A controlled comparison of the word repeating technique with a word association task


University at Albany, State University of New York, Department of Psychology SS399, 1400 Washington Ave., Albany, NY 12222, United States

1. Introduction

Acceptance and Commitment Therapy (ACT) is a treatment approach that aims to encourage acceptance of unwanted or distressful thoughts and feelings and movement toward valued goals (Hayes, Strosahl, & Wilson, 1999). A set of techniques within ACT are used for the purpose of promoting cognitive defusion, which is conceptualized as the disruption of the context required for language to have meaning (Blackledge, 2003). Cognitive defusion techniques are based on Relational Frame Theory (RFT; Hayes, Barnes-Holmes, & Roche, 2001), which posits that relationships between language, cognition, overt behavior, and emotions are both learned (i.e., through direct experience) and derived (i.e., learned without direct experience). These relationships form relational networks, which are webs of interconnected associations that can be activated by any stimulus that exists in that network. For instance, viewing a photograph of a deceased loved one activates the network containing information related to (1) the photograph itself (e.g., when and where it was taken and other aspects of photographs in general), (2) the individual in the photograph (e.g., his or her personality, events that were shared with the individual, and how he or she passed away), and (3) how to act when viewing photographs of this type (e.g., displaying a sad facial expression or crying).

Within RFT, psychopathology is understood as cognitive rigidity in relational networks and their interpretation, termed “thought-action fusion” (Hayes et al., 1999). Thought-action fusion occurs when language, spoken or thought, becomes synonymous with the actual thoughts and feelings to which it is related. If the thought “I am alone because my loved one died” comes to elicit the emotions of sadness and despair, this thought and the related emotions have been “fused.” This fusion may result in distress both from the thoughts themselves and attempts to avoid or suppress the thoughts.

Rather than attempting to rationalize, challenge, change, or decrease the frequency of such thoughts, as is often the goal in other cognitive behavioral therapies, ACT employs cognitive defusion techniques to deliteralize language. That is, the meanings of words are divorced from the words themselves so that a word is just a word and a thought is simply a thought, nothing more. If the thought “I am alone because my loved one died” is viewed simply

* Portions of this study were presented at the 42nd annual meeting of the Association for Behavioral and Cognitive Therapies in Orlando, FL.
* Corresponding author. Tel.: +1 518 437 4446; fax: +1 518 442 4867.
E-mail address: kd8977738@albany.edu (K.P. De Young).
as a combination of words outside of their literal context (i.e., the relations they entail), it loses its potency to evoke related cognitions and emotions. Thus, defusion techniques aim to help individuals become uninvolved observers of mental processes that would otherwise lead them to feel distressed. Within the RFT conceptualization of these techniques, they do not actually change the relationships between stimuli; they encourage individuals to view the relationships between stimuli as unnecessary, thereby changing their function.

One specific method used to achieve cognitive defusion is the word repeating technique (WRT; Hayes et al., 1999). In this exercise, a word or short phrase is quickly repeated aloud until the context required for the word to have literal meaning has changed. Individuals are told to focus on the auditory qualities of the word and the sensations of the word being produced by their mouths in order to demonstrate that words are just a collection of sounds and need not elicit the meanings to which they are related.

Although researchers have only recently studied word repetition as an ACT technique, it has been a topic of scientific inquiry for nearly a century. Basset and Warne published a study on the loss of word meaning with verbal repetition in 1919, and Severance and Washburn published a study on the loss of word meaning upon prolonged visual fixation in 1907. Originally termed semantic (or verbal) satiation, this effect received considerable attention during the 1950s and 1960s with reviews on the topic published in 1964 and 1971 (Amster, 1964; Esposito & Pelton, 1971). Esposito and Pelton's review focused on methods used to study satiation (i.e., the extent to which a word has lost its meaning) and found that much of the research on the subject had major methodological and conceptual flaws that limited conclusions regarding whether any measured changes occurring after word repetition were due to loss of meaning or were the result of other phenomena prompted by repetition. Although this technique has more recently been integrated into the therapeutic repertoire of ACT, an increasingly popular treatment, there persists a relative lack of experimental evidence demonstrating that the effects of the WRT are the result of words having lost their meanings.

The first study to test the effects of the WRT within the ACT framework used an alternating treatment design with a sample of eight female undergraduates (Masuda, Hayes, Sackett, & Twohig, 2004). Masuda et al. found that the degree to which participants believed negative self-referential words described them and the extent to which these words had a negative emotional impact decreased after they repeated these words for 30 s (i.e., performed the WRT) and were lower than ratings following an alternative procedure and a control task. Importantly, although techniques in the alternative procedure were varied (e.g., breathing training and positive self-talk) and the authors acknowledged that it was unknown what strategies participants employed, the goal of these techniques was thought suppression (“the purpose was not to think of the negative thought,” p. 480), which has been found to have detrimental rather than inert or beneficial effects (Wegner, Schneider, Carter, & White, 1987). The authors interpreted these findings as being consistent with RFT; however, their data could not address the process by which the observed effects occurred.

In another study, Masuda et al. (2009) replicated and extended their previous findings using a parametric design in which a non-clinical sample repeated negative self-referential words for 1, 3, 10, 20, or 30 s. They found that word repetition reduced how much participants believed the negative self-referential words and how uncomfortable the words made them feel after a short period as 1 s of repetition. The strength of the effects increased with the duration of the repetition.

Most recently, Masuda et al. (2010) conducted a controlled comparison of the WRT and a distraction task in a non-clinical sample and found that performing the WRT resulted in lower ratings of the degree to which negative self-referential words made individuals feel uncomfortable and the extent to which individuals believed the words described them compared to both the distraction task and the control task. In addition, the authors found that levels of depression did not moderate post-task ratings. The use of a theoretically opposing comparison task (distraction) was a substantial improvement upon prior research, but it remains unclear whether the effects of the WRT are in fact achieved via the loss of word meaning or if another process common to other techniques is responsible.

In the present study, the effects of the WRT on the degree to which individuals believed negative self-referential words accurately described them and the degree to which these words had a negative emotional impact were compared to the effects of another procedure that specifically requires that words keep their meaning in order to elucidate possible explanations of the mechanism through which the WRT exerts its effects. The comparison procedure selected was the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). A computerized word-sorting task that was developed to measure implicit attitudes, preferences, and self-concepts by comparing the latency of word-sorting when different word categories are paired with one another. The theory behind the IAT suggests that words are more quickly sorted into categories when categories that are closely associated with one another are paired together (e.g., good and flower), compared to when they are paired with an antonym of the closely associated category (e.g., bad and flower). Studies have documented that the strength of word associations measured by the IAT weaken over repeated administrations (Nosek, Greenwald, & Banaji, 2007), suggesting that practice at sorting words affects associative strength and perhaps indicating that the relationships between the words change. The IAT was chosen as the comparison task in this study because (1) it requires one to process the literal content of words in order for them to be correctly sorted (i.e., words must keep their meanings), and (2) it could conceivably result in an outcome similar to that of the WRT but through a theoretically different mechanism that is more similar to that which is posited to occur in other therapeutic techniques, such as cognitive restructuring (a technique commonly employed in cognitive behavioral therapies to change the way an individual thinks). If the WRT results in decreases in the extent to which individuals believe negative self-referential words are true of them and in their negative emotional impact via the removal of words from their literal context, one would not expect to observe similar changes following a task that specifically requires that words keep their meanings. However, if both procedures result in similar outcomes, they may instead alter the context of the words by changing their associations with other evaluative or self-referential words rather than by removing their meanings. Gaining knowledge of the mechanisms through which therapeutic techniques act is invaluable for the development and implementation of theoretically consistent treatments.

In addition, because all of the participants in the studies conducted by Masuda et al. (2004, 2009, 2010) received a detailed rationale for engaging in the WRT, it is unknown to what degree suggestibility may have influenced their results. It has been proposed that the rationale for a treatment can be as potent as, or even more potent, than the treatment itself (Fish, 1973). Moreover, Esposito and Pelton (1971) noted that the effects of word repetition were enhanced when participants were provided with a rationale. The present research therefore sought to test the effect of providing individuals with a believable rationale in addition to comparing the WRT and IAT procedures.
Finally, the studies by Masuda et al. (2004, 2009, 2010) evaluated the effect of the WRT on negative self-referential thoughts that were initially rated toward the extreme end of the rating scales. In their most recent studies (Masuda et al., 2009, 2010), participants were excluded if they did not provide initial ratings for negative self-referential thoughts that were sufficiently extreme. Although choosing particularly extreme stimuli for the task is necessary to allow for change in the hypothesized direction and is theoretically important for testing therapeutic techniques in analog samples, this procedure sets the stage for the statistical phenomenon of regression to the mean. It is thus necessary to use an appropriate control condition to prevent over-interpreting change that may be better accounted for by other factors, such as habituation to the response items and demand characteristics that contribute to regression to the mean (c.f. Smith, 2005). The present study employed a control condition for this reason.

2. Method

2.1. Participants

A total of 200 undergraduates participated for credit toward a course requirement. Mean (SD) age of participants was 19.3 (3.6) years. Women comprised 68% of the sample, and the race/ethnicity of participants was as follows: 64.5% Caucasian, 15.5% African-American, 11.5% Latino(a), 5.0% Asian-American, and 3.5% other/missing. Participants were required to be at least 18 years old and native English speakers.

2.2. Materials

2.2.1. Negative word list (NWL)

The NWL was developed for this study and contains 43 negative, potentially self-referential words. These words were generated through an exhaustive search of multiple thesauri for words that individuals could apply to themselves in a self-degradatory fashion. Each word was rated on two dimensions using 7-point scales: believability (i.e., to what degree individuals believed the word to be true of them), from “not at all believable” to “very believable,” and discomfort (i.e., to what degree the word made them feel uncomfortable), from “not at all uncomfortable” to “very uncomfortable.” In a control condition (n = 40), the NWL demonstrated adequate immediate test-retest reliability (from pre- to post-test, \( r = .94, p < .05 \) and \( r = .66, p < .05 \) for discomfort and believability ratings, respectively) and 2-week test-retest reliability (from post-test to follow-up, \( r = .79, p < .05 \) and \( r = .85, p < .05 \) for discomfort and believability ratings, respectively). Based on the control group’s ratings at pre-test, Cronbach’s alpha indicated excellent internal consistency for discomfort ratings (\( \alpha = .97 \)) and believability ratings (\( \alpha = .95 \)). The six words rated most uncomfortable and believable were the following: clumsy, flawed, insecure, crazy, failure, and fat. The full list of all 43 words is available from the first author.

2.2.2. Implicit association test (IAT)

The IAT measures the latency of word-sorting. The relative strength of the association between two pairs of words (e.g., “good and bad” and “flower and insect”) is measured by the time it takes to sort words into these categories when the categories are paired with one another. As an example, “good” and “flower” would appear in the upper left corner of a computer monitor, while the words “bad” and “insect” would be in the upper right corner. Words that can be sorted unambiguously into one of these four categories (e.g., daisy, spider, wonderful, and awful) appear one at a time in the center of the screen. The participant’s task is to press a key with one hand if the word belongs in the “good” or “flower” category and to press a different key with his or her other hand if the word belongs in the “bad” or “insect” category. The pairings are then switched so that one key must be pressed to sort words into the “good” and “insect” categories and another key must be pressed to sort words into the “bad” and “flower” categories. The relative valence of an individual’s attitude toward flowers versus insects is derived by subtracting the latencies of the first pairing from the second.

In the present investigation, the IAT was individually tailored for each participant using four words from the NWL. The two most believable and two most uncomfortable words as rated by each participant were used as the target words for the “bad” category so that an opportunity existed for the ratings of these words to demonstrate change from pre- to post-test. The remaining non-target words were used to test whether any effects of the task generalized beyond the words targeted in the task. The “good” category was comprised of the following words for every participant: wonderful, terrific, awesome, and noble. The self-referential categories were “self” and “other” and were comprised of the words: me, my, mine, and I, and their, them, they, and it, respectively. Participants first completed a practice round with “good” and “bad” and “insect” and “flower” to orient them to the task. Next, they completed four rounds of “self” was “base” and “good” and “base” in which each of these pairings were completed twice. A total of 240 trials (i.e., four blocks of 60 target word presentations) were completed by each participant with the experimenter in the room.

2.2.3. Word repeating technique (WRT)

The WRT is described in detail by Hayes et al. (1999), and their example with the word “milk” was used as the protocol for the WRT conditions in this study. Four words from the NWL list were chosen as target words for this task. Again, the two most uncomfortable and two most believable words as rated by each participant were used as the target words for that participant. Participants began by saying the word “milk” repeatedly for a period of 30 s to orient them to the task. Participants then repeated the target words one at a time for a period of 30 s each. Finally, participants repeated each of the four “good” words used in the IAT conditions. This was done to balance both the exposure to “good” versus “bad” words and the duration of the tasks between conditions. Fifteen seconds into each word repeating trial, the experimenter instructed the participant to repeat the word faster and louder. After 30 s, participants were instructed to stop.

2.3. Design and procedure

Participants were randomized to one of five conditions represented by a 2 × 2 design of word task (IAT or WRT) by rationale (rationale given or not) and a fifth control condition. This resulted in a total of 40 participants per condition. All conditions were equal in terms of their duration (about 5 min), exposure to an experimenter (except for the control condition in which there was less face-to-face interaction), and amount of repeated exposure to word stimuli.

All participants first completed the NWL. They then received verbal instructions regarding how to complete their respective task and, if in a rationale condition, were provided with a theoretically consistent rationale. In the WRT with rationale condition (WRT-R), the description closely resembled the script from the Hayes et al. (1999) “milk” example. For the IAT with rationale (IAT-R) condition, the description of the task was based upon an associational explanation of how words can be interrelated, stressing how repeatedly pairing words together can result in word cueing when certain words come to mind and that pairings can be strengthened or weakened with practice. Both rationales discussed the role...
words and language play in human suffering and related the purpose of the task to alleviating suffering. Control participants were asked to read a brief article about hydrogen fuel cell technology followed by completing five multiple-choice questions about the content of the article. Immediately after completing the task, all participants completed the NWL a second time.

2.4. Data analyses

To examine whether statistically significant changes in believability and discomfort of target and non-target words occurred, repeated measures analyses of variance (ANOVAs) were conducted with the 160 participants in the four experimental conditions using a 2 (IAT vs. WRT) by 2 (rationale vs. no rationale) by 2 (pre-test vs. post-test) analysis. Statistical significance was set at \( p < .05 \).

Finally, experimental groups were compared to the control group. Although the control group was not exposed to target words, the two pairs of words they rated most believable and most uncomfortable (i.e., the words that would have been used as target words, had control participants been in an experimental condition) were used to analyze for regression to the mean. Since the most extremely-rated words were used in both the IAT and WRT procedures, regression to the mean could account for decreases that may have occurred from pre- to post-test. To test whether pre- to post-test changes exceeded the magnitude of regression to the mean, separate 2 (experimental condition vs. control condition) by 2 (pre-test vs. post-test) ANOVAs were conducted comparing ratings of target and non-target discomfort and believability ratings of the control group to each experimental condition. Significant group by time interactions would indicate that the effect in the experimental group exceeded that observed in the control group (i.e., exceeded what would be expected by simple regression to the mean). In order to control for family-wise error rate in these analyses, a Bonferroni correction was employed, setting statistical significance at \( p < .003 \).

3. Results

3.1. Pre- versus post-test word ratings

Means, standard deviations, and effect sizes for pre-test to post-test within group change on word discomfort and believability ratings appear in Table 1. For target word discomfort ratings (Fig. 1A), analyses revealed a significant main effect of time, such that post-test ratings were lower than pre-test ratings \( (F_{1,156} = 55.51, p < .05) \). Pairwise comparisons indicated that both tasks resulted in significantly lower post-test discomfort ratings for target words \[ IAT (F_{1,156} = 12.61, p < .05); WRT (F_{1,156} = 48.79, p < .05) \]. There was no main effect of task \( (F_{1,156} = .29, p = .59) \); however, there was a significant time by task interaction \( (F_{1,156} = 5.90, p < .05) \) indicating that individuals who engaged in the WRT task demonstrated a larger decrease than individuals who engaged in the IAT task. There was no main effect of rationale \( (F_{1,156} = .26, p = .61) \), and no time by rationale \( (F_{1,156} = .30, p = .58) \) or task by rationale \( (F_{1,156} = .26, p = .61) \) interactions.

There was no main effect of time for non-target word believability ratings (Fig. 1D; \( F_{1,156} = 1.43, p = .09 \)). There was also no main effect of task \( (F_{1,156} = .60, p = .44) \) or rationale \( (F_{1,156} = .12, p = .73) \). In addition, the time by task \( (F_{1,156} = .56, p = .46) \), time by rationale \( (F_{1,156} = .12, p = .73) \), and task by rationale \( (F_{1,156} = .28, p = .60) \) interactions were all non-significant.

3.2. Experimental versus control condition: Testing for regression to the mean

None of the changes from pre- to post-test in any of the four experimental conditions exceeded the magnitude of regression to the mean observed in the control group (all \( p's > .003 \)). However, since the Bonferroni procedure is a conservative correction for family-wise error, comparisons were examined at the conventional significance level of \( p < .05 \) as well. At this significance level, only one pre- to post-test change exceeded the magnitude of regression to the mean observed in the control group. For non-target discomfort ratings, changes from pre- to post-test in the WRT with rationale condition exceeded changes in the control condition \( (F_{1,78} = 4.40, p = .04; n^2 = .05) \), demonstrating a small effect.

4. Discussion

The goals of the present research were twofold: (1) to compare the effects of the WRT and the IAT on the extent to which negative self-referential words are believed to be true and have a negative emotional impact, and (2) to examine the importance of providing participants with a rationale for the respective tasks. The IAT was selected as a comparison procedure because it could conceivably result in an outcome comparable to the WRT but through a theoretically different mechanism. While the WRT is hypothesized to function via loss of word meaning, correctly sorting words in the IAT procedure requires one to process the literal content of words, so any effects could not be due to loss of target word meaning.

Concerning the second goal, the effectiveness of the tasks were not found to vary based on whether or not individuals were provided task-specific rationales. Several possibilities exist that might have accounted for this finding. First, the rationales may not have been convincing enough. Both the rationale for the WRT and the IAT were constructed using the Hayes et al. (1999) "milk" example as a template, which has been used in previous research but does not guarantee effective rationales. Given that no data were collected regarding how much participants believed the rationale that was provided to them, it remains unclear whether this factor was responsible. Second, the participants may not have found the rationales to be engaging. The experimenters read aloud the rationale to participants, and portions of the rationales were designed to be interactive, including practicing the task, in order to
engage participants as much as possible. Nevertheless, participants may not have been particularly motivated to fully attend to the study procedures. Finally, the rationales provided may not have been convincing enough to override participants’ experience during the task. Even if participants accepted the task rationale and were convinced of its effectiveness prior to performing it, they may have later discounted these beliefs when completing post-test ratings if their experience during the tasks conflicted with the information provided to them by the rationales.

Regarding the first goal of this study, both techniques were found to result in lower discomfort and believability ratings from pre-test to post-test for target words, with the WRT demonstrating a larger decrease than the IAT in discomfort rating but showing no difference in believability ratings. The WRT also resulted in lower post-test discomfort ratings for non-target words, whereas the IAT did not. Neither task resulted in lower post-test believability ratings for non-target words. If the reductions in the discomfort and believability ratings were the result of loss of word meanings, one would expect to observe these results only for individuals who participated in the WRT task, since target words must have retained their meanings in the IAT conditions.

The differences observed between target and non-target words are intriguing, and may indicate that the effects of the WRT tend to generalize beyond the experimental paradigm more readily than the effects of the IAT, at least in the context of their effects on discomfort. The lack of a generalization effect for the IAT is consistent with its hypothesized mechanism being one of associative strength between the specific words used in the task. It is likely that these relationships must be directly manipulated in order to change. This finding is also consistent with the hypothesized mechanism of the WRT, which is didactic and experiential and is believed to generalize to other stimuli with which the individual has not had direct experience.

These findings should be interpreted in light of the results of the regression to the mean analyses. Contrary to recent findings (Masuda et al., 2010), the WRT procedure did not outperform the control task in post-test target word discomfort or believability ratings. The results of these analyses indicate that neither the WRT nor the IAT demonstrated changes beyond what could be accounted for by statistical regression. Importantly, the failure to demonstrate change beyond statistical regression was not due to a lack of statistical power. Analyses conducted with G*Power

### Table 1
Means and standard deviations for discomfort and believability by condition.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Discomfort</th>
<th>Believability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre (M)</td>
<td>Post (M)</td>
</tr>
<tr>
<td></td>
<td>(SD) (95% CI)</td>
<td>(SD) (95% CI)</td>
</tr>
<tr>
<td>Target</td>
<td>4.02(1.2)</td>
<td>3.58(1.5)</td>
</tr>
<tr>
<td></td>
<td>0.32</td>
<td>0.31</td>
</tr>
<tr>
<td>Non-target</td>
<td>4.21(1.1)</td>
<td>3.71(1.7)</td>
</tr>
<tr>
<td></td>
<td>0.35</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Note: IAT-R = Implicit Associations Test with rationale; IAT = Implicit Associations Test without rationale; WRT-R = Word Repeating Task with rationale; WRT = Word Repeating Task without rationale; M = Mean; SD = Standard Deviation; d = Cohen’s d effect size within group pre-test to post-test change; 95% CI = 95% confidence interval.

Fig. 1. Mean Discomfort and Believability Ratings for Target and Non-Target Words. Note: (A) Target discomfort ratings; (B) Target believability ratings; (C) Non-target discomfort ratings; (D) Non-target believability ratings; IAT-R = Implicit Associations Test with rationale; IAT = Implicit Associations Test without rationale; WRT-R = Word Repeating Task with rationale; WRT = Word Repeating Task without rationale; Error bars indicate standard error of the mean.
(Version 3.0; Erdfelder, Lang, & Buchner, 2007) indicated that a sensitivity to detect moderately small effects (i.e., Cohen’s $f^2 = .10$) existed in these analyses at 80% power.

Notably, a potentially important difference between this study and the studies by Masuda et al. involves the use of the NWL in this study. Since participants in this study rated all words (i.e., both target and non-target) immediately following the task, on average, more time would have elapsed between the end of the task and when the target words were rated, depending upon where in the list the target words appeared, than elapsed in the studies by Masuda et al. It is therefore possible that the effects of the task diminished in those additional seconds that were present in this study but not in the studies by Masuda and colleagues. However, if this is the case, the longer-term therapeutic impact of this technique is questionable.

4.1. Clinical implications

Although this research represents a step toward better understanding the mechanisms through which the WRT functions, these results cannot speak to the clinical utility of the WRT due to the use of a non-clinical sample. Indeed, the purpose of the present investigation was not to provide evidence for the clinical utility of these tasks but rather to determine whether the WRT and IAT tasks, which are believed to operate via theoretically distinct mechanisms, produce similar outcomes in terms of the believability and discomfort of negative self-referential words. In other words, this investigation was not concerned with whether the WRT exerts an effect, which previous research (Masuda et al., 2004, 2009, 2010) has already demonstrated. Instead, this study aimed to shed light on possible explanations for this effect, which may have broad clinical implications. For instance, the purpose of the WRT in ACT is to deliteralize language to reduce suffering, and it is thought that this occurs without changing the relationships between stimuli in a relational network. Furthermore, RFT posits that relief from suffering occurs in the absence of change in these networks; what changes is the individual’s belief in their necessity. If instead this technique works by altering the relationships between stimuli in a relational network, then the mechanism is more consistent with how change is viewed in other popular cognitive behavioral techniques (c.f. Hofmann & Asmundson, 2008). The results of this study have important implications for research on this particular cognitive defusion technique, as well as the verification of the mechanisms through which other therapeutic techniques function. The IAT and WRT tasks resulted in similar decreases in the discomfort and believability of negative self-referential words. This is contrary to what one would expect given the proposed mechanism of the WRT. Since similar results were found for the two tasks, three possible conclusions are that (1) the IAT and WRT tasks function via the same mechanism, (2) there are two distinct mechanisms that produce the same effect, or (3) they do not produce any effect, and thus there is no mechanism. Given that the proposed mechanisms underlying the WRT and IAT tasks are in theoretical opposition (i.e., the IAT requires processing literal word content while the WRT is believed to function via loss of meaning), the likelihood of such similar outcomes resulting from two distinct mechanisms is unlikely. This conclusion is based on Occam’s razor, which states that the most parsimonious explanation tends to be the best explanation. If the tasks indeed produce an effect, it is possible that they do so through a common mechanism, such as exposure. Repeatedly experiencing words that previously evoked strongly negative reactions may extinguish this response over time, whereas avoiding experiencing these words and their associated thoughts may maintain their negative valence through negative reinforcement. Nevertheless, the most parsimonious explanation of the present findings is that the techniques did not exert any effects and were no different from what should have been expected due to regression to the mean.

To the authors’ knowledge, the current investigation is only the fourth to test the effects of the WRT on the believability and negative emotional impact of words. The results of this study are partially consistent with those of Masuda et al. (2004, 2009, 2010), who found the WRT to result in decreases in the believability and negative emotional impact of words in their non-clinical samples. Although the finding that individuals in the present study who participated in the WRT demonstrated significant reductions in the believability and negative emotional impact of negative self-referential words was similar to Masuda et al’s results, the carefully controlled design of the current investigation allowed for an analysis of regression to the mean. This analysis revealed that the decreases observed were no different from those observed in a control group that read a brief neutral article and answered five multiple-choice questions about the article. Although the mechanism through which the WRT exerts its effects remains unclear, it may exert influence through loss of meaning, weakened associations, habituation, or statistical artifact.

4.2. Limitations

The results of this study should be interpreted in light of certain limitations. First, although the suggestion that the IAT and WRT may act via the same mechanism (whether the mechanism is real or akin to statistical artifact) is the most parsimonious explanation for these effects, any conclusions about the actual mechanism(s) are speculative. The only way to support or disconfirm conclusions about their mechanism(s) is to devise a method to measure whether, or to what degree, a word has or does not have meaning. This may prove exceptionally difficult, as suggested by Esposito and Pelton (1971). Second, this study used a non-clinical undergraduate sample. It is possible that the WRT and IAT tasks would produce different outcomes in clinical populations who exhibit higher levels of thought-action fusion and/or stronger negative self-referential word associations. However, the recent results obtained by Masuda et al. (2010), in which the effects of the WRT were not moderated by levels of depression, suggest that this may be unlikely. Third, participants engaged in the WRT and IAT procedures for a relatively brief period of time (i.e., in total about 5 min). Thus, it is possible that participants did not receive a large enough “dose” of either procedure to induce changes of a magnitude detectable by the analyses used in this study. However, given that Masuda et al. (2009) found effects of word repetition after as little as 1 s, the 30 s duration of each trial in this study was likely sufficient.

Despite these limitations, this study has important implications for the status of cognitive defusion and the therapies that employ it (e.g., ACT). If the WRT does indeed work by a different mechanism that the one put forth by proponents, then the theory underlying its use (i.e., RFT) may need to be modified.

References