Consciousness and Cognition 22 (2013) 1318–1331

Contents lists available at ScienceDirect

Consciousness and Cognition

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

Conscious thoughts from reflex-like processes: A new experimental paradigm for consciousness research

Allison K. Allen [a], Kevin Wilkins [b], Adam Gazzaley [c,d], Ezequiel Morsella [a,d,*]

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

ARTICLE INFO

Article history:
Received 6 March 2013

Keywords:
Consciousness
Ironic processing
Imagery
Unconscious processes
Implementation intentions
Action sets

ABSTRACT

The contents of our conscious mind can seem unpredictable, whimsical, and free from external control. When instructed to attend to a stimulus in a work setting, for example, one might find oneself thinking about household chores. Conscious content thus appears different in nature from reflex action. Under the appropriate conditions, reflexes occur predictably, reliably, and via external control. Despite these intuitions, theorists have proposed that, under certain conditions, conscious content resembles reflexes and arises reliably via external control. We introduce the Reflexive Imagery Task, a paradigm in which, as a function of external control, conscious content is triggered reliably and unintentionally: When instructed to not subvocalize the name of a stimulus object, participants reliably failed to suppress the set-related imagery. This stimulus-elicited content is considered 'high-level' content and, in terms of stages of processing, occurs late in the processing stream. We discuss the implications of this paradigm for consciousness research.

1. Introduction

Since the time of Descartes, the concept of the reflex arc has served as a useful mechanistic model for understanding how the nervous system works. The concept was central to the ideas of Galvani, the British Associationists, James, Freud, Ramón y Cajal, Sherrington, Pavlov, Watson, Vygotsky, Skinner, Hebb, and, more recently, LeDoux and Kandel. In the arc, external stimuli reliably elicit stereotypic behavioral responses. For example, in the pupillary reflex, decreased luminance triggers pupil dilation, reliably and unintentionally. Under normal circumstances, reflexes are predictable and under the control of external stimuli. (It should be added that this concept of the simple, automaton-like reflex has been criticized as a "convenient, if not probable, fiction" (Sherrington, 1906, p. 137) and as being incapable of explaining many cognitive functions (Koestler, 1967), such as language learning and use (Chomsky, 1959), maze learning in the rat (Tolman, 1948), and motor control (Lashley, 1951).)[1] Today, courses in psychology and neuroscience continue to present the reflex arc as a helpful, introductory concept.

Despite its vast influence on our understanding of the nervous system, the concept of the reflex seems insufficient to explain the mental phenomena of everyday life. Reflexes differ from our mentations in significant ways (Chomsky, 1988;

* Corresponding author. Address: 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] It should be noted that, though Descartes coined the term 'reflex' and was the first to provide a concrete model of reflex action, he, like James (1890) after him, did not believe that all human cognitive functions could be explained by so simple a process.

1053-8100/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.concog.2013.08.007

James, 1890; Shallice, 1972). For example, compared to reflexes, the contents of our conscious mind can seem unpredictable and insulated from the influence of external control. Regarding predictability, conscious content can appear whimsical: While attending a lecture, for example, one might find oneself thinking about household chores and then about an event from childhood—an unexpected memory triggered by the sight of a candy wrapper. This ever changing nature of conscious content led William James to construe consciousness as a mercurial and meandering stream (James, 1890). Regarding external control, it is obvious that conscious content can be free from its reign. It is commonplace to find oneself thinking about things that are unrelated to current external stimulation. When instructed to pay attention to a salient stimulus in a work setting, for instance, one may instead hum a melody "in one's head," revealing the "off-line" nature of conscious content (Barron, Riby, Greer, & Smallwood, 2011; Fodor, 1975, Fodor, 1983; Smallwood & Schooler, 2006; Wegner & Bargh, 1998). Together, these observations lead one to believe that, at times, the conscious content of our mental life is unpredictable and insulated from external control, far unlike reflexes or the workings of a clock or machine.

Despite these intuitions, scientific enquiries into the nature of the conscious mind reveal that conscious contents are far from being unpredictable and free from external control. This was noted long ago by Helmholtz (1856), who proposed that reflex-like unconscious processes can generate conscious content in a manner that resembles reflexive, unintentional actions. When speaking about such 'unconscious inferences,' Helmholtz was referring, not only to the generation of the conscious contents associated with low-level perceptual processes such as depth perception, but to higher-level, non-perceptual processes such as automatic word reading—an unnatural, intellectual process that requires years of extensive training (Helmholtz, 1856). Helmholtz noted that, when an orthographic stimulus (e.g., HOUSE) is presented to one, the stimulus automatically triggers a conscious representation of the phonological form of the word (i.e., /haus/). Seldom in everyday life is it appreciated that, in this situation, the visual stimulus triggers a conscious content that is very different in nature from that of the environmental stimulation that brought the content into existence: The conscious representation is associated, not with the visual modality, but with audition (Levelt, 1989). Helmholtz, along with many after him (Freud, 1938; James, 1890; Miller, 1959; Vygotsky, 1962; Wegner, 1989), concluded that, under the right circumstances, one could witness conscious thoughts operating in a manner that is as predictable and as susceptible to external control as are reflexes. Little is known about this manner of operation. (Interestingly, it was Neal Miller (1959), known for his behaviorist research, who proposed a forward-looking mechanistic model of conscious content that was based on sensory-response [S – R] models of behavior.)

Conscious content can be generated by Helmholtzian unconscious inferences not only in perception, but in action control, as in the case of action-related urges, which are triggered in a predictable and insuppressible manner in certain stimulus environments. For example, when one holds one's breath while under water, or runs barefoot across the hot desert sand in order to reach water, one cannot help but consciously experience the inclinations to inhale or to avoid touching the hot sand, respectively (Morsella, 2005). Regardless of the adaptiveness of the expressed actions, the conscious strife triggered by the external stimuli cannot be turned off voluntarily (Morsella, 2005; Öhman & Mineka, 2001). In these cases, the externally activated action-related urges are, in a sense, insulated (or *encapsulated*; Fodor, 1983), from voluntary control. Thus, although inclinations triggered by external stimuli can be behaviorally suppressed, they often cannot be mentally suppressed (Bargh & Morsella, 2008). One can think of many examples in which externally triggered conscious contents are more difficult to control than is overt behavior (cf., Bargh & Morsella, 2008).

It has been proposed that, in addition to unconscious inferences, other cognitive mechanisms resembling reflexes are intimately associated with conscious content. Such reflex-like mechanisms have been hypothesized to serve a role in, of all things, voluntary action—the very form of action that, historically, has been juxtaposed with reflexive processing. According to *ideomotor theory*, conscious representations about action effects automatically lead to the (unconscious) motor programs that give rise to those effects (Greenwald, 1970; Harleß, 1861; Hommel, 2009; Hommel, Müsseler, Aschersleben, & Prinz, 2001; James, 1890; Lotze, 1852). Conscious representations of one's finger flexing, for instance, automatically lead to the flexing of one's finger, unless one has activated representations of incompatible action effects, such as that of the finger *not* flexing (James, 1890). In this way, conscious representations of action effects influence motor control, which itself is largely unconscious (Fecteau, Chua, Franks, & Enns, 2001; Goodale & Milner, 2004; Grossberg, 1999; Heath, Neely, Yakimishyn, & Binsted, 2008; Jeannerod, 2006; Liu, Chua, & Enns, 2008; Rosenbaum, 2002; Rossetti, 2001). These representations of action effects can be activated directly by environmental stimuli (Hommel et al., 2001). Of import, in this framework, there is no 'decider' (a homunculus of sorts) deciding which representation of action effects should influence motor control (James, 1890; Morsella, 2009). Instead, in the absence of interference, representations of action effects (which may be activated by environmental stimuli) can automatically lead to the instantiation of the corresponding action effects, in a manner that captures aspects of the concept of the reflex arc (Hommel & Elsner, 2009; James, 1890; Vygotsky, 1962).

In summary, as unpredictable and immune from external control as our conscious mind may seem, conscious content can be activated in a manner that resembles reflexes, under certain stimulus conditions and according to some theoretical frameworks. These activations can be predictable, reliable, unintentional, and driven by external control. Moreover, these activations can result from, not only low-level, unintentional processes, such as those involving perception, but from high-level processes involving more than perceptual processing, such as the activation of conscious phonological forms from orthographic stimuli. Further understanding of the nature of these mechanisms could inform theories about the nature of consciousness, one of the greatest puzzles in science (Crick, 1995; Roach, 2005).

But how can one study these mechanisms experimentally? One might consider investigating the phenomena mentioned by Helmholtz, such as depth perception and automatic word reading. However, in part because these phenomena occur as everyday experiences, it would be unsurprising to demonstrate them in the laboratory (even though the processes underlying such phenomenon are far from trivial and far from simplistic). It would be more striking and informative to demonstrate that conscious contents can be activated via external control predictably and unintentionally even when the contents depend on high-level, sophisticated processes (beyond those of perceptual processing) for their generation. In terms of stages of processing, these contents should be based on stages in the processing stream that are far removed from the perceptuo-semantic processing of the stimulus.

To the benefit of the experimenter, in humans it is easy to find an effect that meets these criteria and can be realized feasibly in the laboratory. In terms of stages of processing, the representation of the phonological form (e.g., /haus/) of the name of an object (e.g., the picture of a house) occurs late (Dell, Martin, Saffran, Schwartz, & Gagnon, 1997; Jackendoff, 1990; Levelt, 1989; Levelt, Roelofs, & Meyer, 1999). There is a consensus in psycholinguistics that this stage of processing depends to some extent on the successful perceptual and semantic processing of the object (Dell et al., 1997; Levelt, 1989). For example, to activate the phonological representation /haus/ after being presented with a picture of a house, there must first occur the perception of the object and then its semantic processing, at least to some extent (cf., Dell et al., 1997; Morsella & Miozzo, 2002). As a 'late stage' process, the representation is more isomorphic with overt action (naming the object) than with the stimulus input (the visual stimulus; Morsella & Bargh, 2010; Sperry, 1952). In the case of 'subvocalizing,' that is, when one 'talks in one's head' and names an object in such a manner, the representation has been referred to as the 'pre-articulatory output' (Slevc & Ferreira, 2006), a term that reveals the intimate link between this kind of representation and overt action (Morsella & Bargh, 2010). (One interesting feature of subvocalization is that it is one of the quickest ways in which one can intentionally change the contents of one's consciousness.)

Unlike depth perception and object perception, this late stage of processing does not always occur automatically: It is not the case that one continuously activates the conscious phonological forms of objects that happen to fall upon the eye, at least not under everyday circumstances. Under certain conditions, however, the activation of this conscious content is likely to arise. For the purpose of illustration, we will present the reader with a demonstration of such a situation. Momentarily, we will present to you an object enclosed within parentheses. Your task is to *not* subvocalize (i.e., 'say in one's head') the name of the object. Here is the stimulus (▲). In our experience and in extensive piloting ($n = 10$), when presented with these instructions (which induce a certain *action set*) and then being presented by this stimulus, most people cannot suppress the conscious experience of the phonological form of the word 'triangle.' (As discussed below, this demonstration is based in part on the pioneering research by Wegner (1989) and Gollwitzer (1999).) Because of this phenomenon and for the reasons listed above, we selected the phonological representation as the conscious content of interest in our experimental approach, the *Reflexive Imagery Task* (RIT).

Before describing the experimental approach, we delineate the criteria that one would strive to satisfy in a paradigm in which external stimuli trigger contents into consciousness in a reflexive and systematic manner.

The paradigm should have…

1. Conscious content associated with a late stage of processing, preferably a stage associated with action production.
2. Effects that are activated unintentionally, predictably, and via external control, so as to resemble reflex action. (That the effect is unintentional renders it more similar to a reflex while also diminishing the likelihood of artifacts from idiographic, strategic processing, demand characteristics, and social desirability.)
3. A reliable and easily replicable effect that, similar to reflexes, occurs in the majority of subjects and in the majority of the experimental trials.
4. Stimulus parameters (e.g., word frequency, number of letters, word valence) that can be varied experimentally. For example, the stimulus can be presented alone or accompanied by other objects.
5. Instruction or action set parameters that can be varied systematically.
6. Effects that are predicted by several frameworks (e.g., by Gollwitzer, 1999; Miller, 1959; Morsella, 2005; Wegner, 1989) and builds incrementally on previous research (e.g., Gaskell, Wells, & Calam, 2001; Smári, 2001; Wegner, Schneider, Carter, & White, 1987).
7. Conscious content that, when used as the dependent measure, is well-studied and has well examined properties (e.g., the phonological form; Miller, 1996).
8. The ability to measure on a trial-by-trial basis the latency and neural aspects (e.g., as in neuroimaging) of the response processes involved.

All of these criteria are met by the RIT, in which, as in our demonstration above, participants are presented with the picture of objects after being instructed to not subvocalize the name of the objects. It is important to note that, when participants are instructed to not subvocalize the name of the object, they have no idea what object will be presented.

When developing the paradigm, we initially considered instructing participants to simply subvocalize the name of object (the *Intentional Naming* experimental condition) or, when presented with each object, to indicate when they found themselves subvocalizing object names in an incidental manner (the *Incidental Naming* condition), but these experimental scenarios would not satisfy *Criterion 2*, because the effects could be based on intentional processing. (Nevertheless, in our study, we provide data from these two empirical conditions.) Hence, our primary condition was the *Unintentional Naming* condition, in which the participant is instructed to not subvocalize the object's name. As discussed below and in the General Discussion, this condition is based in part on the pioneering work of Wegner (1989) on *ironic processing* and of Gollwitzer (1999) on *implementation intentions* (in which sets lead to automatic, stimulus-triggered behavioral dispositions), as well as on psycholinguistic research (Levelt, 1989; Morsella et al., 2009; Slevc & Ferreira, 2006).

When designing the paradigm, we strove to incorporate into it some of the features that Wegner and Schneider (2003) observed to be characteristic of important foundational studies and paradigms: "Classics [studies] share a simplicity that is almost embarrassing. The basic study can usually be described in a sentence... The classic study also can be summarized in a single comparison, at most two, and never, ever goes into the kinds of factorial designs that make analysis of variance teachers so proud. This simplicity promotes another feature of classic studies—what might be called *fertility*, a set of methods that make them easy to replicate and extend... [the study] survives the winnowing pressures that occur in the process of scientific consensual validation. The only way for a study to be validated properly is if just about any team in any lab can carry it off successfully and finds the procedure straightforward enough that it is not daunting" (p. 328). Regarding the depth and fertility of the investigated phenomenon, Wegner and Schneider (2003) add, "The classic studies teach us that not every interesting finding can be fully explained in the same breath with which it is reported." A component of our straightforward paradigm is ironic processing, a topic to which we now turn.

1.1. Ironic processing

Wegner's (1989) research on ironic processing reveals that one is more likely to think about something (e.g., white bears) when instructed to not think about that thing (Wegner et al., 1987; see review in Smári, 2001). In his classic paradigm, participants are instructed to not think about one thing (e.g., white bears) over a period of time (e.g., over one minute; e.g., Gaskell et al., 2001) and must indicate whenever they inadvertently think about that thing. One of Wegner's primary objectives was to shed light on the nature of self-control and the clinical implications of failures in self-control. According to his influential model of the ironic phenomenon (Wegner, 1994), the paradoxical effect arises from two distinct mechanisms that are essential for the control of action and cognition. (The model is based in part on Wiener's (1948) cybernetic theory, in which all forms of control require two mechanisms, one for performing an act and another for checking the outcome of the act against some criterion, standard, or reference value.) The first mechanism is an intentional *operating* process that actively searches for mental contents (e.g., sensations and thoughts) that will maintain a desired mental state (e.g., to be happy). This process tends to be effortful, capacity limited, and consciously mediated (Wegner, 1994). It has been proposed that the neural correlates of this mechanism involve areas of frontal cortex, such as dorsolateral prefrontal cortex (Mitchell et al., 2007).

The second mechanism is an 'ironic' *monitoring* process that automatically scans activated mental contents to detect contents signaling the failure to establish the desired state (e.g., to be happy). The detection of such goal-incongruous content (e.g., a sad memory) is essential for successful cognitive control (Cohen, Dunbar, & McClelland, 1990). Specifically, when the monitor detects contents that signify failed control of the operating mechanism, it increases the likelihood that particular content will enter consciousness, so that the operating process can process the content and change its operations accordingly. This ironic monitor mechanism is usually unconscious, autonomous, and requires little mental effort. Importantly, during mental control, the monitoring process is not reflected in conscious thought. Neuroimaging research suggests that activities of the monitor (and unintentional thoughts) are associated with the anterior cingulate cortex (Mitchell et al., 2007; Wyland, Kelley, Macrae, Gordon, & Heatherton, 2003), a region of the brain located on the medial surface of the frontal lobe and interconnected with many motor areas. This region has been associated with cognitive control (Gazzaley & Nobre, 2011), including the detection of cognitive conflict (Cohen et al., 1990), error-prone processing (Brown & Braver, 2005), and, more inclusively, inefficient processing (Botvinick, 2007).

In most circumstances, the two mechanisms—the operating and monitoring processes—work in harmony. However, such harmony fails when the goal of the system as a whole is to not activate a particular mental content (e.g., content *X*), for the following reasons. First, the operating process can only bring goal-related contents into consciousness and cannot actively exclude contents, so it cannot actively exclude content *X*, for example (Wegner, 1994). Second, the ironic monitor will reflexively bring into consciousness mental contents (e.g., content *X*) incongruent with the goal at hand. Together, these two properties will lead to the activation of content *X* in consciousness. The effect is striking: Even the threat of electric shock cannot compel people to suppress ironic thoughts (see discussion of relevant findings by McGranahan (1940) and by Sears (1943), in Wegner et al. (1987)). Various kinds of ironic effects have been documented in the literature. For example, regarding action control, participants can be tricked into moving in an unintended direction (e.g., to move leftwards when instructed to *not* move leftwards; Wegner, Ansfield, & Pilloff, 1998) as a result of ironic processing. (For reviews of ironic processing and thought suppression, see Wegner, 1989, and Rassin, 2005. For a review of ironic processing in populations suffering from various psychopathologies, see Magee, Harden, & Teachmen, 2012.)

Unlike previous paradigms, the focus of our paradigm is not so much on ironic processing per se (or on implementation intentions or linguistic processing per se), but on the nature of the unintentional action-related imagery that is influenced by a combination of stimulus conditions and action sets. In our experimental approach, ironic processes (and implementation intentions) are used as tools to render our effects unintentional. In addition, they make the experimental session more intellectually engaging than would be the case with just having the Intentional Naming and Incidental Naming conditions.

Little is known about the nuts and bolts of the RIT phenomenon. It should be noted that, at the time at which such an ironic process was first examined experimentally (1987), there was little theory regarding the paradoxical phenomenon (Wegner & Schneider, 2003). Even today the phenomenon is not fully understood (Wegner & Schneider, 2003). It is a rich psychological event that is based on several component stages and processes. Here, we outline just a subset of them, in the order in which they occur. First, there is the *induction of set* (e.g., either to subvocalize the name of presented objects or to not subvocalize the name of the objects). Second, the set must be held in mind during the time between the beginning of an experimental trial and the presentation of the object. Interestingly, during this span, there may be some occurrence of what historically has been referred to as *imageless thought* (cf., Woodworth, 1915), in which action sets can influence ongoing behavioral dispositions even though these sets do not produce any obvious, self-reportable contents in conscious awareness. In other words, the subject can perform the task and remember the instructions even though that information seems to not actively occupy consciousness in any detectable way. Third, there is the presentation of the triggering stimulus, the visual object. Under certain conditions of set, this stimulus will lead to the entry into consciousness of action-related imagery (e.g., subvocalization of the name of the object). This component process of the RIT phenomenon is interesting, because the nature of entry into consciousness of any representation remains one of the greatest enigmas in psychological science (Crick & Koch, 2003).

In our paradigm, the nature of the mental content, as well as the latency of when it enters consciousness, can be examined by use of self-report. Substantial evidence corroborates participants' reports about the occurrence of ironic conscious mental content. For example, various neuroimaging studies have revealed that the occurrence of unintentional, conscious content (including out-of-the blue cognitions, such as *stimulus-independent thought*) is coupled systematically with principled brain activations that, in these studies, were predicted *a priori* and were anticipated to be associated with the content (Mason et al., 2007; McVay & Kane, 2010; Mitchell et al., 2007; Wyland et al., 2003). In addition, behavioral evidence reveals priming effects from the activated, ironic mental content (Najmi & Wegner, 2009), suggesting that the conscious thought occurred as reported by the participant.

Before discussing the theoretical implications of this task and delineating the many potential variants of it, we present the procedures of the basic paradigm. It is our hope that this task will serve as a fecund and reliable paradigm for research on the basic mechanisms giving rise to conscious content.

2. Method

2.1. Participants

San Francisco State University undergraduate students ($n = 32$) participated for course credit.

2.2. Stimuli and apparatus

Stimuli were presented on a 50.8 cm Apple iMac computer with a viewing distance of approximately 48 cm. Stimulus presentation and data recording were controlled by PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). All questions and instructions were written in black 36-point Helevetica font. Fonts and images were displayed on a white background.

Participants were shown a series of 40 well-known objects (Table 1) displayed within in a 5.5 cm wide by 5 cm high area occupying the center of the screen (see Fig. 1). We selected visual objects that yield high 'name agreement' and that had been used successfully in previous research (Morsella & Miozzo, 2002; Snodgrass & Vanderwart, 1980). These 40 objects were

Table 1
List of the visual objects (line drawings).

List 1	List 2
Anchor	Apple
Ball	Balloon
Cake	Bed
Candle	Book
Dog	Cat
Envelope	Chair
Fire	Door
Glasses	Flower
Hand	Fork
Key	Gun
Lamp	Heart
Mouse	House
Pencil	Ladder
Radio	Mouth
Saw	Pipe
Stool	Ring
Tie	Scissors
Trophy	Star
Umbrella	Tree
Wheel	Whistle

divided into two lists of 20 objects. We took care to make sure that the objects within each block would be as dissimilar as possible from each other, both in terms of phonology (e.g., beginning with different phonemes as often as possible) and in terms of perceptuo-semantic features (e.g., no two objects could categorized as the same kind of object). Participants were shown each object only once, in random order. Presentation order of the stimulus lists, and the matching of lists to the three conditions, was fully counterbalanced across participants. (This was done for all but two participants: As mentioned below, in one case, the participant's data had to be excluded from analysis because the participant failed to follow instructions; in the other case, the participant was accidentally run on List A when he or she should have been run on List B.) Importantly, there was no effect of List Type (List A versus B). Specifically, there were no main effects or interaction effects of the factor *List* or *List Sequence* on subvocalization rates, $F\text{s} < 1.8$, $p\text{s} > .18$, or on the other dependent variables, $F\text{s} < 2.4$, $p\text{s} > .10$.

2.3. Procedure

All participants first completed the baseline, Incidental Naming condition (trials = 20), in which, upon being presented with an object, participants indicated if they happened to think of the name of the object. This condition involved no explicit instructions regarding naming or not naming and served to assess (at least to some extent) participants' spontaneous subvocalization rates in response to the stimuli when having no obvious action set toward the stimuli. (Mentioning to participants the possibility of subvocalizing naturally influences the likelihood of subvocalizations, which is a limitation of our baseline condition.) For this baseline condition, instructions presented on the computer screen informed participants that they would be shown a series of visual objects. For this and all the other conditions, participants were instructed to keep their eyes focused on the center of the screen at all times. Participants were instructed to press the space bar with their left hand as soon as they happened to think of the name of the object that was presented on a particular trial. This served as our primary dependent measure. If participants did not happen to think the name of the object, they were instructed to not respond in any way. It was emphasized to participants to press the space bar as quickly as possible as soon as they happened to think of the name of the object. Before beginning every trial, participants were reminded to keep their left hand rested on the space bar at all times. Participants were informed that images would remain on the screen for a fixed period of time, whether or not they pressed the space bar.

Before each trial, a question mark (?) was displayed in the center of the screen, asking participants if they were ready to begin the trial, which they indicated 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 fixation, an object appeared for 4 s, during which time participants could indicate by pressing the space bar if they happened to think the name of the object. Prior to completing the Incidental Naming trials, participants completed one practice trial to familiarize themselves with the task.

After the Incidental Naming condition, half of the participants ($n = 16$) performed the Intentional Naming condition (Participant cohort A) and half performed the Unintentional Naming condition (Participant cohort B), with each condition consisting of 20 trials. The experimental factor of Naming Condition (Intentional versus Unintentional) was a between-subjects factor and not a within-subjects factor because we were attempting to limit artifacts from carry-over effects that would be encountered with the latter.

For each condition, we did not want trial numbers to exceed 20, in order to diminish habituation effects from continuous performance in the experimental conditions, which took place after 20 trials from the Incidental Naming condition. In the Intentional Naming condition, participants were instructed to subvocalize the name of the object. In the Unintentional Naming condition, participants were instructed to *not* subvocalize the name of the object. (The term 'subvocalization' was never used in the instructions because prior research from our laboratory revealed to us that, for many participants, the term is difficult to understand.) Instructions in the second block were identical in all respects to those of Incidental Naming condition except for the following instructions. For the Intentional Naming condition, participants were instructed: "*You will be shown a new series of images. This time, TRY to think of the name of the object that is presented*." For the Unintentional Naming condition, participants were instructed: "*You will be shown a new series of images. Try to NOT think of the name of the object that is presented*." The trials of the second block unfolded in a fashion identical to that of the Incidental Naming condition except for the following. Before each trial, rather than being presented with a question mark (?) prompt, participants were presented with either the reminder, "Think of the Name of the Object" or "Don't Think of the Name of the Object" for the Intentional Naming and Unintentional Naming conditions, respectively. Participants would then press the return key to indicate readiness to begin each trial.

Once participants completed the experiment, they responded to a series of funneled debriefing questions (following the procedures of Bargh and Chartrand, 2000), which included general questions to assess whether (a) participants were aware of the purpose of the study, (b) they had any strategies for completing the task, and (c) anything interfered with their performance on the task. Importantly, funneled debriefing also queried participants about whether or not they used a strategy when trying to perform well in the Intentional Naming and Unintentional conditions. Because the study included participants who were non-native speakers of English, we also included a series of questions to assess whether (a) participants thought the name of the object in a language other than English, (b) they pressed the space bar in such a situation, and (c) they had a strategy for completing the task if they happened to think of the name of the object in more than one language. Participants who spoke languages in addition to English pressed the space bar even if the object name was subvocalized in a language other than English. The trial data from one participant (Participant 7 in Table 2, Section A) were excluded from

Table 2
Number of trials (out of 20) in which subvocalizations occurred as a function of participant and naming condition.

Participant	Incidental	Intentional	Unintentional
Section A			
1	20	20	
2	20	20	
3	20		10
4	20		14
5	20	20	
6	20	20	
7			
8	19		19
9	16	19	
10	20	20	
11	20		19
12	20		
13	20	20	
14	20	20	
15	20		20
16	20		20
17	20	20	
18	20	20	
19	20		3
20	20		18
21	19	20	
22	20	20	
23	19		18
24	20		20
25	20	20	
26	20	20	
27	20		20
28	20		16
29	20	20	
30	20	20	
31	20		20
32	20		20

Variables	Both blocks (40 trials)	Incidental (20 trials)	Intentional (20 trials)	Unintentional (20 trials)
Section B. Correlational analysis of supplementary measures				
Latency, automaticity	−.36[a]	−.40	−.439	−.30
Latency, urge to vocalize	−.04	−.05	.15	−.09
Urge to vocalize, automaticity	.09	.09	−.18	.14

[a] Denotes Fisher's r to z, $p < .05$, given the number of observations.

analysis because the participant did not follow instructions and, during the middle of the baseline condition, ceased inputting any responses or ratings.

2.4. Primary dependent measure: entry into consciousness of the phonological form

Our primary dependent measure was whether participants reported consciousness of the phonological form of the object name, an event which, in the Unintentional condition, can result from unintentional subvocalization. To establish the reliability of the phenomenon, we recorded the number of trials in which such an event occurred. We examined whether there was a significant difference between the proportions of trials having such subvocalizations in the Intentional and Unintentional conditions. For such an assessment, we carried out t-tests (unpaired) on both the raw proportion data and the arcsine transformed version of the proportion data. (Arcsine transformations are often used to statistically normalize data that are in the form of proportions.) In addition to t-tests, to examine this contrast in light of any potential group effects (Cohort A: *Incidental Naming followed by Intentional Naming*; Cohort B: *Incidental Naming followed by Unintentional Naming*), and to include the comparison of the baseline condition, we also conducted a more comprehensive, mixed-design 2×2 ANOVA, in which, in addition to the within-subject contrast between conditions (Baseline Block versus Critical Block [Intentional or Unintentional Naming]), Cohort (A or B) was included as a between-subjects factor. Because this analysis takes into account differences resulting from the peculiarities of the particular cohort, it is a more conservative test than a simple unpaired t-test between the Intentional and Unintentional conditions. Hence, for all reported ANOVAs, Cohort was included as a factor, to assess potential artifacts stemming from the peculiarities of the cohorts of participants.

2.5. Supplementary, trial-by-trial measures

In addition to our primary measure, in order to satisfy *Criterion 8* and begin to learn more about the nature of the responses in the RIT phenomenon, we took the opportunity to include some supplementary dependent variables that could be measured on a trial-by-trial basis. For this initial study, we examined the (a) latency of entry into consciousness of the phonological form, (b) the perceived automaticity of the response, and (c) the degree to which the response influenced dispositions toward overt action, as in the case of the urge to utter the name of the word aloud. (One can imagine variants of the paradigm with other kinds of questions presented on a trial-by-trial basis.) Because participants used their left hand to press the space bar to indicate the experience of a subvocalization, these ratings were inputted with the right hand, using the key pad on the right side of the keyboard. We now discuss each of these supplemental measures.

2.6. Latency of entry into consciousness of phonological forms

Our latency measure consisted of the time span between the onset of the presentation of the stimulus object and the button press made by the participant, indicating consciousness of the phonological form. The continuous, trial-by-trial latency data were then correlated to the other supplementary measures and also submitted to a mixed-design, 2×2 ANOVA, in which, in addition to the within-subject contrast between conditions (Baseline versus Critical Block), Cohort (A or B) was included as a between-subjects factor.

2.7. Perceived automaticity

After each trial, participants were first asked, "*If you thought the name of the object, to what extent did you feel that the thought occurred automatically?*" Participants inputted a rating between 1 and 8, in which 1 indicated that thinking of the name of the object did not occur automatically and that they had complete control over its occurrence, and 8 indicated that it occurred completely automatically and that they had no control over its occurrence. If participants did not think the name of the object, they simply inputted 0. (Zeroes such as these were never included in any subsequent statistical analyses.) These continuous, trial-by-trial data were then correlated to the other supplementary measures, submitted to between-subjects *t*-tests (to compare, for example, the Intentional and Unintentional conditions) and a mixed-design 2×2 ANOVA, which, as explained above, included Cohort as a between-subjects factor.

2.8. Urge to utter object name aloud

The second question participants were asked after each trial was, "*To what extent did you feel the urge to say the object name aloud?*" Urge to utter aloud was included to assess further the well-known inter-relationship between imagery and overt action (Hebb, 1968; Morsella & Bargh, 2010; Slevc & Ferreira, 2006; Sperry, 1952). For this rating, 1 indicated no urge to say the name of the object aloud, and 8 indicated an extremely strong urge to say the name of the object aloud. Again, participants were instructed to input 0 if they did not happen to think the name of the object. These continuous, trial-by-trial urge data were then correlated to the other supplementary measures, submitted to between-subjects *t*-tests (to compare, for example, the Intentional and Unintentional conditions) and a mixed-design 2×2 ANOVA, which, as mentioned above, included Cohort as a between-subjects factor.

Thus, after being presented with a stimulus object, participants were asked the two aforementioned questions. It is important to note that these supplementary measures were added primarily to (a) showcase the kinds of dependent measures (including trial-by-trial measures) that one could include in the paradigm, (b) satisfy *Criterion 8*, and (c) benefit future investigations. The present experiment was not designed to test hypotheses concerning how the supplementary measures of latency, urges to utter the object name aloud, or perceived automaticity are influenced by the naming conditions.

3. Results and discussion

3.1. Primary dependent measure: entry into consciousness of the phonological form

Most importantly, subvocalizations occurred for the majority of participants and for the majority of trials (Table 2, Section A). In the Unintentional condition, only 4 participants experienced subvocalizations on less than 90% of the trials. Participants 3 and 19 appeared to be successful at suppressing the undesired subvocalizations. See descriptive statistics for all dependent measures in Table 3. Regarding the Incidental Naming (Baseline) condition, it was surprising to find such high rates of subvocalizations. The significant interaction pattern in Fig. 2, $F(1, 29) = 8.66$, $p = .006$, reveals that the Incidental Naming condition yielded comparable subvocalizations for both groups of participants (i.e., Cohort A or B), $t(29) < 1$, $p = .73$, and that there was a significant contrast between the Intentional and Unintentional conditions, $t(29) = 2.90$, $p = .007$. (Again, for all ANOVAs, Cohort [i.e., participant group A or B] was included as an experimental factor to assess potential artifacts stemming from the peculiarities of the cohorts of participants.) In this particular ANOVA on the proportions, we used arcsine transformations of the data. In addition to the interaction, there was a main effect of Cohort, $F(1, 29) = 5.92$,

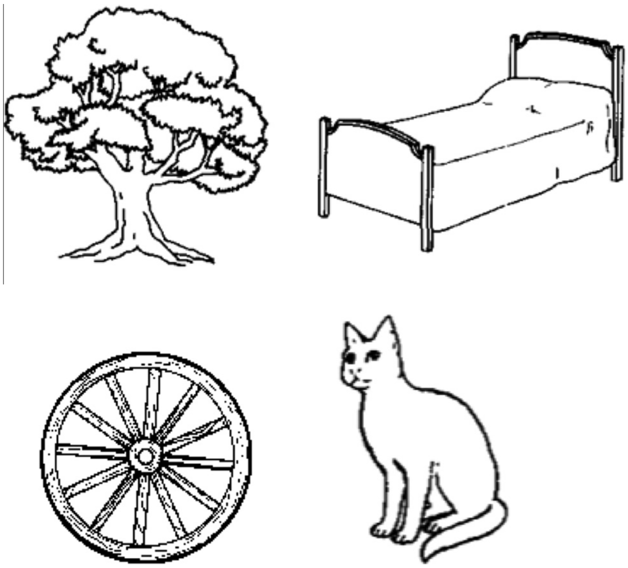

Fig. 1. Sample visual objects. (Not drawn to scale.)

Table 3
By-subject mean values as a function of condition. *SD*s presented in parentheses.

Naming condition	Proportion of trials with subvocalizations	Latency (ms)	Urges to utter name aloud	Automaticity
Incidental (Baseline)	.99 (.04) Range: .80–1.0	706.69 (234.38) *SEM* = 42.10	3.08 (1.98)	6.98 (.94)
Intentional	.99 (.01) Range: .95–1.0	685.82 (257.23) *SEM* = 64.31	2.81 (1.87)	7.35 (.47)
Unintentional	.86 (.24) Range: .15–1.0	1,451.27 (611.42) *SEM* = 157.87	2.65 (1.80)	6.59 (1.15)

p = .02, and a main effect of Block (Baseline Block versus Critical Block), $F(1, 29) = 5.37$, $p = .03$. The same pattern of results is obtained with the untransformed, proportion data. (A by-item analysis, in which visual object is the unit of analysis, leads to similar results, $F [2, 78] = 76.002$, $p < .0001$.) The effects of condition on subvocalization rate survive in ANOVAs in which Stimulus List and Cohort are included as an experimental factor.

3.2. Latency of entry into consciousness of phonological forms

An ANOVA with response latency as the dependent variable revealed an interaction in which the Unintentional condition yielded the longest latencies, $F(1, 29) = 29.87$, $p < .0001$ (Fig. 3), a contrast corroborated by a planned comparison between Unintentional and Intentional conditions, $t(29) = 4.597$, $p < .0001$. There was also a main effect of Cohort, $F(1, 29) = 11.86$, $p = .0018$, and a main effect of Block (Baseline Block versus Critical Block), $F(1, 29) = 27.66$, $p < .0001$. It seems that subvocalizations occur later when they are unintended than when they are intended. One parsimonious explanation for this effect is the following. When instructed to perform a simple action (e.g., naming an object), participants are likely to do it as quickly as possible, potentially leading to floor effects in latency. For instance, on a given trial in the Intentional condition, when a participant is presented with a stimulus object (e.g., the image of a cat), he or she immediately subvocalizes the name intentionally (e.g., ‘‘cat’’). However, more variability would be expected on trials in which participants are instructed to not subvocalize the name of the object. This variability is unlikely to lead to a floor effect in latency, thereby yielding longer mean latencies. In summary, the latency data suggest that participants are performing differently in the Intentional and Unintentional Naming conditions.

3.3. Perceived automaticity

Across the different kinds of Naming conditions, mean ratings of perceived automaticity of the experienced subvocalizations were high, ranging from 6.59 (*SEM* = .30) in one condition block to 7.35 (*SEM* = .19) in another condition (Table 3). Importantly, there was no interaction among the experimental factors represented in Fig. 4, $F(1, 29) = 1.08$, $p = .31$, and there was no main effect of Block (Baseline Block versus Critical Block), $F(1, 29) = .39$, $p = .54$, suggesting that the perceived

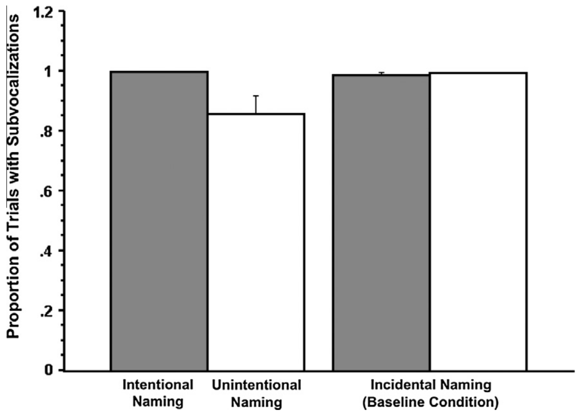

Fig. 2. Proportion of trials in which subvocalizations occurred, as a function of condition. The shaded bars indicate the group of participants that, following the baseline condition, performed the Intentional Naming condition. The unfilled bars indicate the group that, following the baseline condition, then performed the Unintentional Naming condition. Error bars indicate *SEM*s.

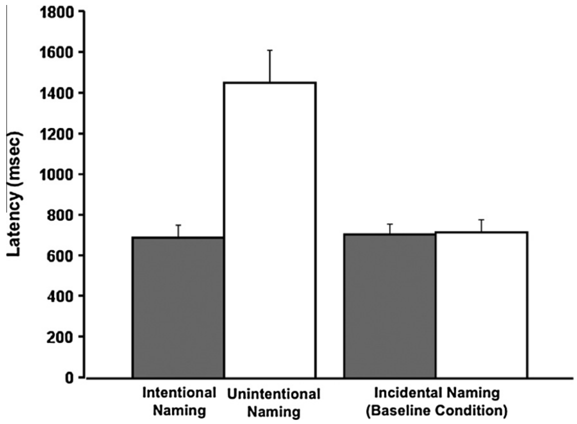

Fig. 3. Mean latency of subvocalization, as a function of condition. The shaded bars indicate the group of participants that, following the baseline condition, performed the Intentional Naming condition. The unfilled bars indicate the group that, following the baseline condition, then performed the Unintentional Naming condition. Error bars indicate *SEM*s.

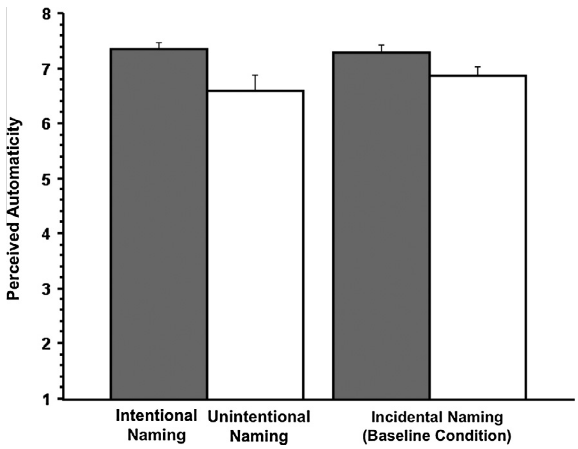

Fig. 4. Mean perceived automaticity of subvocalization, as a function of condition. The shaded bars indicate the group of participants that, following the baseline condition, performed the Intentional Naming condition. The unfilled bars indicate the group that, following the baseline condition, then performed the Unintentional Naming condition. Error bars indicate *SEM*s.

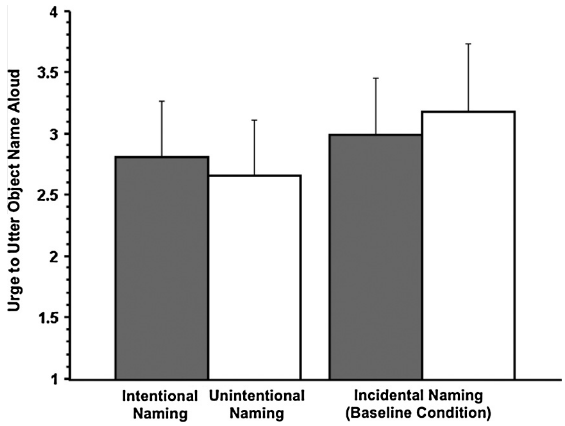

Fig. 5. Mean urge to utter the object name aloud, as a function of condition. The shaded bars indicate the group of participants that, following the baseline condition, performed the Intentional Naming condition. The unfilled bars indicate the group that, following the baseline condition, then performed the Unintentional Naming condition. Error bars indicate *SEM*s.

automaticity of the subvocalization response was comparable in the Intentional Naming and Unintentional Naming conditions, with both conditions being more associated with the high end rather than the low end of the automaticity continuum (i.e., that subvocalization occurred completely automatically with little control over its occurrence). However, there was an unexpected main effect of Cohort, F (1, 29) = 7.72, p = .01. That the ratings of automaticity were so high requires further investigation. In summary, the perceived automaticity of the subvocalization response, for both the Intentional and Unintentional conditions, was on the high end of the spectrum.

3.4. Urge to utter object name aloud

Regarding urges, for all naming conditions, mean urges were significantly higher than a value of 1, which indicates ''no urge to say the name of the object aloud,'' *t*s > 3.5, *p*s < .01 (Table 3, Fig. 5). The lowest mean value was 2.65. This observation begins to reveal a new link between action-related imagery (in our case, the pre-articulatory output; Slevc & Ferreira, 2006) and overt behavior (Morsella & Bargh, 2010). Importantly, there were no interaction effects involving urges, F(1, 29) = 1.10, p = .30 (Fig. 5), and no main effect of Cohort, F(1, 29) = .001, p = .98, but there was an unexpected main effect in which urges to utter were stronger in the baseline condition, F(1, 29) = 4.56, p = .04. One parsimonious explanation for this effect is that, in Block 2, which presumably is after some habituation to the trial events that have transpired in Block 1, urges become weaker. In summary, these data suggest that urges to utter the name aloud are far from nonexistent, even in the Unintentional Naming condition, in which participants are instructed to not subvocalize the name of the presented object.

3.5. Correlational analysis of supplementary measures

For the correlations between the supplementary measures, and the value of correlation coefficients as a function of these condition, see the bottom of Table 2, Section B. The only correlation of interest involving the supplementary measures was an inverse relationship between automaticity and latency, $r = -.36$, Fisher's r to z, $p < .05$: As latency increases, perceived automaticity decreases. This could be explained parsimoniously by that fact that intended acts of subvocalizing tend to occur earlier in the trial, for the reasons mentioned above (see *Latency of Entry into Consciousness of Phonological Form*). It is important to note that this was the mean correlation coefficient per subject, with both blocks of trials (total = 40 trials) collapsed. With the current data, assessing the strength of the correlation by Naming condition relies on an insufficient number of trial data (i.e., 20 trials per condition).

4. General discussion

In everyday life, conscious contents can seem unpredictable and insulated from the direct influence of external stimuli. Our paradigm reveals that, under certain circumstances and in accord with several theoretical approaches (Helmholtz, 1856; Hommel, 2009; Hommel et al., 2001; James, 1890; Miller, 1959; Morsella, 2005; Vygotsky, 1962; Wegner, 1989), conscious contents can be elicited in a manner that is unintentional, reflex-like, and nontrivial. In general, the conscious effects in our paradigm depended on a combination of set and stimulus conditions. Importantly, the paradigm satisfies the eight criteria listed above. The RIT phenomenon is a rich phenomenon that can be mined experimentally in many ways. We will not pretend to understand all aspects of what occurs in the Unintentional condition, an interesting experimental condition that includes the involuntary entry of high-level contents into consciousness.

As mentioned above, in the Unintentional condition, participants may try different strategies to suppress the unintended subvocalizations. Regarding successful suppression, participants may be using various strategies, including that of self-distraction or 'negative cueing' (Wegner et al., 1987; see evidence of positive effects from self-distraction in Hertel and Calcaterra, 2005). When speaking about 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). Occupying the mind with verbal information may indeed delay entry of the subvocalization. In the current study, the funnel debriefing data revealed that 8 participants attempted to suppress the subvocalizations by subvocalizing about other objects or by thinking about something else (e.g., subvocalizing a melody). Some participants reported that they attempted to succeed at the task by looking only at parts of the object and not at the object as a whole. Interestingly, participants who had expert knowledge about certain objects (e.g., trees) seem to have used more precise terms when naming such objects subvocally, according to the reports of one participant. Some participants expressed that, when being told not to think about the object name, the instruction made them more likely to think about the object name. Which strategies are most successful in suppressing the subvocalizing is worthy of future investigation.

Regarding successful suppression, at this stage of understanding, we remain agnostic regarding whether the performance of participants (e.g., as reflected in their latencies) is consistent with 'inhibition' accounts of cognitive control (cf., Aron, 2007; Levy & Wagner, 2011) or with other accounts concerning the control of action and cognition. Regarding the latter, some Jamesian ideomotor approaches would interpret successful suppression as resulting, not from direct inhibition of the undesired action plan, but from the sustained activation of an incompatible action plan (see Hommel, 2009). In short, future investigations could further mine the phenomenon of the RIT and determine whether it is consistent with an inhibition account of control or with other contemporary accounts (e.g., ideomotor accounts).

Interactions with participants and participants' answers to our funnel debriefing questions never led to the identification of a participant who reported that he or she experienced an ironic subvocalization but failed to report it, for some reason. Moreover, in our paradigm, if participants experience the ironic mental content but immediately forget that they did so, because, say, the experience was fleeting, then it is not the kind of robust, notable (and ecologically-valid) instantiation of a conscious thought in which we are interested. In our experiment, a thought event involves a thought that is easy to note and report about on the part of the participant. In everyday life, when one states that one just thought of something, one is usually speaking about this kind of detectable thought. Focusing on such notable thoughts renders our experiment more ecologically valid and easier to replicate.

It is important to note that there are well known limitations regarding self-report measures. For example, self-reports can be inaccurate as a result of participants basing their responses on, not what they actually experience, but heuristics (see Morsella et al., 2009). In addition, as discussed above, inaccurate memory of fleeting mental contents can lead to incorrect self-reports (Block, 2007). Moreover, under certain circumstances (see discussion in Bayne (2013)), there can be what appears to be accurate self-report in the absence of consciousness of the content on which the report is based. Given the striking reliability of the RIT phenomenon at hand, as experienced by both the participants and the experimenters (and perhaps even the reader, in response to our example of the triangle, above), we do not believe that these well known limitations undermine the validity of our primary finding.

The RIT builds on the pioneering investigations by Wegner (1989) on ironic processing. One difference between the RIT and previous research involving ironic processing is that the RIT involves the kinds of punctate stimuli and dependent measures used in psycholinguistic research. In addition, the focus of the present research is not on ironic processing per se, but on how stimuli can activate action-related imagery reliably, via set induction and external control. The RIT builds also on the important research by Gollwitzer (1999) on implementation intentions by demonstrating that, once certain sets are induced, responses to environmental stimuli can resemble reflex-like processes, even when the responses depend on sophisticated unconscious inferences. Future studies could examine the limits of such unintentional responses, thereby revealing the limits of unconscious inferences. For example, it is unlikely that an effect would arise when the instruction is to, say, refrain from doing a complicated operation upon the stimuli (e.g., long division when presented with numbers).

Our findings are consonant with the theorizing by Miller (1959), who proposed that conscious content is more constrained than appears to be the case and that, under the right conditions, content activation can resemble S – R processes. From this standpoint, the unpredictability of conscious content during everyday life may reflect, not so much the working of an *un*-mechanistic and indeterminate system, but the vagaries of quotidian stimulus conditions. Under the correct conditions, conscious content can be highly predictable and externally constrained. Accordingly, it has been theorized that, in nervous function, conscious thoughts should be construed as a highly constrained output (Wundt, 1902/1904), the result of 'multiple constraint satisfaction' (Merker, 2007). In some cases, elicitations of conscious content is easier to predict and control than is overt behavior (Bargh & Morsella, 2008), as discussed above.

Future investigations using this paradigm may examine the neural correlates of the various events involved in the RIT phenomenon (e.g., induction of set, maintenance of set in short term memory, and entry into consciousness of intrusive cognition). Future variants could also be used to investigate the effects of varying stimulus parameters, such as word frequency (e.g., latency of entry of object names with high versus low frequency), valence (e.g., positive versus negative), and the semantic properties of the stimulus objects (e.g., artifacts, food, or animals). In addition, the paradigm may be useful to use with special populations, such as those suffering from aphasia, in which overt naming is impaired but subvocalizing can be spared, and from psychopathological phenomena in which intrusive cognitions play a role (e.g., ruminations and obsessions; see Magee et al., 2012). In future versions of the paradigm, participants could also provide trial-by-trial ratings about aspects of their experience that were not assessed in the present experiment, such as the degree of effort expended on each trial.

It is our hope that the RIT will be used to illuminate a range of psychological phenomena. The paradigm combines component processes investigated in disparate subfields of psychological science, including cognitive control, self-regulation, imagery, working memory, action control, and action sets. More generally, the RIT provides a reliable and mechanistic paradigm for research in psychological science on what is an enigmatic topic—consciousness.

Acknowledgment

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

References

Aron, A. R. (2007). The neural basis of inhibition in cognitive control. *The Neuroscientist, 13*, 214–228.
Bargh, J. A., & Chartrand, T. L. (2000). A practical guide to priming and automaticity research. In H. Reis & C. Judd (Eds.), *Handbook of research methods in social psychology* (pp. 253–285). New York: Cambridge University Press.
Bargh, J. A., & Morsella, E. (2008). The unconscious mind. *Perspectives on Psychological Science, 3*, 73–79.
Bargh, J. A., & Morsella, E. (2010). Unconscious behavioral guidance systems. In C. R. Agnew, D. E. Carlston, W. G. Graziano, & J. R. Kelly (Eds.), *Then a miracle occurs: Focusing on behavior in social psychological theory and research* (pp. 89–118). New York: Oxford University Press.
Barron, E., Riby, L. M., Greer, J., & Smallwood, J. (2011). Absorbed in thought: The effect of mind wandering on the processing of relevant and irrelevant events. *Psychological Science, 22*, 596–601.
Bayne, T. (2013). Agency as a marker of consciousness. In A. Clark, T. Vierkant, & J. Kiverstein (Eds.), *Decomposing the will* (pp. 160–181). Oxford, UK: Oxford University Press.
Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences, 30*, 481–548.
Botvinick, M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective and Behavioral Neuroscience, 7*, 356–366.
Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science, 307*, 1118–1121.
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.
Chomsky, N. (1959). Review of "Verbal Behavior" by Skinner BF. *Language, 35*, 26–58.
Chomsky, N. (1988). *Language and problems of knowledge: The Managua lectures*. Cambridge, MA: MIT Press.
Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*, 332–361.
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. (1995). *The astonishing hypothesis: The scientific search for the soul*. New York: Touchstone.
Crick, F., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience, 6*, 1–8.
Dell, G. S., Martin, N., Saffran, E. M., Schwartz, M. F., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104*, 801–838.
Fecteau, J. H., Chua, R., Franks, I., & Enns, J. T. (2001). Visual awareness and the online modification of action. *Canadian Journal of Experimental Psychology, 55*, 104–110.
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.
Gazzaley, A., & Nobre, A. C. (2011). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Science, 16*, 129–135.
Gaskell, S. L., Wells, A., & Calam, R. (2001). An experimental investigation of thought suppression and anxiety in children. *British Journal of Clinical Psychology, 40*, 45–56.
Goodale, M., & Milner, D. (2004). *Sight unseen: An exploration of conscious and unconscious vision*. New York: Oxford University Press.
Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist, 54*, 493–503.
Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review, 77*, 73–99.
Grossberg, S. (1999). The link between brain learning, attention, and consciousness. *Consciousness and Cognition, 8*, 1–44.
Harleß, E. (1861). Der apparat des willens [The apparatus of the will]. *Zeitshrift für Philosophie und philosophische Kritik, 38*, 499–507.
Heath, M., Neely, K. A., Yakimishyn, J., & Binsted, G. (2008). Visuomotor memory is independent of conscious awareness of target features. *Experimental Brain Research, 188*, 517–527.
Hebb, D. O. (1968). Concerning imagery. *Psychological Review, 75*, 466–477.
Helmholtz, H. v. (1856). Treatise of physiological optics: Concerning the perceptions in general. In T. Shipley (Ed.), *Classics in psychology* (pp. 79–127). New York: Philosophy Library.
Hertel, P. T., & Calcaterra, G. (2005). Intentional forgetting benefits from thought substitution. *Psychonomic Bulletin & Review, 12*, 484–489.
Hommel, B. (2009). Action control according to TEC (theory of event coding). *Psychological Research, 73*, 512–526.
Hommel, B., & Elsner, B. (2009). Acquisition, representation, and control of action. In E. Morsella, J. A. Bargh, & P. M. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 371–398). New York: Oxford University Press.
Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding: A framework for perception and action planning. *Behavioral and Brain Sciences, 24*, 849–937.
Jackendoff, R. S. (1990). *Consciousness and the computational mind*. Cambridge, MA: MIT Press.
James, W. (1890). *The principles of psychology*. New York: Dover.
Jeannerod, M. (2006). *Motor cognition: What action tells the self*. New York: Oxford University Press.
Koestler, A. (1967). *The ghost in the machine*. London: Hutchinson.
Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior. The Hixon symposium* (pp. 112–146). New York: Wiley.
Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.
Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1–75.
Levy, B. J., & Wagner, A. D. (2011). Cognitive control and right ventrolateral prefrontal cortex: Reflexive orienting, motor inhibition, and action updating. *Annals of the New York Academy of Sciences, 1224*, 40–62.

Liu, G., Chua, R., & Enns, J. T. (2008). Attention for perception and action: Task interference for action planning, but not for online control. *Experimental Brain Research, 185*, 709–717.
Lotze, R. H. (1852). *Medizinische Psychologie oder Physiologie der Seele*. Leipzig: Weidmann'sche Buchhandlung.
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*, 393–395.
McGranahan, D. V. (1940). A critical and experimental study of repression. *Journal of Abnormal and Social Psychology, 35*, 212–225.
McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schoolor (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.
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 science, Study 1* (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..
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. (2010). What is an output? *Psychological Inquiry, 21*, 354–370.
Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 555–563.
Morsella, E. (2009). The mechanisms of human action: Introduction and background. In E. Morsella, J. A. Bargh, & P. M. Gollwitzer (Eds.), *The Oxford handbook of human action* (pp. 1–32). New York: Oxford University 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.
Najmi, S., & Wegner, D. M. (2009). Hidden complications of thought suppression. *International Journal of Cognitive Therapy, 2*, 210–223.
Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108*, 483–522.
Rassin, E. (2005). *Thought suppression*. Amsterdam, The Netherlands: Elsevier.
Roach, J. (2005). Journal ranks top 25 unanswered science questions. National Geographic News. Retrieved from news.nationalgeographic.com.
Rosenbaum, D. A. (2002). Motor control. In H. Pashler & S. Yantis (Eds.). *Stevens' handbook of experimental psychology: Sensation and perception* (3rd ed) (vol. 1, pp. 315–339). New York: Wiley.
Rossetti, Y. (2001). Implicit perception in action: Short-lived motor representation of space. In P. G. Grossenbacher (Ed.), *Finding consciousness in the brain: A neurocognitive approach* (pp. 133–181). Netherlands: John Benjamins Publishing.
Sears, R. R. (1943). *Survey of objective studies of psychoanalytic concepts*. New York: Social Science Research Council.
Shallice, T. (1972). Dual functions of consciousness. *Psychological Review, 79*, 383–393.
Sherrington, C. S. (1906). *The integrative action of the nervous system*. New Haven, CT: Yale University Press.
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.
Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin, 132*, 946–958.
Smári, J. (2001). Fifteen years of suppression of white bears and other thoughts: What are the lessons for obsessive–compulsive disorder research and treatment? *Scandinavian Journal of Behavior Therapy, 30*, 147–160.
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.
Sperry, R. W. (1952). Neurology and the mind-brain problem. *American Scientist, 40*, 291–312.
Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55*, 189–208.
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., Ansfield, M., & Pilloff, D. (1998). The putt and the pendulum: Ironic effects of the mental control of action. *Psychological Science, 9*, 196–199.
Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D. T. Gilbert, S. T. Fiske, & G. Lindsey (Eds.). *The handbook of social psychology* (4th ed) (Vol. 1, pp. 446–496). Boston, MA: 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.
Woodworth, R. S. (1915). A revision of imageless thought. *Psychological Review, 22*, 1–27.
Wiener, N. (1948). *Cybernetics*. New York: Wiley.
Wundt, W. (1902/1904). *Principles of physiological psychology* (Translated from the Fifth German Edition (1904) by E. B. Titchener). London: Swan Sonnenschein.
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.