Published in final edited form as:

Brain Res. 2010 February 26; 1316C: 62. doi:10.1016/j.brainres.2009.12.041.

Why the White Bear is Still There: Electrophysiological Evidence for Ironic Semantic Activation during Thought Suppression

Ryan J. Giuliano¹ and Nicole Y. Y. Wicha^{1,2}

¹The University of Texas at San Antonio

²The University of Texas Health Science Center at San Antonio

Abstract

Much research has focused on the paradoxical effects of thought suppression, leading to the viewpoint that increases in unwanted thoughts are due to an ironic monitoring process which increases the activation of the very thoughts one is trying to rid from consciousness. However, it remains unclear from behavioral findings whether suppressed thoughts become more accessible during the act of suppression. In the current study, event-related potentials were recorded while participants suppressed or expressed thoughts of a focus word during a simple lexical decision task. Modulations in the N400 component reported here demonstrate the paradoxical effects occurring at the semantic level during suppression, as well as some evidence for the rebound effect after suppression periods. Interestingly, semantic activation was greater for focus words during suppression than expression, despite differences in the N1 window suggesting that expression elicited greater perceptual processing than suppression. Results provide electrophysiological evidence for the Ironic Process model and support recent claims of asymmetric network activation during thought suppression.

Keywords

thought suppression; semantic processing; executive control; N400; event-related potentials; lexical decision

1. Introduction

Unwanted thoughts have an unusual gravity in that they seem to bounce back into our awareness with even the slightest cue, and once succumbed to, the thoughts seem inescapable. Up to 80% of the general population have experienced intrusive unwanted thoughts in daily life (Rachman and de Silva, 1978), such as when asked to keep a secret and despite our best efforts the guarded words escape during casual conversation before we can stop them. The paradoxical effects of thought suppression were first demonstrated in a study by Wegner and colleagues (1987) where thoughts of a white bear were found more likely to recur for people who initially suppressed thoughts of a white bear than for people who had not been instructed to suppress such thoughts. The finding of a post-suppression rebound effect has since been replicated several times, in a

© 2009 Elsevier B.V. All rights reserved.

Corresponding author: Nicole Y. Y. Wicha, Department of Biology, University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, Tel: (210) 458-7013, Fax: (210) 458-5658, wicha@uthscsa.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

variety of investigations on how actively suppressing a thought increases the likelihood of their subsequent intrusion into conscious awareness (for reviews, see Abramowitz et al., 2001; Wenzlaff and Wegner, 2000). Within clinical populations, thought suppression may have a potential role as a causal factor in a vast spectrum of psychopathologies, where the experience of adverse mental states might prompt a suppressive response, including, but not limited to, posttraumatic stress disorder, obsessive-compulsive disorder, depression, and generalized anxiety disorder (Najmi and Wegner, 2008a). It could be a human instinct to avoid thoughts that lead to the experience of distress, but unfortunately this strategy appears to backfire, resulting in the resurgence of the very thoughts one is attempting to avoid.

The theory of ironic processes of mental control (Wegner, 1994) stipulates that thought suppression manipulates the associative pathways leading to and from a suppressed thought via two distinct processes (see association networks as per Collins and Loftus, 1975). First, a conscious, *effortful operating process* searches for thoughts that are not the suppressed thought, strengthening the activation of associative pathways leading away from the unwanted thought. Second, an unconscious, less effortful *ironic monitoring process* monitors for occurrences of the suppressed thought, in order to reactivate the conscious operating process, priming pathways leading toward the unwanted thought. In the classic “white bear” study (Wegner et al., 1987), participants were unsuccessful at prohibiting thoughts of a white bear from emerging into consciousness during five minutes of stream-of-consciousness reporting, with at least one white bear thought occurring each minute despite explicit instructions to not think about a white bear. When told to try and think about a white bear in a subsequent 5-minute verbalization period, the same participants reported more thoughts of a white bear than participants who hadn’t suppressed initially. Moreover, the frequency of white bear thoughts per minute increased over time in post-suppression periods, whereas white bear thoughts per minute decreased over time in all other conditions. Wegner and Smart (1997) proposed that suppression garners a state of deep cognitive activation whereby suppressed thoughts become increasingly activate, yet remain below the level of conscious awareness and only become conscious during brief intrusions. In contrast, expression of a thought increases its occurrence in conscious awareness without necessarily increasing its deeper activation (Wegner and Smart, 1997), and may actually lead to unpriming of the expressed thoughts (Sparrow and Wegner, 2006). Similarly, most studies show that target thoughts are reported infrequently during suppression and only increase in occurrence during subsequent expression periods, referred to as the rebound effect (e.g., Clark et al., 1991; Clark et al., 1993; Wegner et al., 1987; Wegner and Gold, 1995), although a few studies report that suppression increases the frequency of focus thoughts relative to a control condition of simply monitoring for occurrences of the thought (Lavy and Vandenhout, 1990; Salkovskis and Campbell, 1994). Unfortunately, most studies on thought suppression have relied on self-report measures that are susceptible to participant biases, for example in not reporting thought intrusions during suppression (Wenzlaff and Wegner, 2000). Those studies using dependent measures such as response time have observed a paradoxical increase in thought accessibility during suppression only when a cognitive load or distraction was added to the task (Najmi and Wegner, 2008b; Page et al., 2005; Wegner and Erber, 1992). In accord with the ironic process theory, the imposition of cognitive load would inhibit the conscious operating process, resulting in the increased accessibility of suppressed thoughts (e.g., Arndt et al., 1997; Macrae et al., 1994; Najmi and Wegner, 2008b; Page et al., 2005; Wegner and Erber, 1992). It remains unclear whether suppression leads to immediate increases in thought accessibility in the absence of an imposed cognitive load (Wenzlaff and Wegner, 2000). In the current study we used event-related potentials (ERPs), in particular the N400 component as an on-line measure of semantic processing, to test the hypothesis that an ironic monitoring process facilitates semantic access to suppressed thoughts.

Scalp-recorded ERPs are a powerful tool in language and memory research, offering precise temporal measurements of cognitive processes on the order of milliseconds. The N400, a widely distributed negative-going wave peaking around 400 ms after the onset of any meaningful word, picture, gesture or sound, has been a particularly reliable and useful dependent measure (e.g., Federmeier and Kutas, 2002; Holcomb and McPherson, 1994; Neville et al., 1997; Van Petten and Rheinfelder, 1995). Kutas and Hillyard (1980) observed that the amplitude of the N400 was larger for sentence endings that didn't match the prior context (e.g., "He took a sip from the *transmitter*.") than for those that did (e.g., "...*fountain*."). Since its initial discovery, the N400 has been used more generally as an index of on-line semantic processing (for a review, see Kutas and Federmeier, 2000). The amplitude of the N400 is graded based on the semantic fit and predictability of a stimulus within its prior context, such that the amplitude increases as the level of fit and predictability decreases (e.g., DeLong et al., 2005; Federmeier and Kutas, 1999; Federmeier, 2007; Kutas and Hillyard, 1984; Van Petten and Kutas, 1990; Wicha et al., 2004). When processing meaningful words in a list, the N400 typically has a more anterior maximum over frontocentral sites compared to words in a sentence. However, N400 amplitude shows a similar graded increase as in sentences as the degree of association between the target word and a previous word decreases (Bentin et al., 1985; Brown and Hagoort, 1993; Franklin et al., 2007; Hill et al., 2002; Holcomb, 1988; Holcomb, 1993; Nobre and McCarthy, 1994; Roehm et al., 2007; Rugg, 1985), resulting for example in small N400s for category exemplars preceded by a list of words within their category (Nunez-Pena and Honrubia-Serrano, 2005). Although the exact function of the N400 remains unclear, it appears to be modulated by the relatedness of a stimulus with its prior context, such that a closer semantic relationship between stimulus and context (less difficult integration) elicits reduced N400 amplitude. Overall, these results implicate the N400 in processes reflecting the semantic activation of a word, making it an ideal candidate for examining how thought suppression manipulates access to particular words.

In the current study we take advantage of the N400 component to test the idea that suppressed words are actually more active than normally processed words. We employ a paradigm similar to that developed by Najmi and Wegner (2008b), adapted to the constraints of the ERP methodology. In particular, we use a simple lexical decision task (LDT) preceded by a brief period of free-verbalization in which a word, e.g., *mountain*, was actively suppressed or expressed. During the LDT period, the target word *mountain* appeared along with unrelated real words and nonsense words. We measured the amplitude of the N400 for each word type under the suppression or expression conditions. Following the ironic monitor theory, if suppressed thoughts are more accessible than expressed thoughts, then a suppressed word should elicit a smaller N400 than that same word when it is being expressed. For example, when the word *mountain* is being suppressed, the spread of activation elicited by the visual presentation of *mountain* should be facilitated by the ironic process that monitors semantic access to *mountain*, presumably priming the associative pathways leading towards *mountain* (Najmi and Wegner, 2008b). Since the N400 is sensitive to automatic spread of activation (e.g., Hill et al., 2002), facilitated activation of semantic properties of *mountain* when suppressed should result in a decreased amplitude of the N400 relative to expression conditions. In expression periods following the initial suppression of *mountain*, where the rebound effect is typically observed and hypothetically elicits increased activation of the word, small N400 amplitude should also be observed relative to expression alone periods.

Alternatively, successful suppression of the focus word could lead to increased N400 amplitudes during suppression, according to an interpretation of the N400 as an index of semantic inhibition (Debrulle, 2007). When presented with a distracter-prime-target word sequence, N400 amplitudes evoked by distracters are larger when participants are instructed to ignore than attend to them while making a relatedness judgment about the prime and target. This effect is larger in participants who are quicker to make prime-target relatedness judgments

in critical trials where the distracter is related to the target (Debruille et al., 2008). The inhibition theory argues that the large N400 elicited by incongruous sentence endings reflects the inhibition of information related to the expected ending, rather than the difficulty of integrating the unexpected ending with the prior context (Debruille, 2007). Accordingly, evidence from several studies may support this idea showing relatively small N400s to unexpected words that are congruent with the sentence context, such as 'The cat that fled from the *mice*...' (e.g., Kolk et al., 2003; Kuperberg et al., 2003; Nieuwland and Van Berkum, 2005). Similar inhibition-related negativities have been observed in studies using adaptations of the Think/No-Think paradigm (Anderson and Green, 2001) to examine the suppression of memory retrieval (e.g., Bergstrom et al., 2007; Bergstrom et al., 2009; Mecklinger et al., 2009). In a typical Think/No-Think procedure, participants learn a list of cue-target word pairs, then later are presented with the cue and asked either to recall the target (Think) or to suppress the target from reaching awareness (No-Think). Traditionally, subsequent recall is impaired for No-Think targets relative to Think targets (e.g., Anderson and Green, 2001). Parallel ERP studies have shown the No-Think effect as a reduction in a late-onset parietal positivity believed to reflect inhibited recollection (Bergstrom et al., 2007; Bergstrom et al., 2009; Mecklinger et al., 2009). In addition, Bergstrom and colleagues (2009) observed a frontocentral negativity between 300 and 500 ms after the onset of word cues with a No-Think instruction, which was enhanced to items for which the associated target was later forgotten, similar to observations by Mecklinger and colleagues (2009) of a negativity elicited by words associated with a No-Think instruction.

To date, thought suppression has yet to be investigated with an ERP methodology, despite promising findings from on-line measures of thought suppression that have added much substance to a theory of ironic processes (Najmi and Wegner, 2008b; Page et al., 2005). ERPs provide a valuable measure of real-time cognitive processes without the necessity of a behavioral response. Their millisecond precision in the time domain can elucidate the rapid neural changes that occur in response to a stimulus. Moreover, by using reliable components, such as the N400, that have a fairly well established phenotype, we can test the specific predictions of the theories mentioned above. In order to build upon behavioral findings with more information about language-related processing, an explicit goal of this study was to establish a method for measuring thought suppression's ironic effects using ERPs. Analyses focused on the difference in focus word ERPs between expression and suppression blocks with an emphasis on the N400 time window, although other emerging effects were examined in detail, such as the P300— a positive-going wave that varies in latency but tends to reach a maximum by 300 ms post-stimulus and is thought to reflect categorical decision processes. We predicted that focus words would elicit smaller N400 amplitudes during suppression than expression, reflecting increased semantic accessibility as per the ironic monitoring process. Likewise, similar semantic facilitation effects were expected in expression blocks after suppression relative to those performed before suppression as a result of the post-suppression rebound. Alternately, larger N400 amplitudes for focus words during suppression blocks would reflect increased inhibitory processes.

2. Results

In order to establish a logical connection between performance on the thought suppression task and the ERPs obtained during the lexical decision task, the results were analyzed in a step-wise sequence. First, the number of mentions of the focus thought during the verbalization periods was analyzed as a function of mental control type (expression vs. suppression) and block order (first vs. second). Based on previous findings (e.g., Wegner et al., 1987), it was predicted that mentions of the focus thought would be most prevalent during expression periods occurring after suppression (expression block 2). Then, an initial inspection of the grand-average ERPs was performed as a manipulation check to ensure that any effects of expressing and suppressing focus thoughts were exclusive to the ERPs elicited by focus words. Since focus words were

exclusively affected by the mental control instruction, analyses were then directed at identifying specific differences in the ERPs across expression and suppression blocks.

2.1. Behavioral data

To assess performance during the verbalization periods, a 2 (Order) $\times$ 2 (Mental Control instruction) ANOVA was performed on the number of occurrences of the focus word, as indicated by the ringing of a bell or verbal expression of the word (see Table 1). There was a significant main effect of Order, $F(1, 19) = 35.976, p < .001$, as well as a significant main effect of Mental Control instruction, $F(1, 19) = 74.311, p < .001$. With regard to these effects, thoughts of the focus word occurred more often during the second of the two verbalization periods and more often during expression than suppression. However, these main effects were mitigated by a significant interaction between Order and Mental Control instruction, $F(1, 19) = 40.655, p < .001$, such that increases in occurrence of the focus thought during second blocks relative to first blocks were more pronounced in the expression condition. Similar to previous findings of focus thoughts being most prevalent in expression periods occurring after suppression, there appeared to be only a slight difference in mentions of the focus thought when suppressed in the first block ($M = 21.00, SD = 14.808$) or second block ($M = 22.65, SD = 15.083$), while the focus thought occurred much more often when expressed in the second block ($M = 63.50, SD = 24.595$) than the first block ($M = 49.10, SD = 20.501$). Post-hoc analyses confirmed that the difference in focus thought occurrence during first and second blocks under expression was significant ($p < .001$), while the difference in focus thought occurrence during suppression was not significant between first and second blocks ($p = .077$). These results are consistent with prior findings of a post-suppression rebound (for a review, see Abramowitz et al., 2001), confirming that participants were affected by mental control state. In addition, this confirms that our adaptation of the “white bear” paradigm (e.g., Wegner et al., 1987) to include more than one focus word over the experimental period was feasible for participants.

Accuracy of detecting non-word targets during the LDT was not significantly different between suppression ($M = .913, SD = 0.033$) and expression ($M = .908, SD = 0.034$) blocks, $t(19) = 0.758, p = .458$. The number of false alarms during the LDT was low (focus words $M = 94.38, SD = .06$; distracter words $M = 96.06, SD = .03$) and not significantly different between word type or mental control instruction in a 2 (Mental Control) $\times$ 2 (Word type) repeated-measures ANOVA (Word Type: $F(1, 19) = 1.871, p = .187$; Mental Control: $F(1, 19) = 1.832, p = .192$; Word Type $\times$ Mental Control: $F(1, 19) = 3.770, p = .067$). Since behavioral responses were restricted to non-words, reaction times to focus and distracter words were not obtained.

2.2. Event-Related Potentials

Visual inspection of the data (see Figure 1) reveals typical ERP components for all three word types, seen as the early N1-P2 complex, followed by a negative component starting at approximately 300 ms (N400), and a positive deflection around 450 to 600 ms (P300). An additional component was only observed in focus word trials during suppression blocks, seen in Figure 5 as the sustained prefrontal negativity from 300 to 600 ms (Late Anterior Negativity). As a manipulation check that mental control efforts were exclusive to focus words, we created difference waves of expression minus suppression brain activity (see Figure 2), which reveal that focus words elicited substantially different ERPs during expression and suppression periods. In contrast, the ERPs elicited by distracter words and non-words showed very little variance between mental control conditions. Manual responses were only required for non-word stimuli, therefore subsequent analyses were limited to focus and distracter words, since neither trial type was contaminated by muscle potentials related to the button-press. Mean amplitudes were analyzed during separate time windows post-stimulus onset: 50-150 ms to capture the N1, 260-440 ms to capture the N400, 300-600 ms to capture the Late Anterior Negativity, and 450-650 ms to capture the P300. These particular N400 and P300 windows

were chosen to model the time windows used in a previous study that employed a word list with a similar stimulus duration time and SOA (Nobre and McCarthy, 1994). For each time window, a repeated-measures ANOVA was performed with within-subject factors of Order (first vs. second block), Mental Control (suppression vs. expression), and Word Type (focus vs. distracter), and scalp distribution factors of Hemisphere (left vs. right), Laterality (lateral vs. medial), and Anteriority (prefrontal, frontal, central, occipital) as shown in Figure 6. Greenhouse-Geisser correction was applied to all repeated measures with more than one degree of freedom to correct for errors of sphericity. In the case of significant interactions, simple comparisons were performed with Bonferroni-adjusted degrees of freedom.

2.2.1 50-150 ms epoch (N1)—Figure 1 illustrates the frontocentral N1 elicited by both focus and distracter words (Laterality, $F(1, 19) = 26.943, p < .001$; Anteriority, $F(3, 57) = 35.477, p < .001$; Laterality $\times$ Anteriority, $F(3, 57) = 26.672, p < .001$), with a maximum at prefrontal sites in the right hemisphere (Hemisphere $\times$ Laterality $\times$ Anteriority, $F(3, 57) = 4.621, p = .009$). There was a significant main effect of Mental Control, $F(1, 19) = 5.020, p = .037$, revealing that N1 amplitudes were larger during expression than suppression blocks, predominantly at medial frontocentral electrodes (Mental Control $\times$ Laterality, $F(1, 19) = 9.699, p = .006$; Mental Control $\times$ Anteriority, $F(3, 57) = 5.179, p = .017$). However, this effect was mitigated by a Word Type $\times$ Mental Control interaction, $F(1, 19) = 4.896, p = .039$, such that N1 amplitudes were only larger during expression than suppression for focus words ($p = .019$) and not distracters ($p = .776$), most pronounced at medial electrodes (Mental Control $\times$ Word Type $\times$ Laterality, $F(1, 19) = 5.694, p = .028$). Moreover, a Mental Control $\times$ Order interaction, $F(1, 19) = 4.573, p = .046$, indicated that this difference was only significant for the first blocks ($p = .004$) but not the second ($p = .588$). A Mental Control $\times$ Order $\times$ Hemisphere interaction, $F(1, 19) = 7.955, p = .011$, demonstrates that this latter effect was largest in the right hemisphere ($p = .001$), although still significant in the left hemisphere ($p = .030$). In sum, larger N1 amplitude was observed only for expressed focus words in the first block over right anterior sites, relative to suppressed focus words.

2.2.2 260-440 ms epoch (N400)—As seen in Figure 1, focus and distracter words elicited a right-lateralized frontocentral negativity in the N400 window, as indexed by significant effects of scalp distribution factors (Hemisphere, $F(1, 19) = 5.536, p = .030$; Anteriority, $F(3, 57) = 23.353, p < .001$; Laterality, $F(1, 19) = 11.073, p = .004$; Hemisphere $\times$ Laterality, $F(1, 19) = 5.896, p = .025$; Laterality $\times$ Anteriority, $F(3, 57) = 3.231, p = .047$; Hemisphere $\times$ Laterality $\times$ Anteriority, $F(3, 57) = 5.403, p = .011$). Overall, distracter words elicited larger N400 amplitudes than focus words (Word Type, $F(1, 19) = 11.132, p = .003$), and this difference was largest at medial electrode sites (Word Type $\times$ Laterality, $F(1, 19) = 13.230, p = .002$). A three-way Word Type $\times$ Mental Control $\times$ Order interaction approached significance, $F(1, 19) = 4.262, p = .053$, suggesting that focus words elicited smaller N400 amplitudes during the first block of suppression relative to expression, with no significant difference between suppression and expression in the second block (see Figure 3). This effect reached significance at medial electrodes (Word Type $\times$ Mental Control $\times$ Order $\times$ Laterality, $F(1, 19) = 26.611, p < .001$), where amplitudes were smaller during suppression than expression for focus words in the first block ($p = .004$) and not the second block ($p = .396$), while neither difference was significant for distracter words (first block, $p = .770$; second block, $p = .222$). Closer examination of this interaction revealed that N400 amplitudes for focus and distracter words were only significantly different during suppression first blocks, where amplitudes were smaller for focus words ($p = .002$).

The ERPs shown in Figure 4 further elucidate the interaction between mental control instructions and block order that was observed for focus words, demonstrating a frontocentral N400 during expression first blocks that washed out by expression second blocks, and a frontally-maximal negativity during suppression first blocks that increased in size and

distribution by suppression second blocks. A Word Type $\times$ Mental Control $\times$ Order $\times$ Hemisphere $\times$ Anteriority interaction, $F(3, 57) = 4.775, p = .009$, revealed that N400 amplitudes for focus words increased significantly from suppression first to suppression second blocks at frontal electrodes in the left hemisphere ($p = .045$), whereas the decrease in N400 amplitudes for focus words from expression first to expression second blocks was non-significant. An exploratory analysis of this distributional pattern at individual electrodes found that N400 amplitudes were significantly larger during suppression second than suppression first blocks at the left dorsal and medial frontal sites (LDFr, $p = .018$; LMFr, $p = .030$), and were marginally smaller during expression second than expression first blocks at a right medial central site (RMCE, $p = .058$) and a midline parietal site (MiPa, $p = .062$). In sum, N400 amplitude was larger for expressed than suppressed focus words, only in the first block and at medial central electrodes. Additionally, N400 amplitude was larger for suppressed focus words that were previously expressed than when suppressed from the onset, exclusively at left-lateralized frontal electrodes.

2.2.3 300-600 ms epoch (Late Anterior Negativity)—As seen in Figure 5, a sustained anterior negativity was elicited by focus words during suppression blocks. A marginally significant Word Type $\times$ Mental Control $\times$ Order $\times$ Hemisphere $\times$ Anteriority interaction, $F(3, 57) = 2.912, p = .051$, indicates that this negativity, which was right lateralized and maximal at prefrontal electrodes, was larger during suppression than expression for focus words but not distracters. To further examine this trend, we analyzed mean amplitudes elicited by focus words using a 2 (Mental Control; suppression vs. expression) by 2 (Order; first vs. second block) repeated-measures ANOVA at each prefrontal electrode. There were significant effects of Mental Control at the midline prefrontal (MiPF), $F(1, 19) = 4.578, p = .046$, and right lateral prefrontal (RLPF) electrodes, $F(1, 19) = 4.563, p = .046$, such that focus word amplitudes were more negative during suppression than expression. This sustained negativity was larger during suppression than expression blocks regardless of block order, as indicated by the absence of significant Order or Mental Control by Order effects.

2.2.4 450-650 ms epoch (P300)—During the P300 time window, focus and distracter words elicited a left-lateralized positivity that was maximal at central-posterior electrodes (Laterality, $F(1, 19) = 9.627, p = .006$; Hemisphere, $F(1, 19) = 17.867, p < .001$; Anteriority, $F(3, 57) = 19.794, p < .001$; Hemisphere $\times$ Laterality, $F(1, 19) = 25.836, p < .001$; Hemisphere $\times$ Anteriority, $F(3, 57) = 13.689, p < .001$; Laterality $\times$ Anteriority, $F(3, 57) = 13.122, p < .001$; Hemisphere $\times$ Laterality $\times$ Anteriority, $F(3, 57) = 19.503, p < .001$). A main effect of Word Type, $F(1, 19) = 13.848, p < .001$, indicates that ERPs elicited during this time window were more positive-going for focus words than distracters, regardless of mental control instruction (Figure 1), and this effect was largest over posterior electrodes in the left hemisphere (Word Type $\times$ Hemisphere, $F(1, 19) = 6.560, p = .019$; Word Type $\times$ Laterality, $F(1, 19) = 10.213, p = .005$; Word Type $\times$ Anteriority, $F(3, 57) = 8.370, p = .002$; Word Type $\times$ Hemisphere $\times$ Laterality, $F(1, 19) = 5.258, p = .033$; Word Type $\times$ Hemisphere $\times$ Anteriority, $F(3, 57) = 5.170, p = .004$; Word Type $\times$ Laterality $\times$ Anteriority, $F(3, 57) = 8.233, p = .001$; Word Type $\times$ Hemisphere $\times$ Laterality $\times$ Anteriority, $F(3, 57) = 4.283, p = .011$). Main effects of, and interactions with, Order and Mental Control did not reach significance (F 's < 0.5), demonstrating that this component was only sensitive to word type.

3. Discussion

The present study tested claims that thought suppression facilitates semantic access to a suppressed word through an ironic monitoring process (Najmi and Wegner, 2008b; Wegner, 1994). During the initial block of the lexical decision task, we observed facilitated semantic accessibility for focus words at the N400 component during suppression relative to expression conditions. This was in contrast to self-reported thought occurrences obtained during the

preceding verbalization periods, where occurrences of the focus thought were more frequent during expression than suppression. In contrast to the semantic facilitation observed in initial suppression blocks, modulations in the N400 observed during subsequent blocks suggest that the semantic accessibility of focus words was inhibited during suppression after expression, possibly reflecting that thoughts become more manageable once thoroughly expressed.

As in the classic white bear study on thought suppression (Wegner et al., 1987), focus thought occurrences in this study, as indicated by bell push or overt mention of the focus word, were more frequent in expression than suppression conditions and most prevalent in expression periods during the second block (i.e., following suppression). Along with replicating the post-suppression rebound effect (for a review, see Abramowitz et al., 2001), these findings lend further credence to an interpretation of the observed ERP effects as being the result of the mental control manipulation, and validate the generalization of the ERP effects to previous thought suppression research. Immediately following each verbalization period, semantic accessibility was assessed by recording ERPs time-locked to the visual presentation of words and non-words during a lexical decision task. Instructions adopted from Najmi & Wegner (2008b) asked participants to maintain their mental control efforts from the verbalization period throughout the lexical decision task. Of particular interest was how expression and suppression would differentially modulate the N400 component, believed to be an index of automatic spreading activation (e.g., Hill et al., 2002). For the purposes of this experiment, the N400 was measured as the mean amplitude of brainwaves between 260 to 440 ms post-stimulus onset.

The ERP results reported here suggest that semantic processing of a target thought is indeed modulated by efforts to control its presence in consciousness. As seen in Figure 3, focus words elicited smaller N400s during suppression first blocks than expression first blocks, indicating that a word's semantic representation is more accessible when it is being suppressed than expressed. To our knowledge, this is the first direct evidence of the increased activation of a concept, as suggested by the ironic process theory (Wegner, 1994). Whereas previous studies have alluded to increased accessibility of suppressed thoughts via faster response times while performing a distracter task (e.g., Najmi and Wegner, 2008b; Page et al., 2005; Wegner and Erber, 1992), here we demonstrated the modulation of a brain-response related to semantic processing elicited solely by the visual presentation of the word being suppressed, based on average activity for ten different focus words. The dissociation between the N400 effect and the self-reported thought occurrences during the same blocks supports the view posited by Wegner and Smart (1997) that suppression increases the activation of a thought at a level below conscious awareness, despite being moderately successfully at keeping that thought from crossing the threshold into consciousness.

After initial mental control blocks, we observed a unique finding not reported in previous investigations of thought suppression: focus thoughts appeared to become less accessible during suppression periods after expression. Specifically, focus word N400 amplitudes were larger during suppression after expression than during initial suppression periods, although there was no significant difference between the two blocks of suppression in the number of self-reported thought occurrences (see Table 1). Since this increase in N400 amplitude observed from suppression first to suppression second blocks was largest at frontal sites in the left hemisphere (see Figure 4), similar to the frontocentral negativity observed in tasks measuring knowledge inhibition (Debruille, 1998; Debruille, 2007) and intentional forgetting (Bergstrom et al., 2009; Mecklinger et al., 2009), it seems that focus thoughts were easier to inhibit in post-expression periods. The opposite pattern was observed during expression, such that N400 amplitudes were larger in the first block than the second block (i.e., after suppression), demonstrating an increase in focus thought accessibility during post-suppression blocks that mirrored the rebound effect observed in the verbalization periods. Note that reductions in N400 amplitude due to increased thought accessibility might be difficult to

distinguish from reductions in N400 amplitude that occur with repetition of a stimulus (for a review, see Rugg and Curran, 2007). However, focus words elicited larger N400 amplitudes in the second block of suppression than when suppressed in the first block, opposite of what might be expected for repeated exposure of an item from the first to the second block. Alternately, this effect may reflect an increase in inhibition of semantic processing during second block suppression (Debrulle, 2007). Although novel, the finding of inhibited semantic access during suppression periods after expression is consistent with previous research on “unpriming” thoughts through expression (Sparrow and Wegner, 2006). It is possible that suppressing a thought becomes more feasible after the thought has been thoroughly expressed or unprimed, as opposed to when suppression is attempted from the outset.

Another explanation for the aforementioned negativity is that it reflects processes related to response inhibition, due to the task demands of withholding responses on focus and distracter word trials. Typically referred to as the No-Go N2, several ERP studies have observed increases in the amplitude of this component at frontocentral electrodes on No-Go trials where a manual response must be withheld, relative to Go trials where a response is executed (e.g., Gemba and Sasaki, 1989; Kok, 1986; Pfefferbaum et al., 1985; Sasaki et al., 1993). In this study, focus and distracter words were, in essence, No-Go trials, while non-words constituted the Go trials. Therefore, non-words would be expected to elicit a smaller negativity than focus or distracter words at frontocentral sites, where the No-Go N2 is maximally distributed. However, as seen in Figure 1, the negativity elicited by non-words appears to be just as large at frontal electrode sites as the negativity elicited by focus and distracter words. Note that the absence of a No-Go N2 effect in a simple task like lexical decision is not unusual (Chwilla et al., 1998; Nobre and McCarthy, 1994). Moreover, participants did not seem to experience difficulty inhibiting a motor response to the No-Go trials, as indicated by false alarm rates of less than 5% for both focus and distracter words. Thus, it seems that the frontocentral negativity reported here reflects semantic inhibition associated with the N400 component, rather than response inhibition, as indexed by a No-Go N2.

In addition to the predicted N400 effects, visual inspection of focus word ERPs revealed an unexpected, yet noteworthy sustained anterior negativity from 300 to 600 ms post-stimulus. Interestingly, this effect was exclusively observed during suppression blocks at two medial and right prefrontal electrodes (see Figure 5). Although localization of the neural sources for a scalp-recorded ERP should be interpreted with caution, this observation is consistent with fMRI findings of sustained activation in right-lateralized dorsolateral prefrontal cortex during thought suppression (Mitchell et al., 2007). According to models of thought suppression (Wegner, 1994; Wenzlaff and Wegner, 2000) and more general models of cognitive control (Matsumoto and Tanaka, 2004; Miller and Cohen, 2001), the prefrontal cortex is responsible for the conscious operating process involved in the maintenance of a desired mental state. The sustained negativity observed in this study may reflect the sustained activity associated with suppressing a thought, and might serve as a manipulation check that participants were indeed attempting to suppress their thoughts when instructed to do so.

Another unexpected effect emerged in the 50 to 150 ms time window, traditionally referenced as the N1 component (Hillyard et al., 1973; Luck, 2005). No predictions were made for this window in the current study since the N1 is believed to be an exogenous component that is relatively independent of higher-order cognitive processes (Coles and Rugg, 1995). Modulations in the N1 component are commonly targeted in studies examining discrimination between stimuli based on perceptual features (e.g., Vogel and Luck, 2000) or differences in attention to the stimulus (Hillyard et al., 1973), so the presence of significant differences between expression and suppression conditions described here suggests the top-down control involved in the mental control manipulation had some influence on early sensory processes. As seen in Figures 3 and 4, the N1 component was modulated by the mental control instructions,

and more interestingly this effect was exclusive to focus words and not distracters. The N1 is viewed as an index of processes related to orienting attention towards task-relevant stimuli, such that larger N1 amplitude is found for task-relevant than task-irrelevant stimuli (Luck et al., 1990). Accordingly, the large N1 amplitudes elicited by focus words during expression trials suggest that expressed words elicited increased attentional allocation, and therefore enhanced perceptual processing relative to suppressed words. However, attention allocation also contributes to the size of the P300, with more attention often leading to larger amplitude (for a review, see Polich, 2007). Although focus words elicited much smaller N1 amplitudes during suppression than expression, given that the P300 was significantly larger for focus than distracter words it would be hard to argue that attentional resources were not being allocated to focus words. Instead, perhaps attentional resources may have been allocated to efforts of decreasing, rather than increasing, the amount of perceptual processing of focus words when suppressing them. These hypotheses should be tested with further studies looking directly at the role of attention in focus word suppression.

Finally, during the window between 450 to 650 ms post-stimulus, we observed modulation of a positivity by word type (see Figure 1). Commonly referred to as the P300, this positive component is typically elicited by “oddball” stimuli that have a low probability of occurrence, as well as by stimuli from the task-relevant target category (e.g., Donchin and Coles, 1988). The P300 is thought to reflect categorical processing rather than response generation (Kutas et al., 1977), and its amplitude is larger for a target word that requires a categorical decision, such as the non-words in our lexical decision task. Chwilla and colleagues (1998) found that the P300 amplitude is minimized for words when participants are only required to push a button for non-words in a lexical decision task. In our data the largest positivity was indeed observed for the target non-words over posterior electrodes, consistent with prior research. However, we also observed larger P300 amplitude for focus words than distracters at centroparietal electrode sites. Since the centroparietal P300 is believed to reflect task-relevant processing (Polich, 2007), the larger amplitudes for focus words than distracters suggests that the mental control task forced the focus words to be treated as “targets” even though they were not relevant to the primary task of lexical decision. In addition, we observed a slightly different pattern over frontocentral electrodes, consistent with the subcomponent of the P300, known as the P3a (Polich, 2007). The P3a is more frontally-distributed, has been shown to be dependent on the novelty (or probability of occurrence) of a stimulus and elicited by infrequent non-targets across several modalities (e.g., Fabiani et al., 1998). In this study, focus words comprised only 10% of the trials within each block of the lexical decision task, whereas distracter words were presented on 40% of the trials, with non-words filling out the remaining half. As seen in Figure 1, the frontocentral distribution of the P3a elicited by focus words suggests that this component reflects novelty, in contrast to the task-relevant centroparietal P300. Together, these findings indicate, first that participants were indeed attending to the focus words, and second that focus words were processed as targets even though the contrast between them and distracter words was irrelevant to identifying non-words. Moreover, these effects illustrate one way in which ERPs are a valuable dependent measure, in that they can show effects to multiple stimuli types without a behavioral response. Importantly, there was no difference in the P300 amplitude elicited by focus words during suppression versus expression (see Figure 2). Relative to the observed differences between expression and suppression during the N1 and N400 time windows, the absence of significant differences between expression and suppression throughout the P300 time window suggests that focus words had an equivalent “surprise” value during expression and suppression blocks (Donchin, 1981).

Herein we presented a first attempt at studying thought suppression using a direct measure of brain activity to better understand its underlying cognitive processes. Designing an ERP paradigm to measure the effects of thought suppression was challenging. ERPs require sufficient trials to obtain reliable averages, yet are hampered by repetition effects, which can

in turn lead to an attenuation of N400 amplitude and increases in late positivity (Kim et al., 2001; Rugg, 1985; Van Strien et al., 2005). Most studies looking at thought suppression have used a single suppressed thought (e.g., white bear) and a single suppression exercise (e.g., Najmi and Wegner, 2008b; Wegner et al., 1987; Wegner and Erber, 1992). In our study, to allow for minimal repetitions of each suppressed word, yet a sufficient number of trials to obtain interpretable ERPs, each participant suppressed 10 different words throughout the course of the experiment. There was some question that suppressing more than one word over an experimental period in order to get sufficient data for ERPs would be challenging for a participant, although at least one study has shown that this was possible behaviorally (Page et al., 2005). Additionally, we aimed to make the processing elicited by focus words as implicit as possible by embedding them in a lexical decision task with instructions to only push a button upon the presentation of a non-word. As seen in the second blocks presented in Figure 3, focus words continued to elicit an N400-like component throughout the lexical decision task despite repeated presentation and exposure to each word during verbalization periods. This indicates that the N400 did not exhibit floor effects, similar to a recent study that found preserved N400 effects after 60 repetitions of a word (Debrulle and Renoult, 2009). An extension of this experiment could examine how thought accessibility is affected by distraction and other thought-regulation strategies, which have been shown to be effective at reducing the rebound effect (e.g., Wegner et al., 1987), yet remain to be examined during the act of distraction with an on-line measure. Possible applications include examining the efficacy of different thought regulation techniques and prescription medications at inhibiting the accessibility of problem thoughts in individuals with post-traumatic stress or obsessive-compulsive disorders.

Overall, these data lend support to the hypothesis that the act of suppressing words increases that word's activation at the semantic level, and show that inducing participants into a certain thought process can lead to modulation of brain responses that are believed to be predominantly automatic in nature. In this case, differences in the N400 component, a direct measure of brain activity related to semantic processing, between suppression and expression conditions provide a means to empirically measure the phenomenological experience of thought suppression's paradoxical effects. Perhaps most importantly, the dissociation between the self-reported thought occurrences and ERPs observed here supports the notion that thought suppression manipulations in conjunction with an implicit measurement can be utilized to measure the elusive unconscious mind (Wegner and Smart, 1997). In future studies we will be able to assess factors that influence this process, such as individual variability in the ability to imagine the suppressed thought and the role of attention in the maintenance of these thoughts during secondary tasks. These results add critical information about the nature of suppression's influence on the unconscious, and with thought suppression being considered as a potential causal factor in a broad spectrum of psychopathologies (Najmi and Wegner, 2008a), elucidating the underlying cognitive mechanisms of suppression may lead to breakthroughs that benefit a variety of clinical populations.

4. Experimental Procedure

4.1. Participants

Twenty college undergraduates at the University of Texas San Antonio (12 females, mean age = 19.33, range = 18-22) volunteered to participate in exchange for three hours of course credit. Data collected from two participants was excluded from analysis due to excessive muscle artifacts that persisted through the entire ERP recording session. All participants reported being right-handed, native speakers of English, with no history of neurological disorders, and normal or corrected-to-normal vision.

4.2. Stimuli

Ten focus words were chosen from the MRC Linguistic Database based on frequency ($M = 29.8$; range = 15-51), imageability ($M = 576.0$, range = 509-629), and concreteness ($M = 581.1$, range = 501-663), with the requirement that all words were not tools or animate objects (i.e., people or animals) to avoid the possibility that differences in brain activity between conditions were due to cognitive processing specific to either category (Shapiro and Caramazza, 2003). As an additional precaution against carryover effects between blocks, none of the focus words were associatively related according to the Edinburgh Associative Thesaurus (EAT). Based on these criteria, the focus words were *mountain, uniform, factory, journal, anchor, lamp, branch, pond, envelope, and dawn*. Each of the ten words was the focus of two separate blocks of the lexical decision task (LDT), once while participants were suppressing thoughts of the word and once while concentrating on thoughts of the word, counterbalanced for order across subjects. Forty unrelated English words were matched to the focus words on frequency, imageability, and concreteness, and did not have an associative relationship (according to the EAT) with the focus word with which they were presented or with each other. Fifty non-words were chosen from the ARC Non-word Database (Rastle et al., 2002), resulting in letter strings that followed spelling conventions in order to have a superficial appearance of being valid but were not legal words (e.g., MARPIS).

Each block of the LDT contained a single focus word, four words unrelated to that focus word, and five non-words. In order to extract an ERP with a good signal-to-noise ratio from the continuous EEG it was necessary to present repeated examples for each target word, as well as several different target words across LDT blocks. Therefore, each word and non-word was repeated four times within each block of LDT for a total of 40 trials per block (4 focus, 16 unrelated, 20 non-word), and exactly one half of the trials in each LDT block were non-words in order to set the probability of trials requiring a button press at 50%. There were a total of 80 critical trials across all LDT blocks in the experiment, comprised of 40 suppressed and 40 expressed focus words. Half of the LDT trials for each subject began with a suppression exercise followed by an expression exercise; half had the reverse order. The order of presentation was randomized for each focus word, as well as the order of presentation of the 10 focus words, across participants.

4.3. Procedure

After being fitted with EEG electrodes, participants were seated in the ERP recording chamber. Stimuli were presented in white, 72-point, Arial font on a black background. Each trial of the lexical decision task (LDT) consisted of a fixation cross “+” for 250 ms, the stimulus for 500 ms, followed by a blank screen for 750-1000 ms before the fixation point for the next trial (as per Nobre and McCarthy, 1994). The study took place in a sequence of randomized blocks consisting of a verbalization period and a LDT, with EEG activity recorded throughout each LDT. Before starting the experimental blocks, participants completed practice blocks for each of the tasks they would be performing throughout the session. For the LDT, participants were instructed that they would be presented with a series of letter strings on a computer screen, and their task was to respond only if the letter string was not a valid English word (e.g. MARPIS) by pushing a button with their right hand as quickly and accurately as possible. The button press responses were logged in a data file and analyzed for accuracy offline. After completing the practice LDT, participants completed a brief practice session on verbalizing their thoughts with the following instructions:

For the next five minutes, please verbalize your thoughts out loud. You are encouraged to report everything that comes to mind no matter how much sense it makes. You are not required to explain or justify why you are thinking the thoughts; we only ask that

you describe your thoughts as they come to you. The important thing is that you continue to talk throughout the five-minute period.

Following this practice verbalization period, participants were given an additional set of mental control instructions (Wegner et al., 1987) based on the focus word to which they were randomly assigned. For each focus word, participants were assigned to either suppression followed by expression, or vice-versa. As in previous research on thought suppression (e.g., Wegner et al., 1987), the order of suppression and expression blocks was randomized such that half were suppression then expression, and half were expression then suppression. At this point, participants were given a bell to ring and instructed to keep their hand over it during each verbalization period. Participants assigned to *suppress* were told the following:

For the next two minutes, please verbalize your thoughts as you did before, with one exception. This time, try not to think of the word “mountain”. Every time you do say “mountain” or have “mountain” come to mind, please ring the bell.

Participants in the *expression* condition were given the same instructions, but with specific orders to try to think of the word *mountain*.

Immediately after this second verbalization period, participants were instructed to *continue to try not to think (or try to think) of a mountain* while they performed a block of LDT. Each new LDT block consisted of new focus, non-focus (distracter), and non-word stimuli. Upon completion of 40 trials of LDT, participants were given the opposite instructions for the same focus word, such that those who were assigned to *suppress* were then given instructions to *express* for 2 minutes, and vice-versa:

For the next two minutes, please verbalize your thoughts as you did before, with one exception. This time, try to think (or not to think) of the word “mountain”. Every time you say “mountain” or have “mountain” come to mind, though, please ring the bell.

Immediately after this verbalization period, participants were instructed to continue with their new mental control instructions (*try to think or try not to think*) while they completed another 40 trials of LDT comprised of the same stimuli arranged in a different order. Each participant repeated this procedure until they completed a *suppression* and *expression* block (verbalization plus LDT) for all 10 focus words. All 10 focus words appeared once for a suppression block and once for an expression block for each subject in random order, such that across subjects every focus word appeared half of the time in suppression-expression and half in expression-suppression orders. Each block consisted of approximately 70 seconds of LDT and 2 minutes of verbalization. Considering the total of 20 blocks, including 30 minutes of EEG preparation, the experiment lasted approximately one hour and 45 minutes. Upon completion of all blocks, participants were debriefed and excused from the study.

4.4. ERP recording procedure

Continuous scalp-recorded EEG was acquired using a geodesic array of 26 pre-amplified sintered Ag-AgCl electrodes (BioSemi active electrodes) embedded in a custom electrode cap (Electro-Cap International Inc. using BioSemi electrode holders; see Figure 5). Additional electrodes were placed below and at the outer canthi of the left and right eyes to record blinks and eye movement respectively, and over the left and right mastoids to serve as offline reference. Preamplifiers in each electrode were used to reduce induced noise between the electrode and the amplification/digitization system (BioSemi ActiveTwo, BioSemi B.V., Amsterdam), allowing high electrode impedances. Electrode offsets were kept below 35 mV. A first-order analog anti-aliasing filter with a half-power cutoff at 3.6 kHz was applied (see www.biosemi.com). The data were sampled at 512 Hz (2048 Hz with a decimation factor of ¼) with a bandwidth of DC to 134 Hz, using a fifth order digital sinc filter. Each active electrode

was measured online with respect to a common mode sense (CMS) active electrode producing a monopolar (non-differential) channel, and was referenced offline to the average of the left and right mastoids. Data were processed using BrainVision Analyzer 2 (Brain Products GmbH, Munich). Non-causal Butterworth digital filters were applied with a low cutoff at 0.1 Hz (12 dB/oct) and high cutoff at 30.0 Hz (12 dB/oct). The EEG data were segmented in intervals of 1000 ms time-locked to stimulus-onset with a 200 ms pre-stimulus baseline. DC local detrend for 100 ms blocks was applied (Hennighausen et al., 1993), followed by baseline correction using -200 to 0 ms pre-stimulus. Epochs containing blinks, eye movement and excessive artifacts were removed from the data. Artifact rejection thresholds were adjusted for each participant for tests of maximum amplitude to capture blinks, maximum voltage step per ms to capture voltage spikes, minimum amplitude per 50 ms to capture flat lining and saccades, and maximum amplitude difference in 100 ms to capture signal drift. Average waveforms were then calculated for each condition time-locked to the events of interest.

Acknowledgments

This research was funded by grant SC1 HD060435 from NICHD/NIGMS and startup funds from the College of Science at the University of Texas at San Antonio (UTSA) to N. Wicha. We thank Drs. Ann Eisenberg, Rebekah Smith, and Reiko Graham for comments. We also thank the department of Psychology at UTSA for allowing us to use their participant pool, and all of the participants who made this study possible. This work was conducted by R. Giuliano as part of his Master's degree in Experimental Psychology.

References

- Abramowitz JS, Tolin DF, Street GP. Paradoxical effects of thought suppression: A meta-analysis of controlled studies. *Clinical Psychology Review* 2001;21:683–703. [PubMed: 11434226]
- Anderson MC, Green C. Suppressing unwanted memories by executive control. *Nature* 2001;410:366–369. [PubMed: 11268212]
- Arndt J, Greenberg J, Solomon S, Pyszczynski T, Simon L. Suppression, accessibility of death-related thoughts, and cultural worldview defense: Exploring the psychodynamics of terror management. *Journal of Personality and Social Psychology* 1997;73:5–18. [PubMed: 9216076]
- Bentin S, McCarthy G, Wood CC. Event-Related Potentials, Lexical Decision and Semantic Priming. *Electroencephalography and Clinical Neurophysiology* 1985;60:343–355. [PubMed: 2579801]
- Bergstrom ZM, Velmans M, de Fockert J, Richardson-Klavehn A. ERP evidence for successful voluntary avoidance of conscious recollection. *Brain Research* 2007;1151:119–133. [PubMed: 17428451]
- Bergstrom ZM, de Fockert JW, Richardson-Klavehn A. ERP and behavioural evidence for direct suppression of unwanted memories. *Neuroimage* 2009;48:726–37. [PubMed: 19563900]
- Brown C, Hagoort P. The Processing Nature of the N400 - Evidence from Masked Priming. *Journal of Cognitive Neuroscience* 1993;5:34–44.
- Chwilla DJ, Hagoort P, Brown CM. The mechanism underlying backward priming in a lexical decision task: Spreading activation versus semantic matching. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology* 1998;51:531–560.
- Clark DM, Ball S, Pape D. An Experimental Investigation of Thought Suppression. *Behaviour Research and Therapy* 1991;29:253–257. [PubMed: 1883305]
- Clark DM, Winton E, Thynn L. A Further Experimental Investigation of Thought Suppression. *Behaviour Research and Therapy* 1993;31:207–210. [PubMed: 8442747]
- Coles, MGH.; Rugg, MD. Event-related brain potentials: An introduction. In: Rugg, MD.; Coles, MGH., editors. *Electrophysiology of mind : event-related brain potentials and cognition*. Oxford University Press; Oxford ; New York: 1995. p. 1-26.
- Collins AM, Loftus EF. Spreading Activation Theory of Semantic Processing. *Psychological Review* 1975;82:407–428.
- Debrulle JB. Knowledge inhibition and N400: A study with words that look like common words. *Brain and Language* 1998;62:202–220. [PubMed: 9576822]

- Debruille JB. The N400 potential could index a semantic inhibition. *Brain Research Reviews* 2007;56:472–477. [PubMed: 18001839]
- Debruille JB, Ramirez D, Wolf Y, Schaefer A, Nguyen TV, Bacon BA, Renoult L, Brodeur M. Knowledge inhibition and N400: A within- and a between-subjects study with distractor words. *Brain Research* 2008;1187:167–183. [PubMed: 18035336]
- Debruille JB, Renoult L. Effects of semantic matching and of semantic category on reaction time and N400 that resist numerous repetitions. *Neuropsychologia* 2009;47:506–517. [PubMed: 18983862]
- DeLong KA, Urbach TP, Kutas M. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience* 2005;8:1117–1121.
- Donchin E. Surprise ... Surprise. *Psychophysiology* 1981;18:493–513. [PubMed: 7280146]
- Donchin E, Coles MGH. Is the P300 Component a Manifestation of Context Updating. *Behavioral and Brain Sciences* 1988;11:357–374.
- Fabiani M, Friedman D, Cheng JC. Individual differences in P3 scalp distribution in older adults, and their relationship to frontal lobe function. *Psychophysiology* 1998;35:698–708. [PubMed: 9844431]
- Federmeier KD, Kutas M. Right words and left words: electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research* 1999;8:373–392. [PubMed: 10556614]
- Federmeier KD, Kutas M. Picture the difference: electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia* 2002;40:730–747. [PubMed: 11900725]
- Federmeier KD. Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology* 2007;44:491–505. [PubMed: 17521377]
- Franklin MS, Dien J, Neely JH, Huber E, Waterson LD. Semantic priming modulates the N400, N300, and N400RP. *Clinical Neurophysiology* 2007;118:1053–1068. [PubMed: 17336145]
- Gemba DH, Sasaki K. Potential Related to No-Go Reaction of Go-No-Go Hand Movement Task with Color Discrimination in Human. *Neuroscience Letters* 1989;101:263–268. [PubMed: 2771172]
- Hennighausen E, Heil M, Rosler F. A correction method for DC drift artifacts. *Electroencephalogr Clin Neurophysiol* 1993;86:199–204. [PubMed: 7680996]
- Hill H, Strube M, Roesch-Ely D, Weisbrod M. Automatic vs. controlled processes in semantic priming - differentiation by event-related potentials. *International Journal of Psychophysiology* 2002;44:197–218. [PubMed: 12031295]
- Hillyard SA, Hink RF, Schwent VL, Picton TW. Electrical signs of selective attention in the human brain. *Science* 1973;182:177–80. [PubMed: 4730062]
- Holcomb PJ. Automatic and Attentional Processing - an Event-Related Brain Potential Analysis of Semantic Priming. *Brain and Language* 1988;35:66–85. [PubMed: 3179703]
- Holcomb PJ. Semantic Priming and Stimulus Degradation - Implications for the Role of the N400 in Language Processing. *Psychophysiology* 1993;30:47–61. [PubMed: 8416062]
- Holcomb PJ, Mcpherson WB. Event-Related Brain Potentials Reflect Semantic Priming in an Object Decision Task. *Brain and Cognition* 1994;24:259–276. [PubMed: 8185897]
- Kim MS, Kim JJ, Kwon JS. The effect of immediate and delayed word repetition on event-related potential in a continuous recognition task. *Cognitive Brain Research* 2001;11:387–396. [PubMed: 11339988]
- Kok A. Effects of degradation of visual stimulation on components of the event-related potential (ERP) in go/nogo reaction tasks. *Biol Psychol* 1986;23:21–38. [PubMed: 3790646]
- Kolk HHJ, Chwilla DJ, van Herten M, Oor PJW. Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language* 2003;85:1–36. [PubMed: 12681346]
- Kuperberg GR, Sitnikova T, Caplan D, Holcomb PJ. Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research* 2003;17:117–129. [PubMed: 12763198]
- Kutas M, McCarthy G, Donchin E. Augmenting Mental Chronometry - P300 as a Measure of Stimulus Evaluation Time. *Science* 1977;197:792–795. [PubMed: 887923]
- Kutas M, Hillyard SA. Reading Senseless Sentences - Brain Potentials Reflect Semantic Incongruity. *Science* 1980;207:203–205. [PubMed: 7350657]

- Kutas M, Hillyard SA. Brain Potentials during Reading Reflect Word Expectancy and Semantic Association. *Nature* 1984;307:161–163. [PubMed: 6690995]
- Kutas M, Federmeier KD. Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences* 2000;4:463–470. [PubMed: 11115760]
- Lavy EH, Vandenhoust MA. Thought suppression induces intrusions. *Behavioural Psychotherapy* 1990;18:251–258.
- Luck SJ, Heinze HJ, Mangun GR, Hillyard SA. Visual Event-Related Potentials Index Focused Attention within Bilateral Stimulus Arrays .2. Functional Dissociation of P1 and N1 Components. *Electroencephalography and Clinical Neurophysiology* 1990;75:528–542. [PubMed: 1693897]
- Luck, SJ. An introduction to the event-related potential technique. MIT Press; Cambridge, Mass: 2005.
- Macrae CN, Bodenhausen GV, Milne AB, Jetten J. Out of Mind but Back in Sight - Stereotypes on the Rebound. *Journal of Personality and Social Psychology* 1994;67:808–817.
- Matsumoto K, Tanaka K. Conflict and cognitive control. *Science* 2004;303:969–970. [PubMed: 14963319]
- Mecklinger A, Parra M, Waldhauser GT. ERP correlates of intentional forgetting. *Brain Res* 2009;1255:132–47. [PubMed: 19103178]
- Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience* 2001;24:167–202.
- Mitchell JP, Heatherton TF, Kelley WM, Wyland CL, Wegner DM, Macrae CN. Separating sustained from transient aspects of cognitive control during thought suppression. *Psychological Science* 2007;18:292–297. [PubMed: 17470250]
- Najmi, S.; Wegner, DM. Thought Suppression and Psychopathology. In: Eliot, AJ., editor. *Handbook of approach and avoidance motivation*. Psychology Press; New York: 2008a. p. 447-459.
- Najmi S, Wegner DM. The gravity of unwanted thoughts: Asymmetric priming effects in thought suppression. *Consciousness and Cognition* 2008b;17:114–124. [PubMed: 17339120]
- Neville HJ, Coffey SA, Lawson DS, Fischer A, Emmorey K, Bellugi U. Neural systems mediating American sign language: Effects of sensory experience and age of acquisition. *Brain and Language* 1997;57:285–308. [PubMed: 9126418]
- Nieuwland MS, Van Berkum JJA. Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary change deafness in discourse comprehension. *Journal of Cognitive Neuroscience* 2005:176–176.
- Nobre AC, McCarthy G. Language-Related Erps - Scalp Distributions and Modulation by Word Type and Semantic Priming. *Journal of Cognitive Neuroscience* 1994;6:233–255.
- Nunez-Pena MI, Honrubia-Serrano ML. N400 and category exemplar associative strength. *International Journal of Psychophysiology* 2005;56:45–54. [PubMed: 15725489]
- Page AC, Locke V, Trio M. An online measure of thought suppression. *Journal of Personality and Social Psychology* 2005;88:421–431. [PubMed: 15740437]
- Pfefferbaum A, Ford JM, Weller BJ, Kopell BS. ERPs to response production and inhibition. *Electroencephalogr Clin Neurophysiol* 1985;60:423–34. [PubMed: 2580694]
- Polich J. Updating p300: An integrative theory of P3a and P3b. *Clinical Neurophysiology* 2007;118:2128–2148. [PubMed: 17573239]
- Rachman S, de Silva P. Abnormal and Normal Obsessions. *Behaviour Research and Therapy* 1978;16:233–248. [PubMed: 718588]
- Rastle K, Harrington J, Coltheart M. 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology* 2002;55:1339–1362.
- Roehm D, Bornkessel-Schlesewsky I, Rosler F, Schlewsky M. To predict or not to predict: Influences of task and strategy on the processing of semantic relations. *Journal of Cognitive Neuroscience* 2007;19:1259–1274. [PubMed: 17651001]
- Rugg MD. The Effects of Semantic Priming and Word Repetition on Event-Related Potentials. *Psychophysiology* 1985;22:642–647. [PubMed: 4089090]
- Rugg MD, Curran T. Event-related potentials and recognition memory. *Trends in Cognitive Sciences* 2007;11:251–257. [PubMed: 17481940]

- Salkovskis PM, Campbell P. Thought Suppression Induces Intrusion in Naturally-Occurring Negative Intrusive Thoughts. *Behaviour Research and Therapy* 1994;32:1–8. [PubMed: 8135704]
- Sasaki K, Gemba H, Nambu A, Matsuzaki R. No-go activity in the frontal association cortex of human subjects. *Neurosci Res* 1993;18:249–52. [PubMed: 8127474]
- Shapiro K, Caramazza A. The representation of grammatical categories in the brain. *Trends Cogn Sci* 2003;7:201–206. [PubMed: 12757821]
- Sparrow B, Wegner DM. Unpriming: The deactivation of thoughts through expression. *Journal of Personality and Social Psychology* 2006;91:1009–1019. [PubMed: 17144761]
- Van Petten C, Kutas M. Interactions between sentence context and word frequency in event-related brain potentials. *Mem Cognit* 1990;18:380–93.
- Van Petten C, Rheinfelder H. Conceptual relationships between spoken words and environmental sounds: event-related brain potential measures. *Neuropsychologia* 1995;33:485–508. [PubMed: 7617157]
- Van Strien JW, Hagenbeek RE, Stam CJ, Rombouts SA, Barkhof F. Changes in brain electrical activity during extended continuous word recognition. *Neuroimage* 2005;26:952–9. [PubMed: 15955505]
- Vogel EK, Luck SJ. The visual N1 component as an index of a discrimination process. *Psychophysiology* 2000;37:190–203. [PubMed: 10731769]
- Wegner DM, Schneider DJ, Carter SR, White TL. Paradoxical Effects of Thought Suppression. *Journal of Personality and Social Psychology* 1987;53:5–13. [PubMed: 3612492]
- Wegner DM, Erber R. The Hyperaccessibility of Suppressed Thoughts. *Journal of Personality and Social Psychology* 1992;63:903–912.
- Wegner DM. Ironic Processes of Mental Control. *Psychological Review* 1994;101:34–52. [PubMed: 8121959]
- Wegner DM, Gold DB. Fanning Old Flames - Emotional and Cognitive Effects of Suppressing Thoughts of a Past Relationship. *Journal of Personality and Social Psychology* 1995;68:782–792. [PubMed: 7776182]
- Wegner DM, Smart L. Deep cognitive activation: A new approach to the unconscious. *Journal of Consulting and Clinical Psychology* 1997;65:984–995. [PubMed: 9420359]
- Wenzlaff EM, Wegner DM. Thought suppression. *Annual Review of Psychology* 2000;51:59–91.
- Wicha NY, Moreno EM, Kutas M. Anticipating words and their gender: an event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *J Cogn Neurosci* 2004;16:1272–88. [PubMed: 15453979]

Figure 1.

ERPs time-locked to onset of focus, distracter, and non- words, collapsed across mental control task and block order for the lexical decision task, shown at medial electrode sites. Time is on the x-axis, with zero marking the onset of the stimulus. Amplitude in microvolts is on the y-axis with negative values plotted up. Focus words elicited a frontocentral P3a component, whereas non-word targets elicited a typical centroparietal P300. The centroparietal P300 amplitude was significantly different between non-target stimuli - larger for focus than distracter words.

Effect of Mental Control (Expression - Suppression)

Figure 2. Voltage maps calculated by subtracting ERPs for expression minus suppression trials during the lexical decision task, collapsed across first and second blocks. The difference ERPs show an effect of mental control task for focus words but not distracters or non-words.

Effect of Mental Control Task by Order

Figure 3.

ERPs time-locked to focus word onset showing the effect of mental control task for first blocks (left columns) and second blocks (right columns). Notice the larger N400 amplitude for expression than suppression during first blocks that is absent in second blocks.

Effect of Order by Mental Control Task

Figure 4.

ERPs time-locked to focus word onset showing the effect of block order for expression (left columns) and suppression (right columns). Notice the decrease in N400 amplitude for expression and an inverse effect for suppression with an increase in N400 amplitude from first to second blocks.

Figure 5.

ERPs elicited by focus words during suppression and expression conditions, shown at prefrontal electrode sites. A sustained negativity can be seen during both suppression blocks at the midline and right lateral electrodes relative to expression blocks.

Figure 6.

Bird's eye view of the electrode montage (front of head at top) and regions used for statistical analysis (indicated by grey outlines).

Table 1

Means and Standard Deviations of Focus Thought Occurrence during Verbalization Periods Grouped by Mental Control Condition and Block Order

Condition	<i>M</i>	<i>SD</i>
Suppression		
1 st Blocks	21.00	14.81
2 nd Blocks	22.65	15.08
Expression		
1 st Blocks	49.10	20.50
2 nd Blocks	63.50	24.59

Note. Means derived from the total number of reported thought occurrences per block type.