates and the subsequent presentation of a visual stimulus, renders it difficult for the participant to suppress activations of the phonological form of the object name 'star' (Allen et al., 2013). This 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/). In still more complex variants of the RIT (e.g., Merrick et al., 2015), participants are instructed to (a) not subvocalize the name of visual objects and, in addition, (b) not count the number of letters composing object names. On a substantial proportion of trials (0.30 [$SE = 0.04$]), participants experience both kinds of involuntary thought. For the RIT effect to arise, it seems that there must be activation of the appropriate action set, which is activated somehow by the instructions provided by the experimenter (Allen et al., 2013).

Empirical evidence suggests that, for participants, entry into consciousness of the undesired content 'just happens' and is not simply an artifact of high-level strategic processes on the part of the participant (Bhangal, Merrick, & Morsella, 2015). For example, it seems unlikely that participants are thinking of the object name incidentally, through the following kind of reasoning. "You told me to not think of the name of the object. The object name is *X*. Therefore, I should not think of *X*." The view that the effect is automatic is further supported by the observations that, on some trials, the effect arises too quickly to be due to strategic processing (Allen et al., 2013; Cho et al., 2014) and that the effect still arises under conditions of cognitive load, in which it is difficult for participants to implement strategic processing (Cho et al., 2014). Last, the effect is unlikely to be due to demand characteristics because the subvocalizations are influenced systematically by factors such as word frequency (Bhangal et al., 2015). Such an artifact of experimental demand would require for participants to have a theory regarding how word frequency should influence responses in an experiment.

The nature of the RIT effect and of ironic effects, more generally, is not fully understood (Wegner & Schneider, 2003). Nevertheless, the paradigm provides a unique portal through which to investigate the often encapsulated nature of the generation of conscious contents as well as the limits of these (unconscious) generative mechanisms, which illuminates interesting contrasts in brain function. For example, as mentioned above and as explained by theory (Morsella, 2005), although participants may fail at suppressing the activation of action-related conscious contents (subvocalizing an object name or letter counting; Merrick et al., 2015), participants nonetheless succeed at suppressing overt behavior. For instance, it is clear that participants can succeed at suppressing an overt object-naming response when instructed to not name the stimulus object aloud (Allen et al., 2013). (Interestingly, participants nonetheless report some urges to name the object aloud; Allen et al., 2013.) Thus, the RIT permits one to investigate the interesting contrast between the control of overt behavior (which is suppressible) versus that of the generation of action-related conscious contents (which is often insuppressible). Second, in Merrick et al. (2015), the RIT effects for the act of letter counting occurred more often for short words than for long words. Such a contrast stemming from the RIT might illuminate the limits of the unconscious processes involved in the generation of conscious contents. In other words, the RIT could reveal the limits of unconscious inferences. In this way, the RIT provides an additional method to contrast the capacities of conscious and unconscious processes, one that 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 processes usually operate upon (Bargh & Morsella, 2008).

For the present project, we addressed a criticism of the RIT—that the effect arises only for simple, automatic processes (e.g., certain forms of cued-memory retrieval) and cannot arise for more complex processes (e.g., symbol manipulation). (One of the present authors [E.M.] has on several occasions been informed by other scientists that the latter is unlikely to arise unintentionally.) Can the RIT effect arise when the process under investigation requires symbol manipulation? Symbol manipulation of a sophisticated kind occurs in the childhood game of Pig Latin (e.g., "CAR" becomes "AR-CAY"). To investigate this empirical question, participants in the present study were first trained to perform a word-manipulation task that is similar to Pig Latin, which involves complex symbol manipulations that are associated with processes in frontal cortex (Miller & Cummings, 2007). After training, participants were instructed to not transform stimulus words in this way. We predicted that the RIT effect would still arise under these conditions, even though the effect requires complex symbol manipulations associated with processes in frontal cortex.

Such a striking and counterintuitive effect would be the first demonstration of unintentional, high-level symbol manipulation caused by external control. It is worth emphasizing that such external control is diametrically opposed to the intentions of the participant. In several ways, the finding would inform theories regarding the differences between conscious and unconscious processes. For example, such an RIT effect would shed more light on the sophisticated capacities of the (often encapsulated) processes that engender the contents composing the conscious field. The processing involved could be construed as more elaborate and less amenable to automatic processing than object naming or the kind of word reading discussed by Helmholtz (1856/1925). Second, the finding would also illuminate the limits of the intentional, executive processes which attempt to regulate which contents enter consciousness. Third, it would provide further evidence that the RIT effect is not restricted to (relatively) simple processes such as automatic word-reading and automatic cued-retrieval (Schacter & Tulving, 1994). Fourth, the finding would stem from the kind of incremental research that, building on established paradigms (Allen et al., 2013; Wegner, 1989), is the kind of empirical investigation that, today, is important for advancements in the field of psychological science.

[2] This effect and the associated theorizing are consistent with the research by Wegner and colleagues on *ironic processing*, in which one is more likely to think about something (e.g., white bears) when instructed to not think about that thing (Wegner, 1994; Wegner, Schneider, Carter, & White, 1987). For reviews of ironic processing and thought suppression, see Wegner (1989) and Rassin (2005). For treatment of the relationship between Wegner's classic model and the RIT, see Allen et al. (2013).

3. Method

3.1. Participants

San Francisco State University students ($n = 35$; 27 females; $M_{Age} = 21.9$ years, $SD_{Age} = 4.32$ years) participated for course credit. The involvement of human participants in our project was approved by the Institutional Review Board at San Francisco State University.

3.2. Stimuli and apparatus

Stimuli were presented on an Apple iMac computer monitor (50.8 cm) with a viewing distance of approximately 48 cm. Stimulus presentation and data recording were controlled by PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). Participants inputted their responses to questions and instructions by computer keyboard. All questions and instructions were written in black 36-point Chicago font; all fonts were displayed on a white background. The stimuli consisted of 20 words of well-known object names (Appendix A). All stimuli were displayed in the center of the screen with a subtended visual angle of 5.96° × 6.56° (5 cm × 5.5 cm). These words were culled from stimulus lists that were used successfully in previous RIT research (Allen et al., 2013).

3.3. Procedures

Participants were shown each word only once in random order (trials = 20).[3] From the 35 participants who participated in the study, the data from 32 participants were included in the analysis. The data for 3 participants were excluded from analysis because these participants did not follow instructions.

Participants were first presented with the rules of a word-transformation game that is similar to the childhood language game of Pig Latin: "*Pig Latin is a language game played in English. In order to form the Pig Latin version of a word, the beginning letter of the word is moved to the end of the word and an '-ay' is attached at the end. For example, 'car' is transformed into 'ar-cay,' and 'picture' is transformed into 'icture-pay' in Pig Latin.*" After participants were informed of the rules of Pig Latin and before the critical trials, participants performed five practice trials in which they transformed words following the Pig Latin rule. These practice trials served to corroborate that each participant understood the rule and could perform the transformations. On each practice trial, participants were instructed to keep their eyes focused on the center of the screen at all times. They were then presented with a word (10 s) and instructed to transform the word according to the Pig Latin rule. (In the subsequent, critical trials, participants were instructed to *not* perform such transformations.) Following each practice trial, participants were instructed to utter the name of the transformed word.

After the training, participants completed the critical trials, in which they were instructed to *not* think of the transformed word. Specifically, participants were told that they would be shown a series of words and that they should press the spacebar when they thought (unintentionally) of the Pig Latin version of the presented word. In addition, participants were instructed to press the spacebar for each subsequent time that, during the trial, they happened to think of the transformation of the stimulus word for the duration of the trial. If participants did not happen to think of the transformed word, they did not respond in any way. Participants were instructed to keep their eyes focused on the center of the screen at all times. Following each trial, participants were asked a question about their performance: "*Did you think of the Pig Latin transformation of the word?* (*Y*/*N*)."

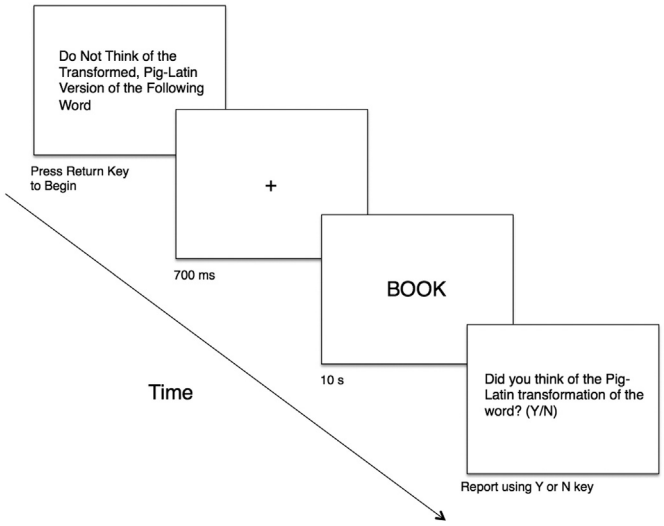

Fig. 1. Schematic depiction of a typical trial. Not drawn to scale.

Participants were reminded of the instruction to not think of the word on each trial. Before presentation of each word, the phrase "*Do Not Think of the Transformed, Pig Latin Version of the Following Word*" was displayed in the center of the screen, serving as a ready prompt; participants indicated their readiness by pressing the return key. Once participants indicated their readiness, a fixation-cross (+) appeared in the center of the screen (700 ms), preparing participants for the presentation of the stimulus. Following the fixation, a stimulus word appeared (10 s). (See Fig. 1.)

Once participants completed the experiment, they responded to a series of funneled debriefing questions (following the procedures of Bargh & Chartrand, 2000), which included general questions to assess whether participants (a) were aware of the purpose of the study, (b) had any strategies for completing the task, and (c) had anything interfere with their performance on the task. Additionally, participants were asked questions to assess whether they (d) understood the instructions for the Pig Latin transformation, and (e) felt like they were good at the Pig Latin transformations before participating in the study. The responses to these funneled debriefing questions did not warrant the removal of any data from analysis.

4. Results

Out of the 20 trials per participant, the mean proportion of trials in which involuntary transformations arose was 0.43 ($SD = 0.25$, $SE = 0.04$, range = 0–1), a proportion that was significantly different from zero, $t(31) = 9.66$, $p < 0.001$. The same result was found with arcsine transformations of the proportion data, $t(31) = 12.82$, $p < 0.001$. (Arcsine transformations are often used to statistically normalize data that are in the form of proportions.) The mean latency of the first transformation was 4095.44 ms ($SD = 1475.06$, $SE = 264.93$; range = 1545.60–6701.00 ms). The mean number of times that participants experienced involuntary transformations per trial was 0.65 ($SD = 0.60$, $SE = 0.11$, range = 0–2.60), which was significantly different from zero, $t(31) = 6.16$, $p < 0.001$.

5. Discussion

Our project yielded another conceptual replication and extension of the RIT, which has proven to be a robust and reliable paradigm. It is striking that, in the current project, the RIT effect still arose, even though the word transformation involves complex symbol manipulations and

[3] The duration of this block of 20 trials was very short (<20 min). During the same session, the subjects also participated in another block of RIT trials, a block of trials that involved a very different type of mental operation and no symbol manipulation of any kind. No noteworthy conclusions could be drawn from these data; hence, they will not be discussed any further. The presentation order of the two RITs was fully counterbalanced across subjects.

processes associated with frontal cortex (Miller & Cummings, 2007). The effect arose on 43% of the trials. It is also striking that the effect still arose despite the fact that the training session was composed of only a handful of trials. This is the first demonstration of unintentional, high-level symbol manipulation caused by external control. It is worth reiterating that such external control was effective even though it was diametrically opposed to the intentions of subjects.

The latency of the effect ($M = 4095.44$ ms, $SD = 1475.06$) tended to be longer than that found in previous studies which examined the latencies of involuntary object naming ($M = 1451.27$ 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). The increased latencies found in the current study may reflect that the process of word manipulation is more complex than that of object naming. The latter could be explained as resulting from cue-based memory retrieval, which is capable of occurring automatically (Schacter & Tulving, 1994; Squire, 1987), but the former is unlikely to result from so (relatively) simple a process. Hence, the present data suggest that the unconscious processes involved in the generation of conscious contents in the RIT can include mechanisms that might be more complex, and that, normally, are less automatic, than that of cue-based memory retrieval. Future investigations could ascertain on a trial-by-trial basis whether participants experience this RIT effect as arising spontaneously (e.g., using the methods in Bhangal et al., 2015).

As mentioned above, the RIT permits one to investigate several informative contrasts in the mind/brain. For example, there is the contrast between the suppressibility of overt behavior versus that of conscious, mental representations of behavioral inclinations (e.g., urges). In the RIT, the latter is more challenging than the former (Allen et al., 2013). This contrast has been accounted for in theoretical frameworks (e.g. Bargh & Morsella, 2008; Krisst, Montemayor, & Morsella, 2015; Morsella, 2005). For example, according to Krisst et al. (2015; and see also Morsella, Godwin, Jantz, Krieger, & Gazzaley, in press), consciousness reflects a stage of processing concerning potential 'action options.' These options are generated in an encapsulated manner. Although activated in the conscious field, these options need not influence action directly (Allen et al., 2013): In action selection, some of these options are selected for production, while others are not. The RIT permits one to investigate additional contrasts. For example, it seems that some experimental conditions are more likely to engender RIT effects than other conditions. Perhaps the effect will not arise for symbol manipulations involving more than one transformation. Our task required only one word transformation. The lack of an RIT effect for processes requiring more than one transformation would reveal the limitations of the mechanisms engendering unconscious contents unintentionally. Regarding such limitations, it is worth considering that, in the Merrick et al. (2015) RIT study, unintentional letter counting was more likely to occur for short words than for long words.

It might also be the case that the RIT effect can arise only for the kinds of processes that one is normally capable of carrying out intentionally. For instance, one can choose to add numbers, subtract numbers, or even play the game of Pig Latin; however, one cannot, by sheer will and without some difficulty, make oneself ecstatic or frightened. It has been proposed that it would not be adaptive for such incentive/motivational states to be influenced directly by intentions or by other conscious contents (e.g., desires and beliefs; Baumeister, Vohs, DeWall, & Zhang, 2007; Öhman & Mineka, 2001). For instance, if one's beliefs could lead one to voluntarily turn off pain, guilt, or hunger, then these negative states would lose their adaptive value (Baumeister et al., 2007). (To influence incentive/motivational states, one may employ *indirect cognitive control*, such as imagining something scary to induce fear; Morsella, Lanska, Berger, & Gazzaley, 2009; Morsella, Larson, & Bargh, 2010.) Hence, the encapsulation of the generation of conscious contents is part of the architecture of the cognitive apparatus. In most cases, this architecture is adaptive. This interpretation, along with theories about encapsulation (Firestone & Scholl, 2014, in press; Pylyshyn, 1984), may have implications for our understanding of the basic mechanisms in psychopathological phenomena (e.g., in obsessions, ruminations, intrusive cognitions, compulsions; Magee, Harden, & Teachmen, 2012; Nolen-Hoeksema, Wisco, & Lyubomirsky, 2008).

One limitation of the present study is that our primary dependent measure is based on the experimental technique of introspection. This self-report technique has well known limitations. For example, trial-by-trial self-reports might be inaccurate because participants may forget or misconstrue what actually transpired on a given trial (Block, 2007) or because participants base their responses on heuristics or beliefs regarding how one should comport oneself in an experiment (see discussion of this limitation in Morsella, Wilson, Berger, Honhongva, Gazzaley, & Bargh, 2009). This limitation could be remedied to some extent in future investigations which couple our dependent measure with neuroimaging that could, under certain conditions, provide objective indices regarding whether or not an RIT effect arose during a given trial. Regarding the validity of the RIT effect, evidence from neuroimaging studies corroborates that, in paradigms in which participants must report about the occurrence of particular conscious contents, it is unlikely that participants confabulate about the occurrence of these mental events (Mason et al., 2007; McVay & Kane, 2010; Mitchell et al., 2007; Wyland, Kelley, Macrae, Gordon, & Heatherton, 2003).

Despite the limitations of the current project and what remains unknown about the nature of the RIT and ironic processing more generally, this project—building on prior research (e.g. Allen et al., 2013; Wegner, 1989) and involving a phenomenon that is robust, multifaceted, and reliable—is the kind of incremental research that is important for progress in the field of psychological science (Nosek, Spies, & Motyl, 2012). In addition, the RIT and new findings reported here are relevant to many subfields of psychological science, including attention (e.g., Lee, Lim, Lee, Kim, & Choi, 2009), self-regulation (e.g., Metcalfe & Mischel, 1999), emotion regulation (e.g., Baumeister et al., 2007), mental imagery (e.g., Kavanagh, Andrade, & May, 2005), mind wandering (e.g. Mason et al., 2007; Smallwood & Schooler, 2006), and psychopathology (e.g., rumination; Nolen-Hoeksema et al., 2008). It is our hope that the present project will serve as a foundation for future investigations regarding the nature of the mechanisms giving rise to involuntary cognitions.

Acknowledgments

This research was supported by the Center for Human Culture and Behavior at San Francisco State University.

Appendix A

Pig Latin stimuli
Ball
Bed
Book
Box
Cake
Cat
Dog
Door
Fire
Fork
Gun
Hand
Heart
Key
Lamp
Mouse
Pipe
Ring
Saw
Top

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.

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., 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.

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*, 1445. http://dx.doi.org/10.3389/fpsyg.2014.01445.

Chomsky, N. (1988). *Language and problems of knowledge: The Managua lectures*. Cambridge, MA: MIT Press.

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.

Firestone, C., & Scholl, B. J. (2014). "Top-down" effects where none should be found: The El Greco fallacy in perception research. *Psychological Science*, *25*, 38–46.

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for 'top-down' effects. *The Behavioral and Brain Sciences* (in press).

Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: MIT Press.

Fodor, J. A. (1983). *Modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.

Freud, S. (1938). In A. A. Brill (Ed.), *The basic writings of Sigmund Freud*. New York: Modern Library (ed. & transl.).

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*, 493–503.

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.

James, W. (1890). *The principles of psychology*. New York: Dover.

Kavanagh, D. J., Andrade, J., & May, J. (2005). Imaginary relish and exquisite torture: The elaborated intrusion theory of desire. *Psychological Review*, *112*, 446–467.

Krisst, L., Montemayor, C., & Morsella, E. (2015). Deconstructing voluntary action: Unconscious and conscious component processes. In B. Eitam, & P. Haggard (Eds.), *The Sense of Agency* (pp. 25–62). New York: Oxford University Press.

Lashley, K. S. (1956). Cerebral organization and behavior. *Proceedings of the Association for Research in Nervouse and Mental Diseases*, *36*, 1–18.

Lee, T., Lim, S., Lee, K., Kim, H., & Choi, J. (2009). Conditioning-induced attentional bias for face stimuli measured with the emotional Stroop task. *Emotion*, *9*, 134–139.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.

Magee, J. C., Harden, K. P., & Teachmen, B. A. (2012). Psychopathology and thought suppression: A quantitative review. *Clinical Psychology Review*, *32*, 189–201.

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* (393–345).

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.

Merrick, C., Godwin, C. A., Geisler, M. W., & Morsella, E. (2014). The olfactory system as the gateway to the neural correlates of consciousness. *Frontiers in Psychology*, *4*, 1011. http://dx.doi.org/10.3389/fpsyg.2013.01011.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, *106*, 3–19.

Miller, B. L., & Cummings, J. L. (2007). *The human frontal lobes: Functions and disorders* (2nd ed.). New York: Guilford Press.

Miller, G. A. (1962). *Psychology: The science of mental life*. New York: Adams, Bannister, & Cox.

Miller, N. E. (1959). Liberalization of basic *S*-R concepts: Extensions to conflict behavior, motivation, and social learning. In S. Koch (Ed.), *Psychology: A study of a science, volume 2* (pp. 196–292). New York: McGraw-Hill.

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. (2016). Homing in on consciousness in the nervous system: An action-based synthesis. *Behavioral and Brain Sciences* (in press).

Morsella, E., Lanska, M., Berger, C. C., & Gazzaley, A. (2009). Indirect cognitive control through top-down activation of perceptual symbols. *European Journal of Social Psychology*, *39*, 1173–1177.

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, UK: 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.

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science*, *3*, 400–424.

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.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, *108*, 483–522.

Olsson, A., & Phelps, E. A. (2004). Learned fear of "unseen" faces after Pavlovian, observational, and instructed fear. *Psychological Science*, *15*, 822–828.

Pylyshyn, Z. W. (1984). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, MA: MIT Press.

Rassin, E. (2005). *Thought suppression*. Amsterdam, The Netherlands: Elsevier.

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *362*, 773–786.

Schacter, D. L., & Tulving, E. (1994). *Memory systems*. Cambridge, MA: The MIT Press.

Seth, A. K. (2007). The functional utility of consciousness depends on content as well as on state. *Behavioral and Brain Sciences*, *30*, 106.

Shallice, T. (1972). Dual functions of consciousness. *Psychological Review*, *79*, 383–393.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*, 946–958.

Squire, L. R. (1987). *Memory and brain*. New York: Oxford University Press.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press.

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.

Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology, vols. 1 and 2* (pp. 446–496) (4th ed.). New York, NY: McGraw-Hill.

Wegner, D. M., & Schneider, D. J. (2003). The white bear story. *Psychological Inquiry*, *14*, 326–329.

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.