THE NATURE, DEVELOPMENT, AND REVERSAL OF POSITIVE ATTITUDES TOWARD NONSUICIDAL SELF-INJURY: IMPLICATIONS FOR PREDICTION AND TREATMENT

Joseph C. Franklin, M.A.

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology (Clinical Psychology).

Chapel Hill
2013

Approved by:

Mitchell J. Prinstein, Ph.D.

Anna Bardone-Cone, Ph.D.

B. Keith Payne, Ph.D.

Gabriel S. Dichter, Ph.D.

Jonathan S. Abramowitz, Ph.D.

ABSTRACT

JOSEPH C. FRANKLIN: The Nature, Development, and Reversal of Positive Attitudes toward Nonsuicidal Self-Injury: Implications for Prediction and Treatment
(Under the direction of Mitchell J. Prinstein, PhD)

Nonsuicidal self-injury (NSSI; e.g., cutting or burning the skin without suicidal intent) is a dangerous and increasingly prevalent health-risk behavior. Despite advances in NSSI research over the past decade, many aspects of NSSI remain poorly understood. In particular, there are few strong predictors of NSSI, it is unclear how positive attitudes toward NSSI develop, and there are no empirically supported treatments for NSSI. In the present study, I addressed these topics with a multi-method, experimental, and longitudinal approach. For Aim 1 of the study, I examined baseline differences between NSSI (n = 58) and control (n = 86) adult participants on NSSI-themed versions of five measures that cover different aspects of attitudes: the implicit association test (IAT); the affect misattribution procedure (AMP); explicit affective ratings; startle eyeblink reactivity; and startle postauricular reactivity. Compared to the control group, the NSSI group displayed significantly more positive attitudes on all five measures. Moreover, AMP scores and explicit ratings prospectively predicted self-cutting frequency over the ensuing six months. For Aim 2, I employed pain offset relief conditioning in an attempt to induce more positive implicit attitudes toward NSSI in the control group. This conditioning significantly diminished startle eyeblink reactivity in the context of NSSI images, but did not significantly affect any other measures. For Aim 3, I tested the ability of aversive conditioning in the NSSI group to reverse positive implicit attitudes toward NSSI and to reduce NSSI behaviors over the subsequent six months. Aversive conditioning normalized startle eyeblink and postauricular reactivity, but did not significantly affect any other measures. Results also provided preliminary support for the hypothesis that aversive conditioning prospectively reduces self-cutting. In conjunction with my other recent studies (Franklin et al., 2010; 2011, 2012, 2013), these findings have prompted a new theoretical framework called the Benefits and Barriers model of NSSI.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF ABBREVIATIONS

APA	American Psychiatric Association
AMP	Affect Misattribution Procedure
DERS	Difficulties in Emotion Regulation Scale
DSM	Diagnostic and Statistical Manual
EMG	Electromyographic
ERS	Emotion Reactivity Scale
IAT	Implicit Association Test
IAPS	International Affective Picture System
MINI	Mini International Neuropsychiatric Interview
NSSI	Nonsuicidal Self-Injury
SITBI	Self-Injurious Thoughts and Behaviors Interview

THE NATURE, DEVELOPMENT, AND REVERSAL OF POSITIVE ATTITUDES TOWARD NONSUICIDAL SELF-INJURY: IMPLICATIONS FOR PREDICTION AND TREATMENT

Nonsuicidal Self-Injury

Nonsuicidal self-injury (NSSI) is self-inflicted tissue damage that is direct, intentional, socially unacceptable, and without suicidal intent (e.g., cutting or burning the skin; Nock, 2010). In contrast to repetitive self-injury in autism spectrum disorders (e.g., repetitive head-banging) and severe self-injury in psychotic disorders (e.g., limb amputation), NSSI is non-repetitive and moderate in intensity. NSSI is surprisingly common, with 1 to 5% of adults, 17% of college students, and 13 to 23% of adolescents reporting that they have engaged in NSSI at least once in their lives (Gratz, 2001; Jacobson & Gould, 2007; Klonsky, 2011). Rates in clinical samples are even higher, with some studies finding that 40% - 61% of adolescent psychiatric inpatients engage in NSSI (Darche, 1990; DiClemente, Ponton, & Harley, 1991). In addition to being a health risk behavior itself (DiClemente et al., 1991), NSSI also may increase the risk of suicidal self-injury (Andover & Gibb, 2010; Joiner, 2005). In fact, recent studies have found that a history of NSSI was a stronger predictor of future suicide attempts than a history of suicide attempts (Asarnow et al., 2011; Guan, Fox, & Prinstein, 2012).

Although large gains have been made in the understanding of NSSI over the last decade, many aspects of this behavior remain poorly understood. Three of the most critical gaps in knowledge about NSSI are (1) how to longitudinally predict NSSI, (2) how positive attitudes toward NSSI develop, and (3) how to effectively treat NSSI.

NSSI prediction. There are very few longitudinal studies of NSSI. Consequently, very little is known about predictors of NSSI beyond general factors (e.g., stress) and specific but relatively weak factors (e.g., history of NSSI). Two recent investigations found that depressive symptoms and stress longitudinally predict NSSI (Guerry & Prinstein, 2010; Hankin & Abela, 2011). Similar to the prediction of suicide attempts, such factors may have high *sensitivity* for predicting NSSI, but they may not have high *specificity* for predicting NSSI (cf. Nock et al., 2010). In other words, such predictors may be limited because nearly everyone who engages in NSSI may have elevated depressive symptoms and stress, but most people with elevated depressive symptoms and stress will never engage in NSSI. Similarly, Glenn and Klonsky (2012) recently found that the frequency of prior NSSI longitudinally predicted future NSSI, but this prediction was relatively weak and the majority of NSSI behaviors in this sample were not severe (e.g., hair-pulling, scratching, biting). Given the limitations of these traditional approaches for the prediction of self-injury, a new approach may be necessary to provide a substantial improvement in the ability to predict future NSSI.

The development of positive attitudes toward NSSI. Knowledge about how positive attitudes toward NSSI develop would provide crucial insights into one of the most perplexing aspects of self-injury: how individuals overcome the natural instinct to avoid stimuli associated with pain and injury (e.g., blood, wounds) so thoroughly that they seek out these stimuli. Such knowledge may establish important new targets for NSSI prevention and treatment efforts; however, there are major obstacles to studying this question. Because people may lack insight into why they engage in NSSI (Janis & Nock, 2009; cf. Nisbett &

Wilson, 1977), it is difficult to examine the development of positive attitudes toward NSSI with self-report methods. Similarly, it is difficult to investigate this question with traditional laboratory methods because such investigations may not be ethically feasible. Such obstacles limit the potential of traditional approaches to yield new knowledge about the development of positive attitudes toward NSSI. A new approach may be necessary to advance knowledge on this topic.

NSSI Treatment. There are currently no empirically supported treatments for NSSI (Nock, 2010). Studies involving behavior therapy, dialectical behavior theory, cognitive therapy, and psychodynamic therapy have all been unable to show significant reductions in NSSI in the experimental group compared to the control group (Levy et al., 2007; Linehan et al., 2006; Lynch & Cozza, 2009; Newman, 2009; Rathus & Miller, 2002; Tyrer et al., 2003). Wood et al. (2001) found that adolescents in a group therapy condition were less likely to engage in NSSI than adolescents assigned to a treatment usual condition. In a replication study, however, Hazell et al. (2009) reported that adolescents assigned to this group therapy condition were *more* likely to engage in NSSI. Given the prevalence and dangerousness of NSSI, it is concerning that there are no empirically supported treatments for this behavior. Even more troubling, NSSI appears to be resistant to all of the major forms of therapy that have been shown to be efficacious for many other forms of psychopathology. Within the realm of traditional therapeutic approaches, few options remain for designing potentially efficacious interventions for NSSI. This suggests the need for a new treatment approach for NSSI.

A New Approach: Attitudes and Conditioning

Attitude research may provide an effective new avenue for studying the prediction, development, and treatment of NSSI. This section will provide a brief overview of attitude research and its potential relevance to NSSI.

Attitude has been defined as “the psychological tendency to evaluate a given entity with some degree of favor or disfavor” (Gawronski & Bodenhausen, 2006, p. 693). There are two major types of attitudes: implicit and explicit. Based on the model proposed by Gawronski and Bodenhausen (2006), implicit attitudes are governed by associative processes. Associative processes are activated automatically when a given stimulus is encountered. They are independent of formal reasoning; an individual may even consider their implicit attitudes to be inaccurate. For example, repeated pairings of a shock with a kitten may eventually lead to a person to develop a negative implicit attitude towards all kittens. This may occur even if this person self-reports that kittens are harmless and pleasant. This latter type of evaluation is called an explicit attitude. Explicit attitudes are based on propositional processes, which are concerned with assessing the truth value of beliefs (Gawronski & Bodenhausen, 2006). In our example, this person may reason that kittens pose no obvious dangers and are regarded by most people as pleasant. Accordingly, this person may have opposing implicit and explicit attitudes toward kittens.

In many instances, implicit attitudes may explain and predict behavior more effectively than explicit attitudes. Continuing the above example, despite having a positive explicit attitude toward kittens, this person may avoid kittens because of a negative implicit attitude toward kittens. In this case, the implicit attitude would more accurately predict behavior than the explicit attitude. Implicit measures are particularly effective for investigating attitudes for which people lack insight or are motivated to conceal. For example, there is a high correspondence between implicit and explicit measures of racial bias for people who are not particularly motivated to appear unprejudiced; in contrast, there is a low correspondence for people who are motivated to appear unprejudiced (e.g., Payne, Cheng, Govorun, & Stewart, 2005). Similarly, Payne, Govorun, and Arbuckle (2008) found that social pressure to under-report drinking alcoholic beverages significantly affected self-report measures of drinking, but not implicit attitudes toward drinking.

People are often motivated to conceal their self-injurious behaviors, and the majority of self-injurers may lack insight into why they are self-inflicting pain or whether they will do it again (Janis & Nock, 2009; Nock, 2010; Nock et al., 2010). Accordingly, implicit measures may be particularly valuable for studying self-injury. Nock et al. (2010) provided a dramatic demonstration of this point. Results showed that implicit associations with suicide were the best predictors of future suicide attempts, far outpacing well-known predictors such as previous attempts, suicidal ideation, depressive symptoms, patient prediction, and clinician prediction. Although similar longitudinal studies have not been conducted on NSSI, initial evidence indicates that implicit associations powerfully discriminate between NSSI and non-NSSI participants (Nock & Banaji, 2007). These studies suggest that implicit attitudes may play an important role in NSSI (Nock, 2010), with implicit measures having the potential to provide unprecedented insight into these behaviors. In particular, information about the development and reversal of positive implicit attitudes toward NSSI may provide a unique window into the development and potential reversal of NSSI. This new approach may generate progress where traditional approaches have found obstacles or stagnation.

In the present study, I employed this new approach to investigate three major aims: (1) the ability of various components of attitudes to discriminate between people with and without a history of NSSI at baseline and to longitudinally predict NSSI; (2) the development of positive attitudes toward NSSI; and (3) the efficacy with which aversive conditioning reverses positive attitudes toward NSSI and reduces future NSSI behaviors. I discuss each of these aims in detail below.

Using Multiple Attitude Components to Longitudinally Predict NSSI

Five attitude measures were utilized in the present study. Three of these measures are traditionally categorized as attitude measures (explicit ratings, affect misattribution procedure [AMP], and implicit association test [IAT]) and the other two are traditionally categorized as psychophysiological measures of motivational direction (startle eyeblink and postauricular reactivity). However, consistent with the views of many social psychologists who employ psychophysiological methods (e.g., Ito & Cacioppo, 2007), I posit that these latter measures are also effective for examining attitudes. Below, I delineate these five measures and discuss how each may provide insight into attitudes toward NSSI.

Measure 1: Explicit affect. As noted above, explicit attitudes arise from propositional reasoning and are typically measured with self-report instruments (cf. Gawronski & Bodenhausen, 2006). In the present study, I measured explicit affective attitudes toward NSSI with a computerized survey that asked participants to rate the valence of NSSI stimuli.

Implicit attitudes. The remaining four measures gauged different components of implicit attitudes. Implicit associations and implicit affect are perhaps the most commonly studied components of implicit attitudes. There is little overlap between these components (e.g., Payne et al., 2008), meaning that simultaneously examining these components maybe provide greater discrimination between NSSI and non-NSSI groups, and may improve the ability to predict future NSSI. Compared to implicit association research, implicit affect research is newer and less common. Indeed, NSSI research has only involved implicit associations (Nock & Banaji, 2007). Interestingly, initial research suggests that implicit affect may more effectively explain behaviors strongly associated with affect regulation, such as drinking alcohol (e.g., Payne et al., 2008). Given that people who engage in NSSI may overcome powerful natural instincts to avoid mutilation stimuli (e.g., blood, wounds), affect may be a particularly important aspect of implicit attitudes towards NSSI.

Measure 2: Implicit associations. Implicit associations are most often measured with a brief behavioral test called the implicit association test (IAT; Greenwald, McGhee, & Schwartz, 1998). The IAT compares the reaction time for classifying one set of semantic stimuli (e.g., "death" and "me") to the reaction time for classifying another set of semantic stimuli (e.g., "life" and "me"). This comparison is quantified as a difference score. Greater difference scores are thought to reflect stronger implicit associations for one set of semantic stimuli compared to another. Implicit associations have been shown to predict future behavior, especially in socially sensitive domains (e.g., prejudice; see Greenwald, Poehlman, Uhlmann, & Banaji, 2008). As noted above, Nock et al. (2010) demonstrated that a death/suicide version of the IAT is currently the best known predictor of future suicide attempts.

Measure 3: Implicit affect – performance-based. In social psychology research, implicit affect is often measured with a brief behavioral test called the Affect Misattribution Procedure (AMP; Payne et al., 2005). The AMP involves presenting a picture (often emotionally evocative) for 75 ms, followed by a blank screen for 125 ms, then a Chinese pictograph for 100 ms, and finally a gray screen that remains until the participant presses a button. Participants are instructed to ignore the emotional picture and to judge whether the Chinese pictograph seems to be more pleasant or less pleasant than the average symbol. Once the gray screen appears, participants press one button for "more pleasant" or another button for "less pleasant." Even though participants are told to ignore the emotional image, the evaluation of the Chinese pictograph is strongly influenced by the nature of the emotion image that preceded the Chinese pictograph. Specifically, the Chinese pictograph is more likely to be judged as pleasant if the picture that preceded it was pleasant (Payne et al., 2005). This is believed to occur because participants misattribute the affect inspired by the emotional picture as being inspired by the Chinese pictograph (see Loersch & Payne, 2011). The AMP has been employed to examine such things as addictive behaviors (Payne, McClernon, & Dobbins, 2007; Payne et al., 2008), prejudice (Imhoff & Rainer, 2009), morals (Hofmann & Baumert, 2010), and voting behavior (Payne et al., 2005, 2010).

Measures 3 and 4: Implicit affect – psychophysiological. Based on constructionist views of emotion (e.g., Barrett, 2012), psychophysiological measures of affect may provide a more direct measure of “core affect” (i.e., a state of pleasure or displeasure combined with arousal and grounded in the body; see Barrett & Bliss-Moreau, 2009) than performance-based measures such as the AMP. This is because psychophysiological measures directly assess the physiological states that constitute core affect. In contrast, performance-based measures include psychological factors that go beyond these physiological states, such as the mental processes involved in making the decision to press a certain key during the AMP. This suggests that both types of measures may index slightly different components of implicit affect, with psychophysiological measures indexing specific components of core affect (e.g., defensive motivation) and performance-based measures gauging overall core affect combined with various psychological factors (e.g., memory, decision-making).

In psychophysiological research, affective-valence startle modulation is often employed to measure motivational aspects of implicit affect. Startle eyeblink is a defensive reflex that occurs in response to a sufficiently intense and sudden stimulus, such as a sudden and loud sound (Blumenthal et al., 2005). Because it is a defensive reflex, startle eyeblink reactivity is heightened by unpleasant stimuli (particularly threatening stimuli; Vaidyanathan et al., 2009) and reduced by pleasant stimuli (see Lang, Bradley, & Cuthbert, 1990). Eyeblink reactivity is quantified as the EMG activity of the orbicularis oculi muscle, which surrounds the eye and contracts to produce an eyeblink (see Blumenthal et al., 2005). Startle eyeblink reactivity is one of the few psychophysiological measures specific to defensive motivation (Lang et al., 1990). In other words, startle eyeblink reactivity tends to increase with the degree to which physiological resources are gathered in response to a perceived threat (cf. Lang, 2010). In the affective-valence startle modulation paradigm, startle eyeblink reactivity reliably increases with the degree to which picture stimuli are perceived as threatening (e.g., Bradley, Cuthbert, & Lang, 1999; Lang, Bradley, & Cuthbert, 1990).

The postauricular reflex is often measured simultaneously with startle eyeblink. The postauricular muscle is located behind the ear and is used by most infant mammals to pull back the ear during nursing; however, it is largely vestigial in humans (Johnson, Valle-Inclán, Geary, & Hackley, 2012). Although it has no apparent functional use in humans, the postauricular muscle can still be primed by pleasant stimuli (especially those related to food, nursing, or reward; see Sandt et al., 2009; Johnson et al., 2012). When an intense stimulus like a startling sound is presented, it synchronizes the motor units of the postauricular muscle, generating a sudden spike in activity that is larger the more primed the muscle is (i.e., during pleasant affective states; see Johnson et al., 2012). Accordingly, postauricular reactivity shows a pattern of affective modulation that is opposite to that of startle eyeblink reactivity (Benning et al., 2004). Startle postauricular reactivity is one of the few psychophysiological measures specific to appetitive motivation (Gable & Harmon-Jones, 2009). In other words, postauricular reactivity tends to reflect the degree to which physiological resources are marshaled to approach a perceived source of nurturance. Interestingly, despite being measured at the same time to the same stimuli and displaying inverse patterns, multiple studies have shown that startle eyeblink and postauricular reactivity display near-zero correlations (e.g., $r = -.07$ in Franklin et al., in press; $r = -.12$ in Sandt et al., 2009). This highlights that value of including multiple measures of affect.

The first aim of the present study was (a) to investigate how attitudes toward NSSI stimuli discriminate between NSSI and Control groups at baseline, and (b) to examine how attitudes toward NSSI prospectively predict NSSI six months later. This portion of the study has the potential to advance knowledge about the nature of attitudes toward NSSI and how to predict future NSSI.

The Development of Positive Attitudes toward NSSI

Although preliminary evidence indicates that people with a history of NSSI display more positive implicit attitudes toward NSSI stimuli (Nock & Banaji, 2007), it is unclear how these attitudes might develop. This knowledge would provide insight into on how people overcome the instinct to avoid mutilation stimuli and how these counter-instinctual attitudes can be prevented or reversed.

Attitude formation and change. Much research has been aimed at examining how attitudes can be formed and altered. This research has usually employed conditioning paradigms such as concurrently pairing pictures (e.g., a neutral object) with other pictures (e.g., an angry face). These studies have demonstrated that such conditioning can decrease racial prejudice (Olson & Fazio, 2006), increase body satisfaction (Martijn, Vanderlinden, Roefs, Huijding, & Jansen, 2010), improve self-esteem (Dijksterhuis et al., 2004), and alter food preferences (Hollands, Prestwich, & Marteau, 2011), among many other things (for a review, see De Houwer, Thomas, & Baeyens, 2001; Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010). Interestingly, paradigms that employ electrocutaneous stimulation (i.e., shocks) as the unconditioned stimuli appear to be the most effective at altering attitudes (see Hofmann et al., 2010). Positive attitudes toward NSSI may develop due to many different naturalistic forms of conditioning. For example, several movies depict admirable characters and attractive celebrities engaging in NSSI (Whitlock, Eckenrode, & Silverman, 2006). In effect, such movies pair pleasant stimuli with NSSI, which should condition positive (or less negative) implicit attitudes toward NSSI. Such social mechanisms may influence attitudes for a subset of people who engage in NSSI, but another conditioning pathway may influence implicit attitudes for nearly everyone who engages in NSSI: pain offset relief conditioning.

Pain offset relief conditioning. Pain has been defined as the unpleasant sensory and emotional experience associated with actual or potential tissue damage (International Association for the Study of Pain Task Force on Taxonomy, 1994). In sharp contrast to the negative affect associated with pain, recent research has shown that pain *offset* leads to emotional relief. That is, the removal of a painful stimulus does not return emotions to a "neutral" state; rather, it seems to generate a pleasant emotional state. People who engage in NSSI may unintentionally (or intentionally in some cases) access this pain offset relief phenomenon in order to reduce stress. As discussed below, stimuli present during NSSI-related pain offset relief (e.g., blood, knives) may gradually gain a more positive valence. This "pain offset relief conditioning" may accordingly represent one pathway through which many individuals acquire positive attitudes toward NSSI.

This phenomenon of pain offset relief is consistent with the opponent process theory of acquired motivation (Solomon, 1980). The opponent process theory proposes that in response to every stimulus there are two major reactions: a primary process and an opponent process (see Figure 1). The primary process reaches maximum intensity soon after stimulus onset and diminishes completely soon after the stimulus offset. Conversely, the opponent process reaches maximum intensity slowly and lingers long after the offset of the stimulus. At any given time, the state of an organism is determined by taking the absolute value of the difference between the intensity of the primary process and the opponent process (see Figure 1). As such, the primary process is most salient immediately after the onset of a stimulus whereas the opponent process is most salient after the offset of a stimulus. For NSSI, the primary processes are pain and fear, and the opponent process is relief. Accordingly, relief should be most salient soon after pain offset during NSSI.

Several laboratory studies have directly demonstrated pain offset relief. For example, Leknes, Brooks, Wiech, and Tracey (2008) found that pain offset diminished physiological arousal and increased self-reported pleasantness in healthy controls. For both NSSI and non-NSSI groups, Franklin et al. (2010) showed that pain offset diminished defensive motivation as assessed by startle eyeblink reactivity. In a follow-up study, Franklin et al (in press) specified that pain offset simultaneously stimulated appetitive motivation (postauricular reactivity) and diminished defensive motivation (startle eyeblink reactivity). In a second follow-up study, Franklin et al. (2013) found a similar pattern in individuals with a history of NSSI and clarified that the degree of pain offset relief was not correlated with NSSI frequency. Taken together, these findings indicate that pain offset relief is a natural mechanism that powerfully regulates affect regardless of self-injury history.

One potential problem with the pain offset relief conditioning hypothesis is that pain may not completely offset during an episode of NSSI. For example, wounds from cutting or burning may hurt for days after the episode. Interestingly, basic pain studies suggest that complete pain offset may not be necessary to generate relief. Grill and Coghill (2002) found that a one degree increase in the intensity of a painful heat stimulus (e.g., from 49 to 50 degrees) led to a one unit increase in perceived pain intensity (on a 1 – 10 scale). Strikingly, a one degree decrease (e.g., from 50 to 49 degrees) led to a 3.5 unit decrease in perceived pain intensity. The authors concluded that pain intensity reduction activates powerful analgesic mechanisms that serve to "amplify awareness of stimulus offset and to reinforce escape behaviors" (p. 2205). Such findings suggest that an individual may induce several instances of pain offset relief (even if the offset is not complete) during a single NSSI session. For example, an individual may cut themselves several times over the course of one minute and feel intense pain throughout that minute. Yet, because of small variations in pain intensity during that minute (e.g., variations in how the knife is cutting through the skin), the individual may experience numerous instances of relief.

Within NSSI episodes, this relief may become associated with stimuli present during pain offset. Consistent with the attitude conditioning research reviewed above, the liking of these surrounding stimuli may be increased due to their association with relief. This may cause individuals who engage in NSSI to develop positive associations with things like blood, knives, wounds, and similar stimuli that people normally find extremely aversive (e.g., Bradley, Codispoti, Cuthbert, & Lang, 2001). Several studies are consistent with this hypothesis. In a study with fruit flies, Tanimoto, Heisenberg, and Gerber (2004) paired one neutral odor with the onset of pain and another with the offset of pain. The flies avoided the odor associated with the onset of pain and approached the odor associated with the offset of pain. In humans, Zanna, Kiesler, and Pilkonis (1970) paired one set of adjectives with the onset of a painful electric shock and another set of adjectives with the offset of shock. Participants developed negative attitudes toward adjectives paired with the onset of shock and positive attitudes toward adjectives paired with the offset of shock. Interestingly, these effects generalized to adjectives that were similar to those included in the word sets. This suggests that this effect involves conditioning with concepts (e.g., ideas related to NSSI) rather than specific stimuli (e.g., a single word or a specific picture).

Andreatta et al. (2010) obtained similar results using neutral shapes as stimuli and startle eyeblink reactivity as the dependent measure. Specifically, startle eyeblink was larger in the presence of shapes paired with pain onset and smaller in the presence of shapes paired with pain offset. Finally, Glenn and Klonsky (2010) found that a large proportion of self-cutters reported that they engage in NSSI in order to see blood, and that blood is what brings them relief. This finding fits well with the above hypothesis because (a) blood is present during pain offset and is likely to become associated with relief and (b) suggesting a learning mechanism, participants with a longer history of NSSI were more likely to report a positive attitude toward blood.

The second aim of the present study was to test the hypothesis that pain offset relief conditioning is one route through which people develop positive attitudes toward NSSI. Such attitudes could greatly increase the likelihood of future NSSI by reducing avoidance of NSSI-related stimuli that are normally regarded as aversive (cf. Bradley et al., 2001). To test this, in a group of people with no history of NSSI, I examined attitudes toward NSSI before and after a pain offset conditioning procedure that paired NSSI stimuli with pain offset. This test has the potential to provide unique insight into how positive attitudes toward NSSI develop and, correspondingly, may establish new treatment targets for NSSI.

The Reversal of Positive Attitudes toward NSSI: A Potential Treatment

As noted above, traditional treatment approaches have not been effective for reducing NSSI (Nock, 2010). Counter-conditioning of attitudes may represent a new route through which NSSI can be reduced. Specifically, given that attitudes strongly predict behavior (e.g., Nock et al., 2010), changes in attitudes may systematically change behavior. Several studies clearly demonstrate that attitudes can be changed (for reviews, see De Houwer et al., 2001; Hofmann et al., 2010). This is true even for heavily ingrained attitudes such as racial prejudice (Olson & Fazio, 2006). More directly relevant to the present study, there is also evidence that conditioning can alter implicit attitudes toward, and prospectively reduce, addictive behaviors. For example, Houben, Schoenmakers, and Wiers (2010) paired beer-related pictures with negative words and pictures within a conditioning procedure. Compared to the control group, the experimental group displayed more negative implicit attitudes toward beer, experienced less craving for beer, and consumed less beer immediately after the manipulation and over the course of the week following the manipulation.

The third aim of the present study was to reduce or reverse positive attitudes toward NSSI in participants with a history of NSSI (who presumably have positive implicit attitudes toward NSSI, cf. Nock & Banaji, 2007), and to determine whether this reduces NSSI over the following six months. The manipulation involved pairing NSSI stimuli with the onset of pain (cf. Andreatta et al., 2010). Half of the NSSI participants received this manipulation and the other half received a control manipulation. If the experimental NSSI group displays reduced NSSI, this would provide empirical support for aversive conditioning as a potential treatment for NSSI. Conditioning with painful stimuli may not be ideal for therapy, but meta-analytic evidence indicates that painful electrocutaneous stimulation generates the largest effect sizes for changing attitudes (Cohen's $d = 1.16$ for painful electrocutaneous stimulation vs. $d = .5$ for other forms of unconditioned stimuli; Hofmann et al., 2010). Accordingly, painful stimuli may be optimal for investigating the feasibility of conditioning procedures as a treatment for NSSI. This may open the door for future investigations with forms of conditioning that may be more amenable to therapy. For example, pairing NSSI stimuli with pictures of snakes or spiders may be sufficient to reverse positive implicit associations with NSSI and to reduce NSSI behaviors (cf. Houben et al., 2010).

General Overview of Experimental Design

All participants were adults recruited through Introductory Psychology classes and advertisements posted in the community, university, and hospital. Because a recent national survey found that there were no NSSI differences based on gender, ethnicity, or income level (Klonsky, 2011), participants were not selected based on these factors. Given that self-cutting is the prototypical NSSI behavior and attitude measures were specific to self-cutting, NSSI participants were selected only if they had a history of self-cutting. Although the present findings may have implications for all NSSI behaviors, the present primary hypotheses focused on self-cutting. An overview of the experimental design is provided in Figure 2. After recruitment, assessment, and startle habituation there were five major parts of the experimental portion of the study.

In part one, I administered the five attitude measures (i.e., the IAT, AMP, startle eyeblink, and startle postauricular, and explicit ratings) to all participants. This partially addressed the first aim of the present study by testing baseline differences in these measures between the NSSI (i.e., self-cutting) and Control groups.

Part two was the conditioning phase of the experiment. Portions of the NSSI group underwent an aversive conditioning procedure whereas portions of the control group underwent a pain offset relief conditioning procedure. In part three, I re-administered the five attitude measures to all participants post-conditioning. Parts two and three addressed the second aim of the present study by examining the effect of pain offset relief conditioning on attitudes toward self-cutting in participants without a history of self-cutting. Parts two and three also partially addressed the third aim of the experiment by testing the ability of aversive conditioning to reverse positive implicit attitudes toward self-cutting in participants with a history of self-cutting. Part four of the experiment involved administering an aversive conditioning procedure to the portion of the control group that received pain offset relief conditioning. The purpose of this was to counteract any positive implicit attitudes toward self-cutting engendered by pain offset relief conditioning.

Part five of the experiment consisted of a six month follow-up survey for all NSSI participants to assess NSSI behaviors. This part of the experiment addressed portions of the first and third aims of the present study. Specifically, this permitted a test of how well attitude measures prospectively predicted future self-cutting behaviors. Additionally, this allowed for a preliminary test of the efficacy of aversive conditioning as a method for reducing self-cutting behaviors.

Summary

Millions of people engage in NSSI each year, with great costs to mental and physical health. Unfortunately, this trend is likely to continue because knowledge about several fundamental aspects of NSSI remains inchoate. The present study has the potential to greatly advance knowledge about NSSI prediction, development, and treatment. To accomplish these aims, I conducted a novel study with a multi-method, experimental, and longitudinal approach. The findings of the present study may lay the foundation for much-needed improvements in the understanding of this perplexing behavior.

Methods

Participants

General Inclusion/Exclusion Criteria. Participants were excluded if they were below age 18 or endorsed auditory problems, blindness, or psychosis. Although adolescents report high rates of NSSI (Jacobson & Gould, 2007), for two major reasons, the present study only included adults. First, it is less feasible to conduct a study involving electrocutaneous stimulation and NSSI stimuli in adolescents (e.g., harder to obtain consent from parents to enact this procedure on their children). Second, among adults who endorse NSSI, the average onset age is 16 (Klonsky, 2011). This gives adult participants more time to develop stronger positive attitudes toward NSSI (cf. Glenn & Klonsky, 2010), thereby increasing the likelihood of detecting significant differences in NSSI attitudes. This also provided a more stringent test of the efficacy of the aversive conditioning procedure: if this intervention can reverse heavily ingrained attitudes toward NSSI, it should be even more powerful for individuals who recently began engaging in NSSI. Because the NSSI stimuli only depicted cutting (see below), individuals who engaged in forms of NSSI other than self-cutting were excluded from the study (unless these behaviors were in addition to cutting). As mentioned above, a recent national survey of randomly sampled adults in the United States found that NSSI status and function did not vary with gender, ethnicity, or household income (Klonsky, 2011). Accordingly, participants were not selected or excluded based on these demographic variables. Posthoc tests indicated that no demographic variable (age, gender, ethnicity) was significantly associated with any of the five self-cutting attitude variable (all *p*s > .05).

Control Group. The control group consisted of 86 (46 females) college undergraduates who participated to partially fulfill an introductory psychology research option. Participants were only included in this group if they reported no history of any form of NSSI, as assessed by the Self-Injurious Thoughts and Behaviors Interview (SITBI, Nock et al., 2007; see below). This group had an average age of 19.24 (Mdn = 19.00; SD = 1.31) and an ethnic composition of 67.4% Caucasian American, 11.6% African American, 11.6% Asian American, and 9.3% mixed/other. Nearly half of these participants (n = 40) were sorted into the control-offset subgroup and received pain offset relief conditioning. The remaining participants (n = 46) comprised the control-control subgroup and underwent the random conditioning procedure. Participants were randomly sorted into groups based on the week that they participated in the present study. This is because it required an hour to alter the conditioning setup and to ensure that the setup worked properly after being altered (e.g., to make sure that the timing between shocks and picture presentations were accurate for all trials). Because participants were often scheduled consecutively throughout the day, it was not feasible to randomly assign conditioning procedures on a participant-by-participant basis. Accordingly, the conditioning procedure alternated each week, resulting unequal subgroup sizes.

NSSI Group. The NSSI group consisted of 58 (43 females) participants recruited from campus, community, and hospital advertisements that offered $75 for study participation. This group had an average age of 23.24 (Mdn = 21.00; SD = 6.90) and an ethnic composition of 63.8% Caucasian American, 6.9% African American, 13.8% Asian American, 3.4% Hispanic American, and 12.1% mixed/other. Participants were only included in this group if they reported a history of self-cutting as assessed by the SITBI. Self-cutting was selected as a necessary behavior because: (a) our experimental tests were specific to cutting; (b) in contrast to behaviors such as biting, cutting is an unambiguously severe NSSI behavior; and (c) several studies have reported that cutting is the most common NSSI behavior (e.g., Nock & Prinstein, 2004). The mean number of lifetime self-cutting episodes for this group was 146.02 (Mdn = 25; SD = 469.25; Range = 1 to 3000). The mean number of months since the last self-cutting episode was 12.12 (Mdn = 1.00; SD = 27.06; Range = 0 to 162).

Over half of these participants were sorted into the NSSI-aversive subgroup (n = 33) and received aversive conditioning. The remaining participants constituted the NSSI-control subgroup (n = 25) and underwent the random conditioning procedure. As with the control group, the conditioning procedure alternated each week, producing unequal subgroup sizes. Six months after the lab visit, NSSI participants were offered $10 for completion of an online survey to assess a variety of factors related to their self-injury; 49 participants completed this follow-up survey.

Self-Report Measures

Self-Injurious Thoughts and Behaviors Interview (SITBI; Nock, Holmberg, Photos, & Michel, 2007). The SITBI is a structured interview that measures the presence, frequency, and characteristics of various types of self-injurious thoughts and behaviors. It has modules for suicidal ideation, suicide plans, suicide gestures, suicide attempts, and NSSI. The SITBI has strong interrater reliability (average K = .99; r = 1.0) and test-retest reliability (average K = .70; intraclass correlation coefficient = .44) over a six month interval (Nock et al., 2007). The SITBI shows strong construct validity, converging with other measures of suicidal ideation (average K = .54), suicide attempts (average K = .65), and NSSI (average K = .87). In the present study, I utilized the NSSI behaviors module, which consists of ten questions.

Mini International Neuropsychiatric Interview (MINI; Sheehan et al., 1998). The MINI is a short, structured diagnostic interview that assesses constructs from the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV). The MINI has shown high interrater reliability (K = 1.0) and test-retest reliability (K = .87), and shows strong convergent construct validity with longer structured clinical interviews of general psychopathology (Sheehan et al., 1995, 1998) In the present study, all participants were administered the MINI modules for mood, anxiety, substance abuse/dependence, eating, and psychotic disorders.

Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004). The DERS is a 36-item self-report questionnaire about emotion regulation (and dysregulation), with higher scores indicating poorer regulation. All items are scored on a 1 (“almost never”) to 5 (“almost always”) Likert scale, resulting in a possible score range of 36 to 205 (M = 77.99; SD = 20.72 for women; M = 80.66; SD = 18.79 for men). The DERS contains six subscales: Nonacceptance (e.g., “when I’m upset, I feel guilty for feeling that way”); Goals (e.g., “when I’m upset, I have difficulty concentrating”); Impulse (e.g., “when I’m upset, I lose control over my behaviors”); Awareness (e.g., “I am attentive to my feelings” [reverse scored]); Strategies (e.g., “when I’m upset, I believe I’ll end up feeling very depressed”); and Clarity (e.g., “I have difficulty making sense of my feelings”). The DERS shows high internal consistency, good test-retest reliability, adequate construct and predictive validity, and is positively correlated with NSSI in both men and women (see Gratz & Roemer, 2004). Because NSSI and Control groups differed on emotion dysregulation (see below), including this measure permitted me to statistically control for emotion dysregulation, thereby ruling out this variable as a potential explanation for group differences.

Emotion Reactivity Scale (ERS; Nock, Wedig, Holmberg, & Hooley, 2008). The ERS measures emotion reactivity, a component of emotion regulation that specifically involves emotional sensitivity, intensity, and persistence. The ERS is a 21-item self-report questionnaire that displays strong internal consistency, convergent and divergent construct validity, and criterion-related validity (Nock et al., 2008). The ERS is positively correlated with self-injury and has been shown to mediate the association between psychopathology and self-injury (Nock et al., 2008). Similar to the DERS, NSSI and control groups differed on emotion reactivity (see below) and controlling for ERS scores eliminated this difference as a potential explanation for group differences.

Online follow-up NSSI survey. This survey was created specifically for this study and designed to assess several aspects of self-injury. This survey was essentially the NSSI module of the SITBI plus several additional questions. Participants were asked about the nature and frequency of their self-injury (a) over their entire lifetime; (b) during the six months prior to coming into the laboratory (the SITBI did not include this specific assessment); and (c) during the six months after coming into the laboratory.

Brief Behavioral Tests

The NSSI Implicit Association Test (IAT; Nock & Banaji, 2007). The IAT is a brief computer-based task that measures implicit associations. The NSSI IAT (identity version) assesses the strength of the association that participants hold between themselves and self-cutting (Nock & Banaji, 2007). As depicted in Figure 3, participants are presented with four words or phrases at the top of the screen. Two of these words are opposing concepts (e.g., Cutting, No Cutting) and two of these words are opposing attributes (e.g., Me, Not Me). The screen also contains a picture that is either related to self-cutting (e.g., pictures of skin that had been cut) or neutral (e.g., non-injured skin). The pictures in the present study were identical to those of Nock and Banaji (2007). Participants are asked to press one key (e.g., "e") if the concept on the left correctly classifies the picture (e.g., "Cutting" for a picture of skin that had been cut), and another key (e.g., "i") if the concept on the right correctly classifies the picture (e.g., "Not Cutting" for a picture of non-injured skin). Reaction time is recorded for each trial on which a correct classification was made. If an incorrect classification is made, the trial is repeated.

The NSSI IAT is composed of two blocks. In one block, NSSI concepts (e.g., cutting) are on the same side as self-related attributes (e.g., "me"). Compared to non-NSSI participants, participants with a history of self-cutting classify self-cutting pictures more quickly in this kind of block (Nock & Banaji, 2007). In the other block, self-cutting concepts are not on the same side as self-related concepts. In this kind of block, non-NSSI participants are faster at correctly classifying pictures related to self-cutting (Nock & Banaji, 2007). The strength of association between cutting and oneself is quantified by taking a standardized difference score (i.e., response latencies for the Cutting/Me block subtracted from the Cutting/Not Me block) for each participant (Greenwald, Nosek, & Banaji, 2003). Positive difference scores reflect stronger associations.

The Affect Misattribution Procedure (AMP; Payne et al., 2005). The AMP is a brief computer-based task that measures implicit affect. On each trial of the AMP, the computer flashes an emotionally evocative picture for 75ms, a blank screen for 125ms, a Chinese symbol for 100ms, and finally a gray screen until the participant presses a key. Participants are instructed to press one key if they judge the Chinese symbol to be more pleasant than the average symbol and another key if they judge it to be less pleasant that the average symbol. Participants are instructed to ignore the emotionally evocative pictures during their judgments of the Chinese symbols. Despite this instruction, evaluations are influenced by the nature of the picture, with more pleasant pictures generating more pleasant evaluations of subsequent Chinese symbols (Payne et al., 2005). The dependent variable for the AMP is the proportion of trials on which a positive evaluation occurs to the total number of trials within a given picture category (note: results were nearly identical when AMP scores were calculated as the percent difference from neutral category responses). The AMP typically includes standard pleasant, neutral, and unpleasant picture stimuli categories. In the present study I created a version of the AMP designed specifically to examine implicit affect toward self-cutting. In addition to pleasant, neutral, and unpleasant categories, the NSSI AMP included a category of NSSI pictures that depicted self-cutting. Each category consisted of 12 pictures, each presented twice during the AMP (cf. Payne et al., 2005, 2008). Pictures were presented in a random order. This resulted in 96 total trials. Although the NSSI AMP included four picture categories, the primary hypotheses of the present study only involved the NSSI category.

Explicit affective rating survey. Explicit affect toward the same 48 images used in the NSSI AMP was assessed with a computerized self-report survey. On each trial, a picture was presented and participants were asked to rate it on a scale of -4 (extremely unpleasant) to 0 (neutral) to +4 (extremely pleasant). After participants made their ratings, they clicked on a button to move on to the next trial. Ratings were transformed into a 0 to 9 score and then divided by 9. This produced a proportion score that facilitated comparisons with AMP scores.

Picture stimuli. The pleasant, neutral, and unpleasant pictures were chosen from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2005). The NSSI pictures were developed specifically for this project with a combination of digital art and theatrical makeup techniques. These images depicted various features of a self-cutting episode (e.g., knife approaching the skin, knife on the skin, knife having just cut the skin, multiple self-cutting wounds with blood). Posthoc analyses revealed that all NSSI picture subtypes (e.g., blood vs. no blood) were similarly related to the major variables of interest.

To establish the construct validity of these images (i.e., negatively valenced, regarded as less negative by individuals with a history of NSSI), 304 Introductory Psychology students evaluated these pictures as part of a separate pilot study conducted before the present study began. On a scale ranging from very pleasant (+9) to very unpleasant (0), all participants regarded the self-cutting images as very unpleasant, but the individuals with no history of

NSSI (n = 273; ratings: M = 1.52, SD = .54) rated these images as significantly more unpleasant than individuals with a history of NSSI (n = 31; ratings: M = 1.74, SD = .63), $F(1, 303) = 4.58$, $p = .03$, $d = .37$. These preliminary findings support the construct validity of the self-cutting images, even with an NSSI group that self-reported infrequent (M = 5.48 episodes; SD = 7.84 episodes; Median = 2 episodes) and nonspecific behaviors (primary form: cutting = 51.6%, scraping = 19.4%, burning = 6.5%, hitting = 12.9%, and other = 9.7%). In conjunction with the fact that the self-cutting images were specific to self-cutting, these factors may have diminished group differences in the evaluation of these images.

Psychophysiological Measures

Startle eyeblink reactivity. Startle eyeblink reactivity typically has been regarded as psychophysiological measure of two closely related constructs: defensive motivation and negative affective valence (Lang et al., 1990; Vaidyanathan et al., 2009). Although some regard motivation and affective valence as essentially the same (Lang, 2010), others note differences between these constructs (e.g., Carver & Harmon-Jones, 2009). Based on the sum of the evidence in the startle eyeblink reactivity literature, I conceptualize startle eyeblink reactivity as reflecting the degree to which physiological resources are marshaled to defend against a perceived threat (cf. Lang, 2010).

Following recommended guidelines (Blumenthal et al., 2005), I measured reactivity by placing two electrodes on the orbicularis oculi, the muscle that surrounds the eye and generates a blink when it contracts. A third electrode was placed on the temple to serve as a ground. Eyeblink is measured because it is among the fastest and most habituation-resistant components of the startle reflex (Blumenthal et al., 2005). Startle was elicited by a 100dB(A) stimulus delivered through headphones. During the 13-trial habituation block (see

Lane, Franklin, & Curran, 2013), startle stimuli were delivered randomly every 13 – 23s in the absence of foreground stimuli. During the affective-valence startle modulation blocks, startle stimuli were presented between 3.5s and 6.5s after the onset of IAPS and self-cutting pictures (cf. Bradley et al., 1999; Benning et al., 2004). To facilitate comparisons across measures, these pictures were identical to those used in the AMP. Similar to the AMP, primary hypotheses for startle modulation only concerned the self-cutting picture category.

Startle Postauricular Reactivity. Startle postauricular reactivity is a psychophysiological measure of appetitive motivation (Gable & Harmon-Jones, 2009; Johnson et al., 2012). Similar to the affect versus motivation interpretations noted for startle eyeblink reactivity above, appetitive motivation overlaps substantially with the construct of positive affect (e.g., Lang, 2010), but there are important differences between these constructs (see Carver & Harmon-Jones, 2009). In particular, whereas positive affect may refer to any pleasant state, appetitive motivation specifically reflects the degree to which resources are gathered to approach an object in the environment.

Following recommended guidelines, I measured postauricular reactivity by placing two electrodes along the postauricular muscle behind the ear – one on the backside of the pinna and the other on the scalp (Benning et al., 2004). Because startle eyeblink and postauricular reactivity were measured concurrently, startle postauricular reactivity used the same ground electrode, startle stimulus, and foreground stimuli as startle eyeblink reactivity.

Equipment. All equipment and methods were in accord with currently accepted guidelines (Blumenthal et al., 2005). Startle stimuli were 100dB(A) broadband white noises (20hz – 20kHz) with an immediate rise/fall time and a duration of 50ms. These stimuli were generated by Adobe Audition. Picture stimuli and startle stimuli were delivered by SuperLab, with pictures being displayed on a monitor in front of the participants and startle stimuli being presented through Sennheiser PX200 headphones on the participants. InVivo Metric electrodes with an 11mm outer diameter and a 4mm inner diameter transmitted electromyographic (EMG) activity from the eyeblink and postauricular muscles to Biopac EMG amplifiers and a Biopac MP150 work station, which concerted muscle activity to digital information. The MP150 sampled the data at a rate of 1,000Hz and stored four versions of the EMG data: raw unfiltered EMG, filtered EMG in a passband of 28 – 500Hz, a rectification of the filtered EMG signal, and a rectified derivation of the filtered signal smoothed with a five-sample boxcar filter. This latter form of the EMG data was presented in Acqknowledge and was the basis for analyzed EMG data.

EMG scoring. EMG data were scored offline in Acqknowledge. For startle eyeblink reactivity, the scoring window was 20 – 120 ms after on the onset of the startle stimulus (Blumenthal et al., 2005). For startle postauricular reactivity, the scoring window was 8 – 20 ms after the onset of the startle stimulus (Benning et al., 2004). Response magnitude for both types of startle reactivity was the average of the difference between peak and onset voltage of the smoothed EMG. Trials on which a response cannot be detected were assigned a value of 0. Trials on which a response could not be detected even if it were presented (e.g., EMG noise contamination due to a yawn) were scored as “missing.”

Quantification of startle eyeblink and postauricular reactivity on NSSI trials. There is no universally accepted way to quantify affective-valence startle modulation data within stimulus categories (e.g., pleasant, neutral, unpleasant). Some studies simply average raw reactivity within categories, others compare reactivity in specific categories to reactivity on neutral trials (or non-picture trials), and still others transform data into standardized z- or t-scores (see Blumenthal et al., 2005 and Vaidyanathan et al., 2009). In the present study, I used a combination of the latter two methods in an attempt to quantify reactivity in a way that was standardized, clearly interpretable, and robust against potential group and conditioning effects on overall reactivity (e.g., higher overall startle eyeblink reactivity in the NSSI group). Specifically, reactivity for pleasant, unpleasant, and NSSI categories were calculated as: [Valenced Category Mean – Neutral Category Mean]/Neutral Category Mean. This transformed reactivity in each valenced category into a "percent difference from reactivity on neutral trials" score. This method was more conservative compared to other calculation methods (e.g., raw reactivity), but all methods produced very similar patterns of reactivity data.

Electrocutaneous stimulation. Electrocutaneous stimulation was generated by a Biopac STM100C, routed through a Biopac STMISOE, and delivered to the nondominant forearm via two disposable Biopac EL-503 electrodes connected to the STIMISOE with two Biopac 110 lead wires. Electrocutaneous stimulus intensity varied across participants depending on the intensity that each participant rated as 'moderate' (see procedure below), but each stimulus lasted 200ms and had an immediate rise/fall time (cf. Andreatta et al., 2010).

Procedure

Figure 2 provides a general overview of the study procedure. Below, the procedure of each portion of the study is described in detail. In total, the experiment lasted approximately 2.5 hours for each participant.

Interviews and questionnaires. After signing an informed consent form, two structured interviews (the MINI and the NSSI module of the SITBI) and two questionnaires (the DERS and ERS) were administered. Depending on the degree of psychopathology and NSSI history, this portion lasted 15 to 30 minutes.

Startle preparation and habituation. To prepare for startle measurement, two electrodes were placed around the eye, two behind the ear, and one on the temple (cf. Benning et al., 2004; Blumenthal et al., 2005). These remained on until the end of the experiment. Headphones were then placed on the participants. The first block was the startle habituation block, which consisted of 15 trials of 100dB stimuli being presented at intervals of 13 to 23secs. The purpose of this block was to bring startle reactivity in each participant to an asymptotic level. Traditionally, startle studies have employed very few habituation trials (i.e., 0 to 3); however, in a recent empirical study of startle habituation, my colleagues and I determined that startle eyeblink reactivity reaches an asymptote around 13 trials (Lane, Franklin, & Curran, 2013). This block lasted approximately ten minutes.

Baseline attitude measurement. Next, I administered the five attitude measures (AMP, IAT, startle eyeblink/postauricular, and explicit ratings). Together, it took approximately 40 minutes to administer these measures.

Electrocutaneous conditioning procedures. All groups underwent a conditioning procedure. Each conditioning procedure involved randomly presenting 18 IAPS pictures (6 each from pleasant, neutral, and unpleasant categories) and 18 self-cutting pictures (the 12 from the AMP/AVSM/explicit tasks and the 6 from the IAT). Each image was presented for 2 seconds and the intertrial intervals varied from 15 to 25 seconds. Each procedure lasted approximately 15 minutes and involved 18 shocks. These shocks consisted of 200ms 150Hz electrocutaneous stimulations delivered via two electrodes attached to the left bicep. Similar to the methods of previous studies (Andreatta et al., 2010; Franklin et al., in press), stimulus intensity was adjusted for each participant until they rated the shocks as a “30” on a 0 (no pain) to 100 (worst pain imaginable) scale. This level was assessed twice to ensure the validity of the ratings. The only differences among the procedures were the contingencies between the shocks and the pictures.

The aversive conditioning procedure was administered to the NSSI-aversive group. This procedure involved the presentation of a shock immediately after a self-cutting picture disappeared from the screen. This established a contingency whereby the self-cutting pictures predicted a painful shock.

The pain offset relief conditioning procedure was administered to the control-offset group. This procedure involved the presentation of a self-cutting image 6 seconds after each shock. This interval was chosen because a prior pain offset relief conditioning study (Andreatta et al., 2010) and my own basic pain offset relief work (Franklin et al., in press) indicated that pain offset relief is most powerful around 6 seconds after pain offset (and disappears around 14 seconds). The procedure established a contingency where self-cutting images were associated with relief.

The random conditioning procedure was administered to the control-control and NSSI-control groups. This procedure was designed to provide a strong experimental control for the procedures above, ruling out the contributions of pain assessment, electrode attachment, shock anticipation, and number of shocks to group differences. In this procedure, six shocks occurred six seconds before normative images (i.e., two shocks before two pleasant, neutral, and unpleasant pictures), and six shocks occurred immediately after normative images. For the remaining six shocks, three occurred six seconds before self-cutting images and three occurred immediately after self-cutting images.

Post-conditioning assessment with implicit measures. Post-conditioning assessment was identical to the baseline assessment.

Aversive conditioning for the control-offset group. Because the conditioning procedure for the control-offset group may have induced a positive implicit attitude toward self-cutting, the aversive conditioning procedure was administered to obviate any such effects. This procedure was identical to that administered to the NSSI-aversive group.

Compensation. After completing their final task, participants were debriefed, compensated (credit for Introductory Psychology students, $75 for other participants), and allowed to leave.

Six month follow-up procedure. Contact information was collected from all NSSI participants at baseline (name, email, phone number). During debriefing, participants were informed of a six-month follow-up survey option for which they would be compensated $10. Six months after their laboratory visit, each NSSI participant was contacted and asked to complete the online follow-up survey (see above). Upon completion of this survey, participants were mailed their compensation.

Data Analytic Plan

Preliminary analyses. Before examining the primary hypotheses, I conducted several preliminary analyses. First, I attempted to replicate the finding of increased ERS and DERS scores in individuals with a history of NSSI (Gratz & Roemer, 2008; Nock et al., 2008). This finding would indicate that ERS/DERS scores should be employed as covariates in analyses that examine differences between NSSI and control groups. Second, as a manipulation check, I attempted to replicate the linear patterns across pleasant, neutral, and unpleasant categories for startle eyeblink, postauricular, AMP, and explicit affective ratings measures. Third, I examined zero-order correlations among the five attitude measures, ERS/DERS scores, self-cutting frequency (raw number of lifetime, past year, past six months, past month, and post-lab past six months; note: class interval and rank-order analyses yielded the same results), and self-cutting recency (raw number of months since last self-cutting episode; note: year-based and rank-order analyses yielded the same results).

Common Features of Major Analyses

All major analyses that examined differences between control and NSSI groups included ERS and DERS scores as covariates. Analyses that included diagnoses as a covariate produced nearly-identical results to analyses without this covariate. Additionally, consistent with prior studies (e.g., Nock et al., 2008), any effects of diagnoses were completely redundant with ERS and DERS scores. Accordingly, diagnoses were not included as covariates in the presented analyses. Longitudinal analyses controlled for self-cutting frequency over the six months pre-lab visit. Most 'prior self-cutting frequency' variables were strongly intercorrelated (see Table 1). Given that the outcome variable in prospective analyses was self-cutting during the six months post-lab visit, the 'self-cutting frequency six months pre-lab visit' was the most natural conceptual covariate. However, as noted in the results section, analyses that included other self-cutting variables as covariates yielded similar results.

For concurrent analyses, the dependent variables were the five attitude measures (IAT, AMP, explicit affective ratings, startle eyeblink, and startle postauricular). Although the four affective measures included categories for pleasant, neutral, and unpleasant stimuli, primary hypotheses only involved the NSSI category. For prospective analyses, the dependent variable was self-cutting frequency during the six months after the laboratory visit (and in exploratory analyses, a composite of all NSSI behaviors during the six months after the laboratory visit). All alpha levels were .05.

Baseline Differences and Prediction

Group differences on implicit measures at baseline. It was hypothesized that the NSSI group would display significantly more positive affect toward self-cutting on the AMP, explicit affective ratings, startle eyeblink reactivity, and startle postauricular reactivity, and significantly more positive associations with self-cutting on the IAT. This hypothesis was first tested with a between-groups multivariate analysis of covariance (MANCOVA). Pending a significant multivariate effect, follow-up between-groups analyses of covariance (ANCOVAs) were conducted for each individual dependent variable.

Longitudinal prediction of self-cutting from baseline implicit measures. It was hypothesized that all five attitude variables would significantly and independently predict self-cutting frequency six months after the lab visit. To test this hypothesis, first I examined zero-order correlations among self-cutting frequency six months post-lab visit, potential covariates (ERS and DERS scores, pre-lab visit self-cutting frequency variables), and the five attitude variables. To control for skew, both self-cutting frequency variables were transformed into ordinal variables (note: results were very similar with raw frequency). Pending significant zero-order correlations, covariates were entered into a hierarchical multiple regression analysis on successive steps and attitude variables were entered on the final step.

Conditioning Effects

Effects on attitude measures. It was hypothesized that aversive conditioning (NSSI-aversive group) would normalize attitudes, pain offset relief conditioning (control-offset group) would generate more positive attitudes toward self-cutting, and random conditioning (control-control and NSSI-control groups) would not affect attitudes. For each measure, these hypotheses were tested with a 3 x 2 (conditioning type x measurement time) mixed ANCOVA. Pending a significant interaction effect, for each conditioning type, follow-up paired-samples t-tests were conducted to examine attitude change for each conditioning type.

Aversive conditioning effects on future self-cutting. It was hypothesized that the aversive conditioning would reduce future self-cutting behaviors. To test this, a 2 x 2 (NSSI subgroup x NSSI assessment window) mixed ANCOVA was conducted to assess changes in self-cutting rates pre- and post-conditioning across the two NSSI subgroups. Pending a significant interaction effect, follow-up paired-samples t-tests were conducted to examine self-cutting frequency change in each group.

Results

Diagnoses

Control group. Within this group, the MINI indicated low levels of current Axis I psychopathology: major depressive disorder (n = 2); panic disorder (n = 2); social phobia (n = 6); alcohol abuse (n = 6); alcohol dependence (n = 12); substance abuse (n = 2); substance dependence (n = 2); generalized anxiety disorder (n = 6); and no cases of self-injury, bipolar disorder, obsessive-compulsive disorder, posttraumatic stress disorder, eating disorders, or psychotic disorders. Diagnoses were not significantly associated with any attitude variables (ps >.05).

NSSI group. The MINI indicated moderate levels of current Axis I psychopathology in this group: major depressive disorder (n = 16); suicide attempts (n = 13); bipolar disorder (n = 4); panic disorder (n = 16); social phobia (n = 9); posttraumatic stress disorder (n = 9);

alcohol abuse (n = 5); alcohol dependence (n = 7); substance abuse (n = 4); substance dependence (n = 1); bulimia nervosa (n = 2); generalized anxiety disorder (n = 18); and no cases of obsessive-compulsive disorder, anorexia nervosa, or psychotic disorders. Diagnoses were not significantly associated with any attitude variables (*p*s >.05).

Analyses that included diagnoses as covariates yielded nearly identical results; accordingly, to maximize statistical power, diagnoses were not included as a covariate in the presented analyses.

Preliminary Analyses

Emotion dysregulation and reactivity. A one-way MANOVA showed that there was a multivariate effect of group (NSSI vs. control) on emotional abnormalities, $F(2, 141) = 30.23$, $p<.001$. Follow-up one-way ANOVAs indicated that the NSSI group displayed higher scores on both the DERS ($F[1, 142] = 49.79$, $p<.001$, $d = 1.16$) and the ERS ($F[1, 142] = 48.65$, $p<.001$, $d = 1.15$). The present means (see Figure 4) and patterns are consistent with several prior studies that have found higher levels of emotion dysregulation and reactivity in individuals with a history of NSSI (cf. Gratz & Roemer, 2008; Nock et al., 2008).

Valence effects on normative stimuli. The normative stimuli in the present study were not directly relevant to the primary hypotheses. As a manipulation check, however, I attempted to replicate the valence effects on normative stimuli (i.e., pleasant, neutral, and unpleasant) that many prior studies have found with the AMP, explicit affective ratings, startle eyeblink reactivity, and postauricular reactivity.

A mixed 2x3 (group [NSSI vs. control] x valence [pleasant, neutral, unpleasant]) ANOVA indicated that there was no significant interaction effect of group and valence on AMP scores ($p = .81$); however, there was a significant main effect of valence, $F(2, 280) =$

105.35, $p<.001$, with contrast effects indicating that this was a significant quadratic effect ($F[1, 140] = 48.63$, $p<.001$). Specifically, the proportion of positive responses diminished from pleasant (M = .71, SD = .20) to neutral (M = .64, SD = .19) to unpleasant categories (M = .36, SD = .24). These findings demonstrate that both groups displayed the normative valence patterns observed in other studies (e.g., Payne et al., 2005).

Echoing AMP results, a 2x3 (group x valence) mixed ANOVA revealed that there was no significant interaction effect of group and valence on explicit affective ratings (p = .15), but that there was a significant main effect of valence ($F[2, 280] = 866.64$, $p<.001$. Contrast effects also indicated that this was a significant quadratic effect ($F[1, 140] = 5.60$, $p = .02$), as pleasant affective ratings diminished from pleasant (M = .77, SD = .10) to neutral (M = .55, SD = .06) to unpleasant categories (M = .31, SD = .10). These results show that both groups displayed the expected valence effects across normative stimuli.

Similarly, a mixed 2x2 (group x valence [pleasant relative to neutral, and unpleasant relative to neutral]) ANOVA found that there was no significant interaction effect of group and valence on postauricular reactivity ($p = .58$), but there was a significant main effect of valence ($F[1, 138] = 5.91$, $p = .02$). Specifically, relative to reactivity on neutral trials, postauricular reactivity diminished from pleasant (M = 4.10%, SD = 34.57%) to unpleasant categories (M = -2.34%, SD = 30.23%). A mixed ANOVA revealed that there was a significant interaction effect of group and valence on startle eyeblink reactivity ($F[1, 139] = 12.12$, $p = .001$). Relative to neutral trials, both groups displayed the expected startle eyeblink reactivity increase from pleasant (M = -9.72%, SD = 32.12%) to unpleasant categories (M = 11.46%, SD = 39.48%); however, the NSSI group displayed a more exaggerated version of this pattern. Together, these results replicate previous findings of valence effects on startle eyeblink and postauricular reactivity (e.g., Sandt et al., 2009).

Intercorrelations among attitudes, ERS/DERS, and NSSI variables. As shown in Table 1, there were several significant associations among attitudes, ERS/DERS, and NSSI variables. Specifically, AMP scores and explicit ratings were moderately positively correlated with one another, but weakly correlated with the other three major variables. As expected, startle eyeblink reactivity was negatively correlated with implicit affect, explicit affect, and implicit associations; curiously, however, postauricular reactivity evidenced a similar pattern.

Also shown in Table 1, several variables were also associated with self-cutting recency and frequency (assessed in various temporal windows). In particular, more positive implicit and explicit affect toward self-cutting stimuli tended to be associated with more recent and frequent self-cutting, and greater startle eyeblink reactivity was associated with less recent self-cutting. Surprisingly, IAT scores were not associated with self-cutting recency or frequency.

Self-cutting frequency variables were generally strongly intercorrelated (see Table 1). Of particular interest were the associations among variables assessed during the lab visit (lifetime, past year, and past month frequency) and variables assessed during the follow-up survey (six months pre-lab, six months post-lab). Despite being assessed at different times, these follow-up variables were strongly correlated with ‘past year self-cutting frequency’ (which was assessed during the lab visit). This is especially important because it supports the construct validity of the ‘self-cutting frequency six months pre-lab visit’ variable, which was questionable because it was assessed during the follow-up rather than during the lab visit.

However, to avoid potential retrospective reporting biases associated with this variable, all analyses involving this variable were duplicated with the 'past year self-cutting frequency' variable.

Although not part of the primary hypotheses, I conducted exploratory posthoc tests to examine associations between attitude measures and 'composite NSSI frequency' before and after the lab visit. These two composite NSSI variables were a sum of all reported severe NSSI behaviors (i.e., cutting, burning, scraping, hitting, etc.). Surprisingly, although the present attitude measures were specific to self-cutting, they were more strongly correlated with the composite NSSI frequency variables than the self-cutting frequency variables (see Table 1). In fact, unlike self-cutting frequency variables, these composite NSSI frequency variables were significantly negatively correlated with startle eyeblink reactivity (indicating that diminished defensive motivation was associated with more frequent NSSI). These results suggest that the present cutting-focused attitude variables may access general attitudes toward NSSI rather than specific attitudes toward self-cutting.

Baseline Group Differences and Prediction

Baseline group differences. A one-way MANCOVA controlling for ERS/DERS scores revealed a significant multivariate main effect of group on the five baseline variables, $F(5, 119) = 9.39$, $p<.001$. Follow-up one-way ANCOVAs clarified that, compared to the control group in the context of self-cutting stimuli, the NSSI group displayed more positive AMP responses ($F[1, 136] = 12.39$, $p<.001$, $d = .60$), more pleasant explicit ratings ($F[1, 141] = 28.72$, $p<.001$, $d = .86$), more positive implicit associations with the self ($F[1, 141] = 15.58$, $p<.001$, $d = .68$), greater postauricular reactivity ($F[1, 138] = 3.50$, $p = .04$, $d = .29$), and diminished startle eyeblink reactivity ($F[1, 139] = 3.95$, $p = .04$, $d = .35$). These findings suggest that individuals with a history of NSSI processes self-cutting stimuli as more pleasant, less threatening, and more strongly related to the self (see Figures 5-7).

Prospective prediction of self-cutting from baseline attitudes. Zero-order correlations revealed that several past self-cutting frequency variables, AMP scores, and explicit affective ratings were all significantly positively associated with self-cutting frequency during the six months after the lab visit (see Table 1). No other correlations reached significance.

Planned hierarchical regression analyses called for ERS and DERS scores to be included on the second step; however, neither of these variables was correlated with any self-cutting frequency variables (see Table 1). Although including these variables led to nearly identical results, to maximize power, the presented analyses do not include these variables. Similarly, to conserve power, only attitude variables with significant zero-order correlations with post-lab self-cutting frequency were included in analyses (analyses including all variables produced very similar results). Additionally, due to the skewed nature of the self-cutting variables, these variables were transformed into rank-order scores. Analyses with non-transformed variables generated nearly-identical results in terms of the predictive power of the attitude variables; however, non-transformed variables tended to diminish the predictive power of prior self-cutting variables. Finally, 'self-cutting six months pre-lab' was chosen as the prior self-injury variable for Step 1 because: (a) conceptually, this six month window matched the six month window of the outcome variable; and (b) it provided the most stringent test for attitude variables (ΔR^2 for the other self-cutting variables ranged from .06 to .17; compared to .19 for the chosen variable). Although the validity of this variable is questionable because it was assessed during the follow-up, it was strongly correlated with ‘self-cutting frequency during the year pre-lab visit’ (see Table 1) and results were nearly identical regardless of which prior self-cutting frequency variables was included.

As shown in Table 2, self-cutting during the six months pre-lab visit significantly predicted self-cutting during the six months post-lab visit. In step 2, AMP scores and explicit affective ratings combined to predict self-cutting above and beyond this powerful covariate; however, neither of these attitude variables independently predicted self-cutting on this step. These results indicate that, although implicit and explicit affect toward self-cutting are somewhat redundant, their combined effect is a strong predictor of future self-cutting that explains variance beyond prior self-cutting.

Prospective prediction of composite NSSI from baseline attitude variables. Although not part of primary hypotheses, inspired by strong zero-order correlations (see Table 1), I conducted exploratory posthoc analyses to examine the ability of attitude measures to predict future general NSSI behaviors (i.e., cutting, along with burning, scraping, hitting, etc.). Echoing the preceding analysis, NSSI variables were transformed into rank-order scores, DERS and ERS scores were not included in analyses, and general NSSI behaviors six months pre-lab visit were included on Step 1 (note: analyses that were non-transformed, included ERS/DERS, and employed other composite NSSI assessments yielded similar results). Diverging from the preceding analysis, startle eyeblink reactivity was included as a predictor on step 2 given its strong association with composite NSSI variables (see Table 1).

As shown in Table 3, composite NSSI frequency during the six months pre-lab visit was a powerful predictor of future NSSI behavior, accounting for nearly 40% of the variance. On step 2, AMP scores, explicit affective ratings, and startle eyeblink reactivity combined to account for significant variance over and above step 1. When simultaneously entered with the other attitude variables on step 2, AMP scores were a significant predictor, and explicit affect ratings (p = .16) and startle eyeblink reactivity (p = .08) trended toward significance. Overall, these results indicate that these three attitude measures – particularly the AMP – are valuable prospective predictors of future NSSI behaviors.

Conditioning Effects

Startle eyeblink reactivity. A mixed ANCOVA indicated that there was a significant interaction effect of conditioning and measurement time on startle eyeblink reactivity, F(2, 125) = 4.29, p = .02. Follow-up paired-samples t-tests indicated that aversive conditioning significantly increased startle eyeblink reactivity in the NSSI-aversive group (t[33] = 3.03, p<.01, d = .55), pain offset relief conditioning significantly diminished startle eyeblink reactivity in the control-offset group (t[35] = 1.83, p = .03, d = .42), and random conditioning had no effect on startle eyeblink reactivity in the control-control and NSSI-control groups (p = .88) (see Figure 8). These findings suggest that pain offset relief conditioning may play a role in reducing defensive motivation toward self-cutting stimuli and aversive conditioning may reverse this process.

Postauricular reactivity. A mixed ANCOVA revealed that there was a significant interaction effect of conditioning and measurement time on postauricular reactivity, F(2, 126) = 2.45, p = .04. Follow-up paired-samples t-tests indicated that aversive conditioning significantly diminished postauricular reactivity in the NSSI-aversive group (t[29] = 1.83, p = .04, d = .42), but pain offset relief conditioning (p = .89) and random conditioning (p = .97) did not significantly affect postauricular reactivity in the other groups (see Figure 9). These findings suggest that aversive conditioning may reverse heightened appetitive motivation toward self-cutting stimuli.

Effects on AMP, IAT, and explicit affective ratings. There were no significant interaction effects of conditioning and measurement time on the AMP, IAT, or explicit affective ratings (*p*s >.05). Unexpectedly, however, analyses indicated a main effect of measurement time for two of these measures. There was no main effect of measurement time on AMP scores (p = .65; pre: M = .33, SD = .25; post: M = .34, SD = .26), but there was for IAT scores (F[1, 134] = 3.64, p = .04, d = .16) and explicit affective ratings (F[1, 134] = 40.59, p<.001, d = .31). Specifically, IAT scores (pre: M = -.16, SD = .51; post: M = -.07, SD = .49) and explicit affective ratings (pre: M = 2.52, SD = 1.42; post: M = 2.94, SD = 1.29) were more positive after conditioning. Given that these effects were not moderated by conditioning type, these findings suggest that mere repeated exposure (see Zajonc, 2001) to self-cutting stimuli diminished negative affect toward, and increased implicit identification with, self-cutting.

Prospective Effect of Aversive Conditioning on NSSI

Follow-up surveys. Forty-nine participants in NSSI group completed follow-up surveys (of a possible 58; for an 84.48% response rate). Of these, 48.8% reported at least one post-lab visit cutting episode, with these individuals accounting for 75 total cutting episodes. In comparison, 77.6% of these follow-up participants reported a cutting episode in the six months pre-lab visit, producing a total of 228 cutting episodes. Similarly, 56.2% of follow-up participants reported at least one post-lab visit NSSI behavior of any kind, for a total of 209 NSSI episodes. In contrast, 89.1% of follow-up participants reported engaging in some form of NSSI at least once during the six months pre-lab visit, producing a total of 489 NSSI episodes. Alleviating concerns about potential ‘NSSI triggering’ caused by the present study, these findings indicate that the study may have prompted a general decrease in NSSI behaviors.

Effects of aversive conditioning on self-cutting. A 2x2 (NSSI subgroup x NSSI assessment window) mixed ANCOVA controlling for ERS/DERS scores did not reveal a significant interaction effect of NSSI subgroup and NSSI assessment window on self-cutting frequency. However, this effect did trend toward significance, $F(1,46) = 3.57$, $p = .07$. Means indicated that, compared to the NSSI-control group, the NSSI-aversive group showed a steeper reduction in self-cutting frequency during the six months after the lab visit (see Figure 10). Follow-up paired-samples t-tests revealed that both the NSSI-aversive ($t[31] = 3.62$, $p = .001$) and NSSI-control groups ($t[16] = 3.19$, $p = .01$) displayed significant reductions in self-cutting frequency after the lab visit.

Given the potential problems with the ‘self-cutting during the six months pre-lab’ (assessed during the follow-up) variable employed in the previous analysis (see above), I attempted to replicate the above findings using the ‘self-cutting during the past year’ (assessed during the lab visit) variable. To account for the fact that this assessment window is 12 months instead of 6, I divided this variable by two. A similar 2x2 (NSSI subgroup x NSSI assessment window) mixed ANCOVA controlling for ERS/DERS scores revealed a significant interaction effect of NSSI subgroup and NSSI assessment window on self-cutting frequency ($F[1, 46] = 6.47$, $p = .01$). Means revealed that self-cutting frequency in the NSSI-aversive group declined steeply, but that self-cutting frequency in the NSSI-control group increased slightly (see Figure 11). Follow-up paired samples t-tests revealed a significant reduction in self-cutting frequency in the NSSI-aversive group ($t[31] = 3.71$, $p<.001$), but not in the NSSI-control group ($t[16] = -.08$, $p = .93$).

Although not part of the primary hypotheses, I conducted exploratory posthoc tests to determine whether the preceding effects would apply to NSSI more generally. A similar 2x2 (NSSI subgroup x NSSI assessment window) mixed ANCOVA controlling for ERS/DERS scores revealed a significant interaction effect of NSSI subgroup and NSSI assessment window on composite NSSI frequency ($F[1, 46] = 9.02$, $p <.001$). As shown in Figure 12, means indicated a steep decline in NSSI behaviors in the NSSI-aversive group ($t[31] = 4.00$, $p<.001$), and a smaller decline in NSSI behaviors in the NSSI-control group ($t[16] = 2.59$, $p = .01$).

Altogether, these findings suggest that aversive conditioning that pairs painful shocks with self-cutting images may longitudinally reduce NSSI behaviors compared to a random conditioning procedure.

Discussion

NSSI is a dangerous, prevalent, and poorly understood behavior. The present study was designed to advance knowledge about the nature, development, prediction, and potential treatment of NSSI. Results indicated that individuals with a history of NSSI possess substantially more positive attitudes toward self-cutting; pain offset relief conditioning and mere repeated exposure may contribute to the development of these abnormal attitudes; and aversive conditioning may normalize certain physiological components of these attitudes. Moreover, results indicated that implicit and explicit affect were strong longitudinal predictors of self-cutting and that aversive conditioning may prospectively reduce self-cutting behaviors.

In conjunction with my other recent work (Franklin et al., 2010, 2011, 2012, 2013), these findings have prompted what I call the Benefits and Barriers Model of NSSI. The central thesis of this model is that NSSI has several short-term benefits (e.g., affect regulation, social reinforcement), but most people never access these benefits because of powerful barriers to NSSI (e.g., pain, aversion to mutilation stimuli, violation of social norms). When these barriers are eroded, however, the model predicts that individuals will be more likely to engage in NSSI. Consistent with the benefits aspect of the model, a large literature has shown that NSSI carries affective and social benefits (e.g., Klonsky, 2007; Nock, 2010; Nock & Prinstein, 2004). Moreover, my recent laboratory studies have shown that one of the primary mechanisms of affect regulation in NSSI – pain offset relief – is equally powerful for non-NSSI and NSSI individuals (even individuals who have engaged in NSSI thousands of times; Franklin et al., 2010, 2013). Similarly, non-NSSI individuals may be vulnerable to the social reinforcement of painful behaviors (Franklin et al., in preparation; cf. Prkachin & Craig, 1986).

These findings suggest that many NSSI benefits may be natural and normal; this places the emphasis on NSSI barriers as the most crucial factors for distinguishing between individuals who do and do not engage in NSSI, predicting NSSI, and treating NSSI. Providing initial support for this idea, recent studies have shown that one major barrier to NSSI, pain perception, effectively distinguishes between non-NSSI and NSSI groups (e.g., for pain tolerance, average d = .70; Hooley et al., 2010; Franklin et al., 2011, 2012). Furthermore, diminished pain perception in NSSI is correlated with NSSI frequency (Hooley et al., 2010; St. Germain & Hooley, in press). This indicates that smaller NSSI barriers translate into more frequent NSSI, although the directionality of this effect remains unclear.

Within the Benefits and Barriers Model of NSSI, the present study could be conceived as a test of two NSSI barriers: (1) the instinctive aversion to self-mutilation stimuli, as assessed by the four affective measures; and (2) the identification barrier, as assessed by the identity version of the NSSI IAT. Specifically, the present study tested: (a) whether these barriers are eroded in individuals who engage in self-cutting; (b) the degree to which this erosion predicts future self-cutting; (c) the ability of pain offset relief conditioning to erode these barriers; (d) the ability of aversive conditioning to rebuild these barriers; and (e) the degree to which rebuilding these barriers prospectively reduces self-cutting. Below, the findings of the present study are discussed in this context.

The Aversion Barrier to NSSI

Consistent with the Benefits and Barriers Model of NSSI, findings across the four affective measures indicated that the "aversion barrier" was diminished in the NSSI group. These effects remained strong even after controlling for emotion dysregulation and reactivity. Effect sizes ranged from large (explicit affective ratings) to moderate (AMP) to small (startle eyeblink and postauricular reactivity).

Patterns across and among affective measures. Suggesting that these four measures captured partially unique features of affect, the intercorrelations among these measures generally were low (see Table 1). The exception to this pattern was the positive moderate correlation between the AMP and explicit affective ratings. Previous studies have also found moderate correlations between measures of implicit and explicit affect (see Payne, Burkley, & Stokes, 2008), and the nature of the present explicit affective ratings task may have facilitated this correlation. This task was self-paced and many participants required only a few seconds to rate each image. This lack of deliberation suggests that many participants primarily relied on their basic affective associations (i.e., "gut feelings" in the terminology of Gawronski & Bodenhausen, 2011) with pictures and did not thoroughly engage propositional reasoning processes. As implicit affect reflects basic affective associations (Gawronski & Bodenhausen, 2006, 2011), the moderate correlation between the AMP and the explicit affective ratings task is not surprising.

The negative correlation between postauricular reactivity and the AMP, however, was unexpected (though it should be noted that this was a small effect). On the surface, this correlation suggests that greater positive implicit affect (AMP) toward self-cutting stimuli was associated with less appetitive motivation toward these same stimuli (postauricular reactivity). One potential alternative explanation for this unexpected pattern is that some participants may have grimaced during the presentation of self-cutting pictures. Johnson et al. (2012) found that grimacing increased overall postauricular reactivity because it generates mechanical priming (vs. affective priming) of the postauricular muscle. In some participants in the present study, extreme negative reactions to self-cutting stimuli may have provoked a grimace, which in turn may have promoted a larger postauricular reaction. This possibility would also help to explain the relatively small group differences in postauricular activity during self-cutting images. Specifically, although greater appetitive motivation may have driven larger mean postauricular responses in the NSSI group, grimacing may have evoked large postauricular responses in some participants in both groups.

Although there were no a priori hypotheses about which measures would most powerfully differentiate the NSSI and control groups, I did not expect that the explicit affective ratings measure would so thoroughly out-perform all other measures. This is because previous studies have indicated that, due to potential factors such as a lack of insight or a desire to conceal, self-report measures are not ideal for investigating NSSI-related phenomena (Franklin et al., 2010; Janis & Nock, 2009). One potential explanation for the strength of this measure is that it captured multiple abnormal processes. Specifically, this measure may have accessed variance due to abnormal affective associations (e.g., self-cutting and pleasant), propositional reasoning based on these associations (e.g., "it seems true that I find self-cutting to be pleasant"), and social behavior stemming from these associative and reasoning processes (e.g., "I am willing to let the researchers know that I find self-cutting images to be pleasant"). This explanation would fit with a psychological constructionist (e.g., Barrett, 2012) interpretation of the overall pattern of affective results. Assuming that NSSI behaviors are constructed from multiple elements (affective, cognitive, social, etc.), measures that access more of these elements will explain more variance in NSSI behaviors. In the present study, it could be proposed that effect sizes diminished from a measure of multiple elements (explicit affective ratings) to overall implicit affect (AMP) to specific motivational components of implicit affect (startle eyeblink and postauricular).

Associations with self-cutting frequency and recency. Echoing findings with the "pain barrier" (Hooley et al., 2010; St. Germain & Hooley, in press), results indicated that AMP scores and explicit affective ratings were significantly correlated with greater self-cutting recency and frequency (see Table 1). Interestingly, exploratory analyses indicated that a composite of the frequency of all reported NSSI behaviors (vs. self-cutting alone) was particularly strongly associated with AMP scores and explicit affective ratings (see Table 1). Overall, these findings support the hypothesis that NSSI behaviors vary with the degree to which NSSI barriers are diminished. These baseline correlations cannot elucidate the directionality of these associations; however, the conditioning and prediction results (discussed below) suggest a transactional relationship between barriers and behaviors. In particular, consistent with the idea that engaging in NSSI can contribute to barrier erosion, pain offset relief conditioning diminished startle eyeblink reactivity in the context of self-cutting images in control participants. On the other end of the transaction, consistent with the idea that eroded barriers generate more NSSI behaviors, diminished aversion to self-cutting images significantly predicted future self-cutting.

Although startle eyeblink reactivity was significantly inversely associated with self-cutting recency, it was not associated with self-cutting frequency (see Table 1). The reason for this lack of association is unclear. Consistent with the psychological constructionist interpretation above, however, it may be that overall affect is a more effective determinant of NSSI behaviors than specific components such as defensive motivation. In-line with the grimace explanation noted above (cf. Johnson et al., 2012), postauricular reactivity was significantly inversely associated with self-cutting frequency. Although speculative, this explanation would be consistent with the hypothesis that greater aversion to self-cutting stimuli (as indicated by a grimace) is associated with less frequent self-cutting.

The Identification Barrier to NSSI

Replicating findings of greater implicit identification with self-cutting in an adolescent sample (Nock & Banaji, 2007), results showed that, compared to the control group, the NSSI group displayed significantly more positive implicit associations between the self and cutting.

Associations with affective measures. Consistent with prior studies (e.g., Payne et al., 2008), associations between the IAT and affective measures were small but consistent. Specifically, IAT correlations with explicit affective ratings and startle eyeblink reactivity reached significance, and correlations with the AMP and postauricular reactivity were in the expected directions. Taken together, this pattern of associations indicates that implicit identification with self-cutting is largely separate from affective attitudes toward self-cutting. In terms of the Benefits and Barriers Model, this suggests that there is little overlap between the aversion and identification barriers: erosion of one barrier does not necessarily imply erosion of another barrier. Although reduction in a single barrier may be sufficient for NSSI to occur, it is likely that NSSI is most probable when multiple barriers are eroded.

Association with self-cutting recency and frequency. Associations between IAT scores and self-cutting recency/frequency were non-significant (see Table 1). These null findings were surprising, especially considering findings that scores on a death/suicide version of the IAT prospectively predict suicide attempts (Nock et al., 2010). However, the present findings are consistent with a recently published study that found that an NSSI version of the IAT did not prospectively predict NSSI (Glenn & Klonsky, 2012). This latter study and the present findings suggest the possibility that – unlike implicit associations with death/suicide – implicit identification with self-cutting may not be associated with past or future NSSI frequency. These findings contrast with the moderate correlations between self-cutting recency/frequency and the AMP and explicit affective ratings.

This pattern of associations implies that conditioning mechanisms may play a role in affective attitudes toward self-cutting (see below), but not in implicit identification with self-cutting. Instead, mechanisms such as spatiotemporal contiguity of the self and cutting (see Gawronski & Bodenhausen, 2011) may contribute to implicit identification with self-cutting. The unexpected finding of significantly stronger implicit identification with self-cutting across the entire sample from pre-conditioning to post-conditioning is consistent with this possibility. Participants viewed the IAT-based self-cutting stimuli six times and the other self-cutting stimuli seven times (a total of 120 self-cutting image views) before the second IAT. The simple act of repeatedly viewing these images may have established stronger implicit associations between the self and cutting. Accordingly, because individuals do not have to engage in self-cutting in order to develop stronger implicit identification with self-cutting, implicit identification with self-cutting may not be strongly associated with the frequency or recency of actual self-cutting behaviors. Rather, implicit identification with self-cutting may be strongly associated with the frequency and recency of activities like viewing self-cutting images online (a surprisingly prevalent phenomenon, see Lewis, Heath, Michal, & Duggan, 2012).

Another possible explanation for the null association between IAT scores and self-cutting recency/frequency is the fact that all individuals in the NSSI group had to self-identify as someone who engages in self-cutting in order to qualify for the NSSI group. Specifically, individuals self-identified during their response to study advertisements, completion of the informed consent form, and the SITBI. This self-identification may have heightened implicit identification with self-cutting in all NSSI participants. Accordingly, even if implicit identification with self-cutting really does vary with self-cutting recency/frequency, this association may have been heavily contaminated and mitigated by the self-identification necessitated for study participation. Given the requirements of participating in an NSSI study, this hypothesis (and potential problem) may be difficult to test.

Prospective Prediction of Self-Cutting

There are few longitudinal predictors of NSSI that are both specific and strong. The ability of the five attitude measures to distinguish between the NSSI and control groups suggested that these measures might hold promise as NSSI predictors. Echoing findings on associations between these attitude measures and lifetime self-cutting frequency, only AMP scores and explicit ratings were moderately positively correlated with the frequency of self-cutting during the six months after the lab visit (see Table 2). Also echoing associations between attitude measures and lifetime self-cutting frequency, IAT scores, startle eyeblink reactivity, and postauricular reactivity were weakly and nonsignificantly correlated with future self-cutting. In similar previous studies (e.g., Glenn & Klonsky, 2012), self-cutting frequency during the six months *after* the lab visit was moderately correlated with self-cutting frequency *before* the lab visit (see Table 2). Unlike predictors in previous studies (e.g., Glenn & Klonsky, 2010), however, the present results indicated that the combined effect of AMP scores and explicit ratings significantly predicted future self-cutting frequency above and beyond prior self-cutting frequency.

Exploratory analyses on a composite of all NSSI behaviors (vs. self-cutting alone) echoed these findings (see Table 3). Likely owing to the greater variability and frequency of this variable, these composite results were more powerful than self-cutting results. Specifically, the combined effect of AMP scores, explicit affective ratings, and startle eyeblink reactivity significantly predicted future NSSI behaviors on step 2 (after controlling for prior NSSI behaviors on step 1). Moreover, AMP scores were a significant unique predictor, and explicit affective ratings and startle eyeblink reactivity trended toward significance. These findings indicate that the preceding self-cutting results may extend to NSSI behaviors more generally.

Consistent with the Benefits and Barriers Model, these prospective findings demonstrate that an eroded aversion barrier (at least on the explicit and performance-based levels) is a prospective risk factor for self-cutting. The present predictors may be particularly valuable because: (a) they are specific to NSSI (e.g., stimuli depict self-cutting and powerfully distinguish between NSSI and non-NSSI groups after controlling for emotional abnormalities); (b) they have the power to predict future self-cutting above and beyond self-cutting frequency in the preceding six months; and (c) they can be quickly, easily, and economically assessed with little training. Future studies should investigate whether other barriers associated with NSSI frequency (e.g., pain tolerance; St. Germain & Hooley, in press) predict future NSSI above and beyond the diminished aversion to NSSI.

Conditioning Effects on Attitudes toward Self-Cutting

I employed two conditioning paradigms to test hypotheses about potential mechanisms of the development and reversal of positive attitudes toward self-cutting. Although conditioning effects are relatively easily obtained when the conditioning target is neutral, it is difficult to condition new attitudes toward familiar or valenced stimuli (e.g., Cacioppo, Marshall-Goodell, Tassinary, & Petty, 1992). Indeed, a meta-analytic review by Hofmann et al. (2010) found that neutral conditioning targets produced moderate conditioning effect sizes (d = .55) whereas valenced targets produced small effects (d = .19). These findings suggest that whereas attitude formation is easy, attitude change is difficult, particularly for strongly valenced and unambiguous stimuli like wounds and blood. Consistent with these findings, there were no conditioning effects on explicit affective ratings, IAT scores, or AMP scores (although two of these measures did show a mere repeated exposure effect). However, there were significant conditioning effects on the two psychophysiological measures. Few previous studies have examined attitude change effects with psychophysiological measures (see Hofmann et al., 2010); the present findings suggest that these attitude components may be more sensitive to attitude change. The reason for this effect is unclear, but consistent with certain attitude theories (Gawronski & Bodenhausen, 2006, 2011; Hofmann et al., 2010), these measures may more directly assess basic affective associations. Whereas these basic associations may shift easily with new environmental contingencies with familiar/valenced stimuli, higher-order associations and propositional reasoning may be more static. In other words, although our gut reactions to familiar stimuli may change, we may not have insight into these changes (cf. Nisbett & Wilson, 1977).

Pain offset relief conditioning. Given that NSSI necessarily involves pain offset relief (Franklin et al., 2013) and stimuli present during pain offset become associated with relief (Andreatta et al., 2010), it is possible that pain offset relief conditioning contributes to the development of positive attitudes toward self-cutting. Although pain offset relief conditioning did not affect most measures, startle eyeblink reactivity results supported this hypothesis. Specifically, pain offset relief conditioning significantly diminished startle eyeblink reactivity in the context of self-cutting, but random conditioning did not affect reactivity levels. This indicates that pain offset relief conditioning during episodes of self-cutting may diminish instinctive defensive motivation toward self-cutting stimuli. In terms of the Benefits and Barriers Model, this suggests that pain offset relief conditioning may partially erode the aversion barrier to NSSI, thereby facilitating future NSSI behaviors.

It is possible that pain offset relief conditioning during NSSI episodes has no effect on any other aspect of attitudes toward NSSI (as suggested by the present results). Nevertheless, it is also possible that 18 trials of pain offset relief conditioning in a laboratory setting was insufficient to alter other attitude components toward such strongly valenced and unambiguous stimuli. Accordingly, the present study cannot rule out the possibility that experiencing hundreds of NSSI episodes (each of which may include numerous instances of pain offset relief) affects other attitude components via pain offset relief conditioning. Regardless of this intriguing possibility, the present study provides experimental support for the hypothesis that pain offset relief conditioning is one viable explanation for the development of at least one aspect of positive attitudes toward self-cutting. As noted above, this mechanism is consistent with evidence that self-cutting recency/frequency are associated with more positive affective attitudes toward self-cutting. Specifically, more frequent/recent episodes may represent more frequent/recent pain offset relief conditioning, which in turn may generate more positive affective attitudes toward self-cutting.

Aversive conditioning. If mechanisms such as pain offset relief conditioning can induce more positive attitudes toward self-cutting, it may be that opposing mechanisms such as aversive conditioning can normalize these attitudes. Although aversive conditioning did not affect explicit affective ratings, IAT scores, or AMP, it did affect the psychophysiological measures. Specifically, aversive conditioning enhanced startle eyeblink reactivity and diminished postauricular reactivity – essentially normalizing psychophysiological reactions toward self-cutting stimuli in individuals with a history of self-cutting. In contrast, random conditioning had no effect on these measures. These findings demonstrate that it is possible to rebuild some components of the NSSI aversion barrier. This is consistent with recent evidence that a brief cognitive intervention targeting self-criticism experimentally normalizes the “pain barrier” to NSSI (Hooley & St. Germain, in preparation). Together, these findings suggest that it may be possible to treat NSSI by rebuilding NSSI barriers. Given that there are currently no empirically supported treatments for NSSI (despite attempts with a range of techniques; see Nock, 2010), this is a particularly intriguing possibility. In the next section, I discuss the prospective effect of aversive conditioning on self-cutting.

Similar to pain offset relief conditioning, it is possible that different forms of aversive conditioning may be able to normalize the non-physiological components of attitudes toward self-cutting. For example, the present study cannot rule out the possibility aversive conditioning with more intense shocks (e.g., with an intensity of “60” instead of “30”), a larger number of trials, or with personalized stimuli may have been able to alter other attitude components. Future studies may benefit from investigating these and other aversive conditioning techniques.

Mere repeated exposure. Results unexpectedly revealed a second mechanism that may contribute to the development of positive (or at least less negative) attitudes toward self-cutting: mere repeated exposure. Across hundreds of studies, the mere repeated exposure of a stimulus has been shown to increase positive affect toward that stimulus and similar stimuli. Zajonc (2001) proposed that this effect exists because the stimulus (a conditioned stimulus) is repeatedly paired with the absence of a noxious consequence (an unconditioned stimulus). Consistent with this effect, results revealed that explicit affective ratings became less negative across all groups from pre- to post-conditioning measurements. As noted above, a similar effect was observed for IAT scores, with this effect being explained by a similar mechanism involving repeated pairings of the self and self-cutting images. Mean AMP levels moved in a similar direction from pre- to post-conditioning measurement, but this effect was nonsignificant and extremely small (d = .04).

Overall, these findings show that the aversion and identification barriers to NSSI can be eroded by mere passive exposure to self-cutting images. It should be noted, however, that these effects were small and, on average, participants still explicitly rated self-cutting images as very unpleasant and maintained negative associations between the self and cutting. Nevertheless, these findings suggest that mere repeated exposure to NSSI stimuli via the media, internet, or peers (Lewis et al., 2012; Whitlock et al., 2006) may increase NSSI risk by contributing to the erosion of NSSI barriers. Within actual NSSI episodes, where more vivid and personally-salient stimuli (e.g., knives, one's own blood and wounds) are present, this effect may be more potent.

Prospective Effects of Aversive Conditioning on Self-Cutting Behaviors

As noted above, attempts with a range of traditional therapeutic techniques have failed to produce an effective treatment strategy for NSSI. According to the Benefits and Barriers Model, NSSI is less likely to occur when the barriers to NSSI are high (i.e., near or exceeding the levels observed in non-NSSI individuals). Although the aversion barrier is eroded in individuals with a history of NSSI (see Figures 5 – 7), aversive conditioning may partially rebuild this barrier (see Figure 8). It follows that this elevated barrier should translate into fewer future NSSI episodes. The present results do not represent definitive evidence for this hypothesis, but all three analyses provide intriguing partial support for this idea (see Figures 10 – 12).

However, each of these analyses possessed a flaw that obscured its support for this hypothesis. First, analyses relevant to Figure 10 included the 'self-cutting frequency six months pre-lab' as its pre-conditioning variable, which is potentially problematic given that this variable was assessed at follow-up (though see above for evidence of its construct validity). Second and similarly, the analyses relevant to Figure 12 employed the ‘composite NSSI six months pre-lab’ variable assessed at follow-up as the pre-conditioning variable. Third, analyses relevant to Figure 11 included the ‘self-cutting frequency one year pre-lab’ as its pre-conditioning variable, but this window was twice as long as the window for the post-conditioning variable (i.e., self-cutting frequency six months post-lab). To compensate, this pre-conditioning value was divided by two; however, the adjustment was artificial and its validity is suspect (though analyses without adjusting the 12-month variable yielded the hypothesized significant interaction effect as well). Additionally, for each of these analyses, the NSSI-aversive group began at a much higher pre-conditioning value than the NSSI-control group; these groups tended to have very similar post-conditioning values. As such, it is possible that the NSSI-aversive group simply evidenced a regression to the mean rather than a treatment effect.

Although it is important to acknowledge these substantial limitations, it should also be noted that the present study is only the second to find a greater reduction in NSSI behaviors in a ‘treatment condition’ compared to a ‘control condition.’ Wood et al. (2001) made the only prior finding of this kind in a study of the effect of group therapy on NSSI; however, nullifying this finding, this same research group later found that group therapy increased NSSI (Hazell et al., 2009). The present study did not produce a new treatment for NSSI, but it did provide a promising future direction for research aimed at finding an effective treatment for NSSI.

Although the present findings involving aversive conditioning with painful electric shocks are promising, I do not recommend this “shock therapy” as a large-scale treatment for NSSI. Rather, I have two treatment recommendations inspired by the present results. First, consistent with the Benefits and Barriers Model of NSSI, treatments should focus on rebuilding *all* NSSI barriers, not just the aversion barrier. For example, treatments may additionally focus on rebuilding the pain barrier (e.g., Franklin et al., 2012; possibly by targeting self-criticism, see Hooley & St. Germain, in preparation), implicit identification barrier (present results; Nock & Banaji, 2007), social norms barrier (e.g., Prinstein et al., 2010), and many others that have not yet been elucidated by research.

Second, consistent with the recommendations of Kazdin and Blaise (2011), these barriers should be rebuilt with techniques that are not only effective (e.g., the present shock paradigm), but also available, affordable, and palatable to the majority of the public. Computerized evaluative conditioning paradigms may meet these criteria (see Hofmann et al., 2010). These flexible paradigms involve pairing target stimuli (e.g., NSSI image) with valenced stimuli (e.g., pleasant/unpleasant words or pictures). If these simple paradigms were converted into phone applications, millions of people would have access to this treatment mechanism; moreover, if these applications were placed within a game-like format, individuals may be intrinsically motivated to regularly self-administer these treatments. In sum, the present shock paradigm may not represent a viable large-scale treatment option, but it does provide a blueprint for how to develop potentially effective large-scale treatments for NSSI.

Limitations and Future Directions

The present study's findings should be interpreted in accord with its limitations, which establish directions for future research. The present sample was larger and more severe than any previous NSSI laboratory study, but a larger sample would have permitted more detailed analyses of follow-up data. Similarly, because self-cutting is the most common severe NSSI behavior (e.g., Nock & Prinstein, 2004), the present study focused on individuals who engaged in self-cutting and all attitude measures were specific to self-cutting. Although exploratory posthoc findings are consistent with the idea that self-cutting findings extend to other forms of NSSI (e.g., Table 3, Figure 12), larger empirical studies focused on non-cutting behaviors are required to support this hypothesis.

Second, the five attitude measures in the present study covered a range of processes, but by no means do these measures cover all aspects of attitudes toward NSSI. Other explicit (e.g., questionnaire about attitudes toward NSSI), implicit (e.g., the more emotion-based NSSI IAT, see Nock & Banaji, 2007), and psychophysiological measures (e.g., brain imaging, physiological arousal) may have yielded results that were unique from the measures included in the present study. Moreover, there were no completely passive attitude measures in the present study: all measures required self-report, performance, or the presentation of a loud sound. It could be argued that the active elements of these measures may have disrupted (or at least changed in some way) the affective experience induced by self-cutting images. Future studies may accordingly benefit from measuring attitudes toward NSSI with other measures, especially passive measures (e.g., cardiac output, total peripheral resistance).

Third, the order of measures was not randomized in the present study. Pilot testing in 25 participants revealed that participants tended to become inattentive when the affective valence startle modulation block (which took ~20 minutes and involved only 2-3 trials per minute) was contiguous with either the startle habituation block or the conditioning procedure. Similarly, some participants displayed a tendency to “press through” the implicit measures when they were administered at the very end of the study (after 2.5 to 3 hours of other procedures). Optimal data collection occurred (i.e., high attentiveness and effort on all measures) when measures were administered in the following order: AMP, IAT, affective-valence startle modulation, and explicit affective ratings. Accordingly, this was the measurement order for all participants. It is possible that this order affected results; however, the high degree of consistency of all measures pre- and post-random conditioning indicates that these measures were robust to the effects of a major experimental procedure that involved painful shocks. Correspondingly, these measures may have been robust to the potentially more minor effects of measurement order. Nonetheless, the present study could not evaluate the effects of measurement order on results. This emphasizes the need for replication of the present results in different experimental paradigms.

Fourth, the present results may not generalize to other procedures designed to change attitudes toward NSSI; however, the present findings establish a foundation for future studies aimed at altering these attitudes. The present 18-shock trial relief/aversive conditioning procedure was based on Andreatta et al. (2010), who included 16-shock trial conditioning procedures to alter startle eyeblink reactivity to neutral stimuli. Consistent with findings that it is harder to change an attitude than to form an attitude (Cacioppo et al., 1992; Hofmann et al., 2010), even with the inclusion of two extra conditioning trials, the startle eyeblink effect sizes in the present study (aversive conditioning: d = .55; relief conditioning d = .42) were smaller than those of Andreatta and colleagues (aversive conditioning d = .92; relief conditioning d = .61). This implies that procedures that use fewer trials and/or less intense shocks are unlikely to see NSSI attitude changes, but procedures that use many more trials and/or more intense shocks may see large shifts in NSSI attitudes. Similarly, given that shocks generates the largest effects on attitude formation/change (Hofmann et al., 2010), traditional picture- or word-based evaluative conditioning procedures may not alter NSSI attitudes unless they are administered several times or include a high number of NSSI-relevant trials (e.g., over 100). Future studies should build on the present findings to create treatments for NSSI that are both effective and able to be delivered on a large scale.

Fifth, the random conditioning procedure employed in the present study may have obscured some results. To provide a stringent experimental control from the aversive and pain offset relief conditioning procedures, it was necessary to present the same pictures and number of shocks to all participants who did not undergo these procedures. However, it would have been inadvisable for the presentation of shocks in this “random conditioning” procedure to occur at truly random intervals during the presentation of pictures. This is because each shock would still occur either a few seconds before or after (or during) pictures, meaning that each shock would still have the effect of producing either pain offset relief conditioning or aversive conditioning of picture stimuli. With this method, inevitably some “random conditioning” participants would have actually received a conditioning procedure that was essentially aversive or pain offset relief conditioning.

To control for this possibility across all assigned to random conditioning participants, each participant received an equal number of pain offset relief and aversive conditioning trials. Unfortunately, given that these two forms of conditioning are not equally powerful (Andreatta et al., [2010] and the present results both indicate that aversive conditioning is more powerful) the random conditioning procedure may have produced a light form of aversive conditioning. This may have obscured some aversive conditioning effects and may help to explain why the NSSI-control group displayed a reduction in self-cutting behaviors in the six months after the lab visit (albeit a smaller reduction that the NSSI-aversive group displayed). Future studies may benefit from including a range of control groups (e.g., no shock, fewer shocks, etc.) to more effectively isolate the effects of aversive conditioning on attitudes toward NSSI.

Sixth, a related limitation is that all participants may have received a non-shock form of aversive conditioning during the two affective-valence startle modulation blocks. Many studies use quick, loud sounds as aversive stimuli; correspondingly, the 100dB 50ms white noise bursts used throughout the startle modulation blocks may have represented a form of aversive conditioning. If this is true, then all participants received 24 sound-based aversive conditioning pairings with self-cutting images (12 per startle modulation block). However, any conditioning effects may have been mitigated by the fact that these sounds were paired with all types of pictures and occurred during picture (rather than immediately after). Nonetheless, this possible sound-based aversive conditioning may help to explain the significant prospective reduction in self-cutting behaviors in the NSSI-control group. Future investigations may benefit from examining the effects of aversive conditioning without including startle-based measures.

Seventh and finally, the regression techniques employed to analyze the longitudinal data were not optimal. More advanced strategies such as multilevel modeling may have been more powerful and effective. This is particularly true for analyses examining the effects of conditioning on NSSI six months after the lab visit. Given the nested nature of these data and the strict assumptions of regression techniques, multilevel modeling may have provided a clearer analysis. Accordingly, in the future I plan to re-examine these data with more advanced statistical techniques.

Conclusion

NSSI is a deadly behavior that affects millions of people each year. Despite a recent surge in research, many crucial aspects of NSSI remain poorly understood. These gaps in knowledge have obstructed the understanding, assessment, prediction, and treatment of NSSI. The present study attempted to improve this state of affairs with an experimental, psychophysiological, and longitudinal investigation of the nature, development, and reversal of positive attitudes toward NSSI. Results revealed that people who engage in self-cutting possess positive attitudes toward self-cutting across a range of measures, some of which are robust predictors of future self-cutting; pain offset relief conditioning may partially explain the development of these attitudes; and aversive conditioning may reverse these attitudes and reduce future self-cutting behaviors. These findings prompted a new model of NSSI – the Benefits and Barriers Model – that has the potential to stimulate research that eventually produces effective large-scale treatments for NSSI.

Table 1. *Correlations among attitudes, ERS/DERS, and NSSI recency/frequency*.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. AMP	-														
2. Explicit	.42^	-													
3. IAT	.10	.21**	-												
4.Postauricular	.17*	-.13	.08	-											
5. Eyeblink	-.14*	-.18*	-.14*	-.05	-										
6. DERS	.17*	.24**	.13	.12	-.14*	-									
7. ERS	.23*	.28^	.10	.07	-.03	.72^	-								
8. Recency	-.23*	-.30**	-.13	.08	-.41**	-.33**	-.15	-							
9. Cut Freq – Lifetime	.34**	.41^	.04	-.30	-.04	.05	.21	-.08	-						
10. Cut Freq – Last year	.23*	.34**	.03	-.18	-.02	-.05	-.18	-.11	.54^	-					
11. Cut Freq – Past 6 mo.	.22	.18	.10	-.01	-.13	.23*	-.08	-.16	-.04	.84^	-				
12. Cut Freq- Past mo.	.36**	.37^	.08	-.27	-.03	-.01	-.09	-.18	.77^	.84^	.10	-			
13. Cut Freq – Future 6 mo.	.34**	.32**	.14	.07	-.10	.08	.00	-.14	.04	.60^	.60^	.07	-		
14. All NSSI Freq-6mo. Pre	.46^	.35**	-.06	-.03	-.29*	.35*	.16	-2.1	.12	.69^	.68^	.21	.59^	-	
15. All NSSI Freq – 6 mo. Post	.49^	.44**	.04	.06	-.30*	.13	.06	-.17	.06	.51^	.37^	.15	.83^	.69^	-

Note. * = *p*<. 05; ** = *p*<.01; ^ = *p*<.001; Correlations among baseline measures include the entire sample, but correlations with NSSI recency/frequency only include the NSSI group (n = 58); freq = raw frequency; 'All NSSI freq' = Composite NSSI variable that was the sum of all reported NSSI behaviors (cutting, burning, scraping, hitting, etc.).

Table 2. *Hierarchical multiple regression analyses predicting self-cutting frequency during the six months post-lab visit.*

Step and Variables	$\Delta F(p)$	ΔR^2	*b* (SE)	*t*(*p*)
Step 1	10.53 (.001)***	.19		
Self-Cutting Frequency 6 mo. Pre-Lab			.39 (.12)	3.25 (.001)***
Step 2	3.47 (.04)*	.12		
AMP Scores			9.84 (6.82)	1.44 (.16)
Explicit Affective Ratings			1.13 (1.05)	1.07 (.29)

Note. Self-cutting frequency variables were transformed into rank-order scores; * = $p<.05$; *** = $p<.001$

Table 3. *Hierarchical multiple regression analyses predicting composite NSSI frequency during the six months post-lab visit.*

Step and Variables	$\Delta F(p)$	ΔR^2	*b*(SE)	*t*(*p*)
Step 1	27.91(.000)***	.40		
Composite NSSI Frequency 6 mo. Pre-Lab			.60 (.11)	5.28 (.000)***
Step 2	3.80 (.02)*	.14		
AMP Scores			16.94 (6.82)	2.79 (.009)**
Explicit Affective Ratings			-1.26 (.95)	-1.33 (.16)
Startle Eyeblink Reactivity			-7.76 (.4.31)	-1.78 (.08)

Note. Self-cutting frequency variables were transformed into rank-order scores; * = $p<.05$; *** = $p<.001$; Composite NSSI variables were the sum of all reported NSSI behaviors (cutting, burning, scraping, hitting, etc.)

Figure 1. *Model of primary and opponent processes presented separately (top) and combined (bottom).*

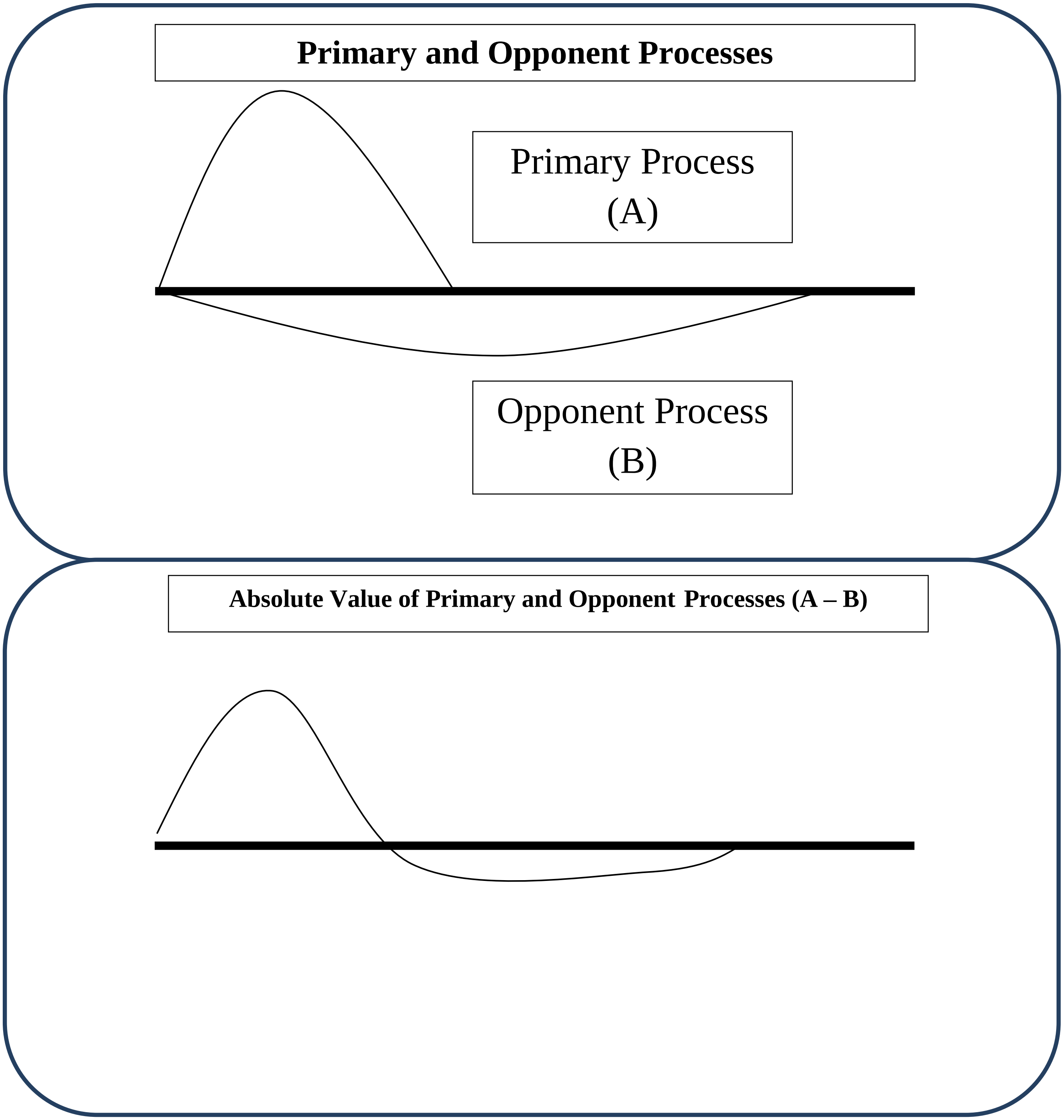


Figure 2. *Overview of Experimental Procedure.*

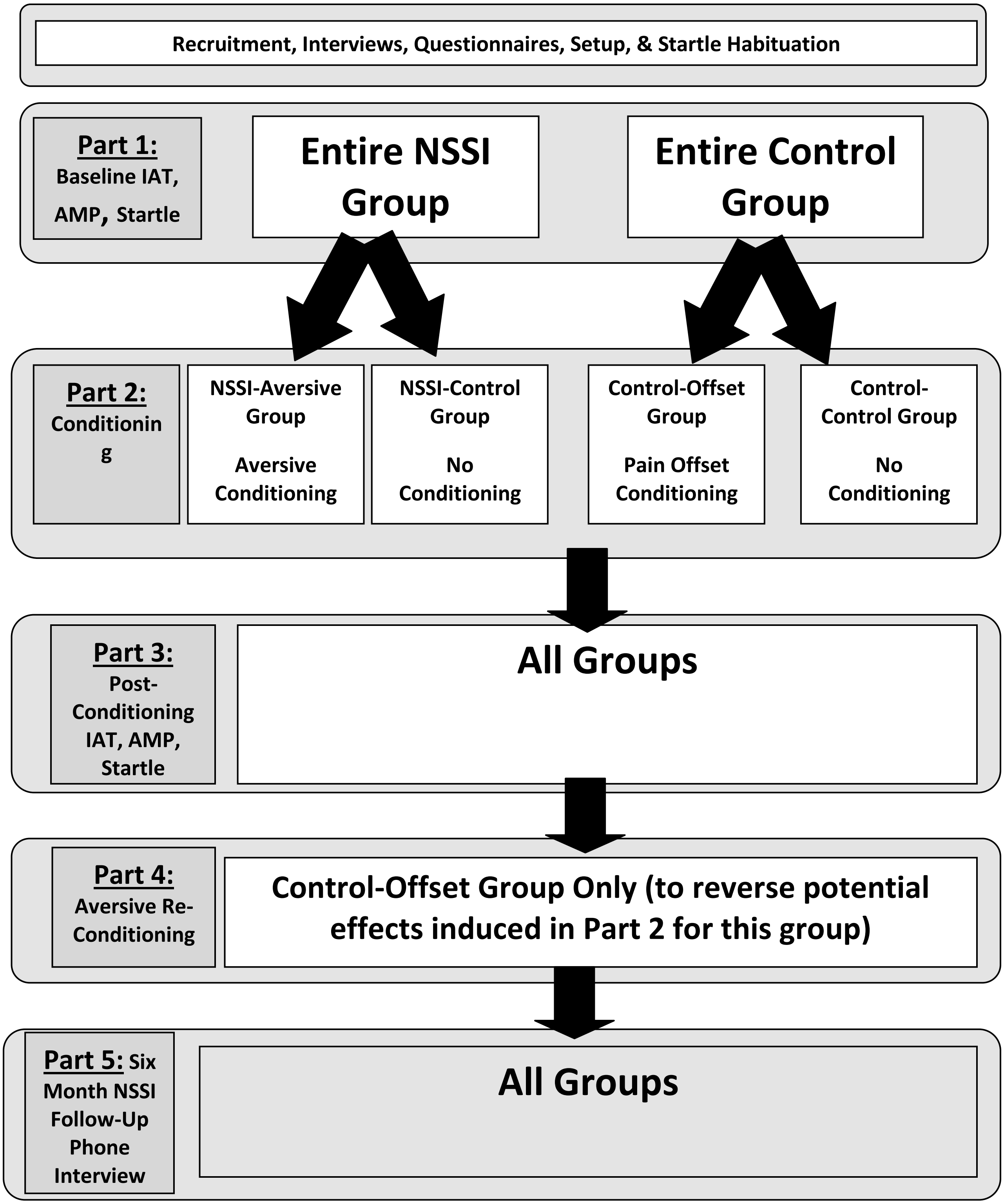
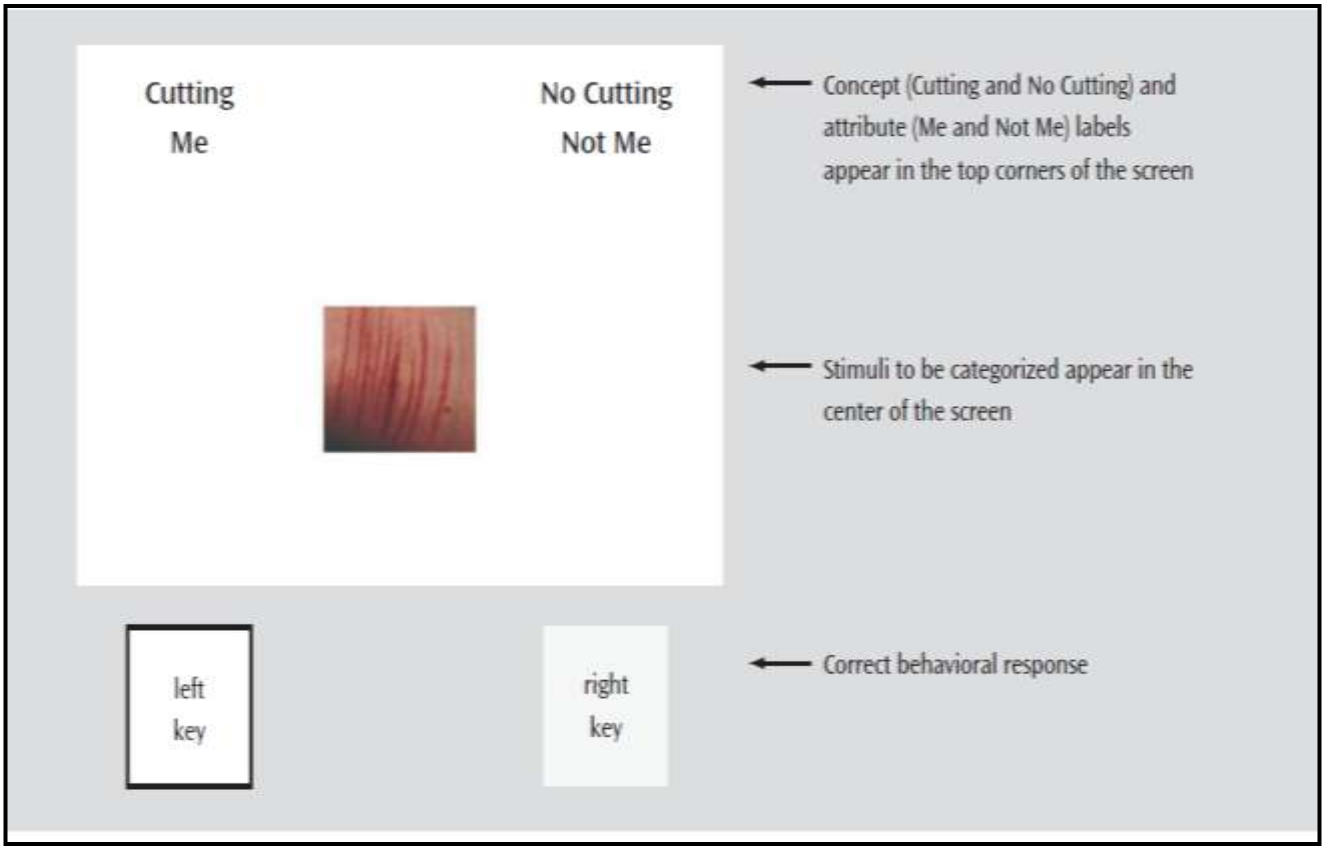
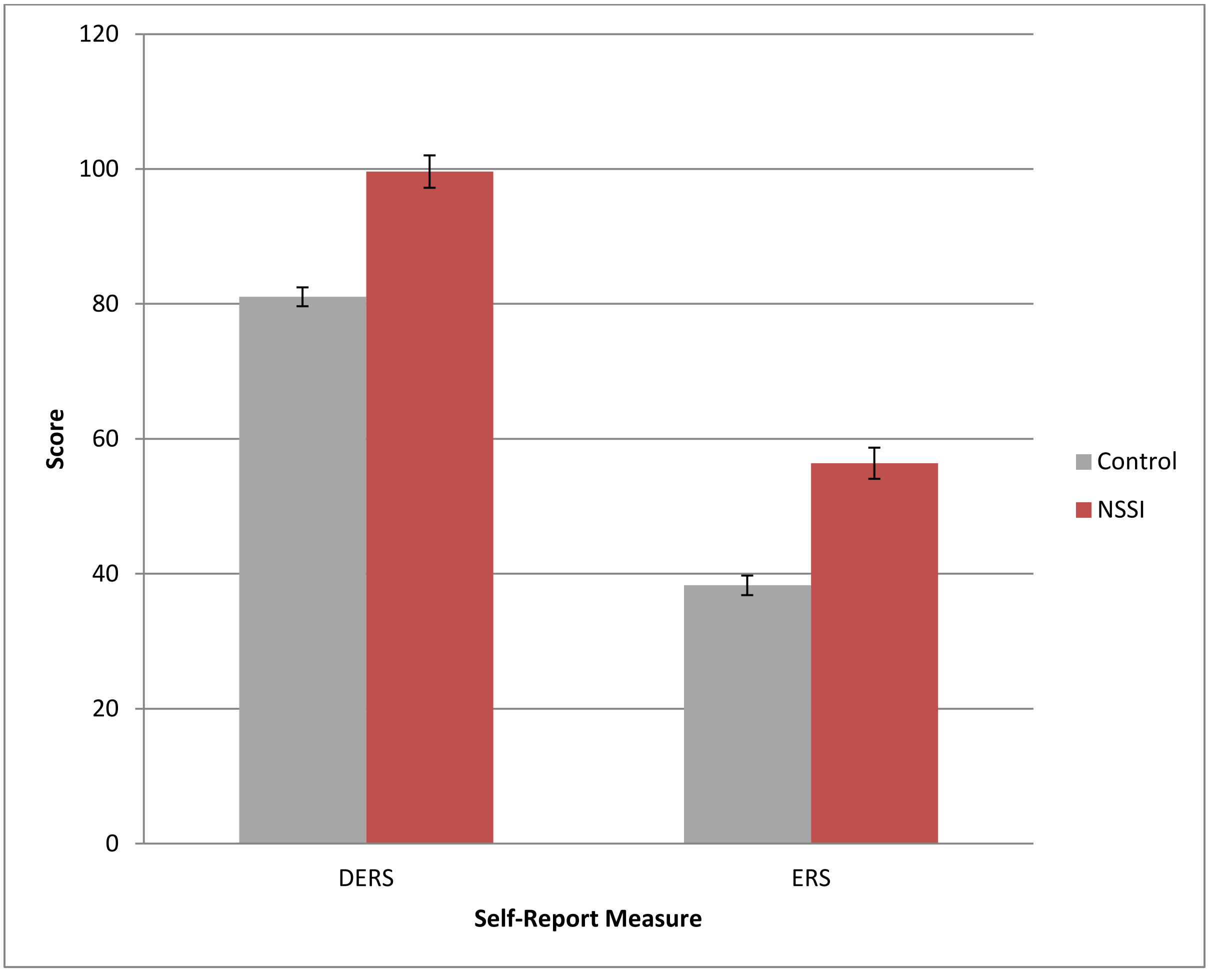


Figure 3. *Example of screen during the NSSI version of the IAT*.



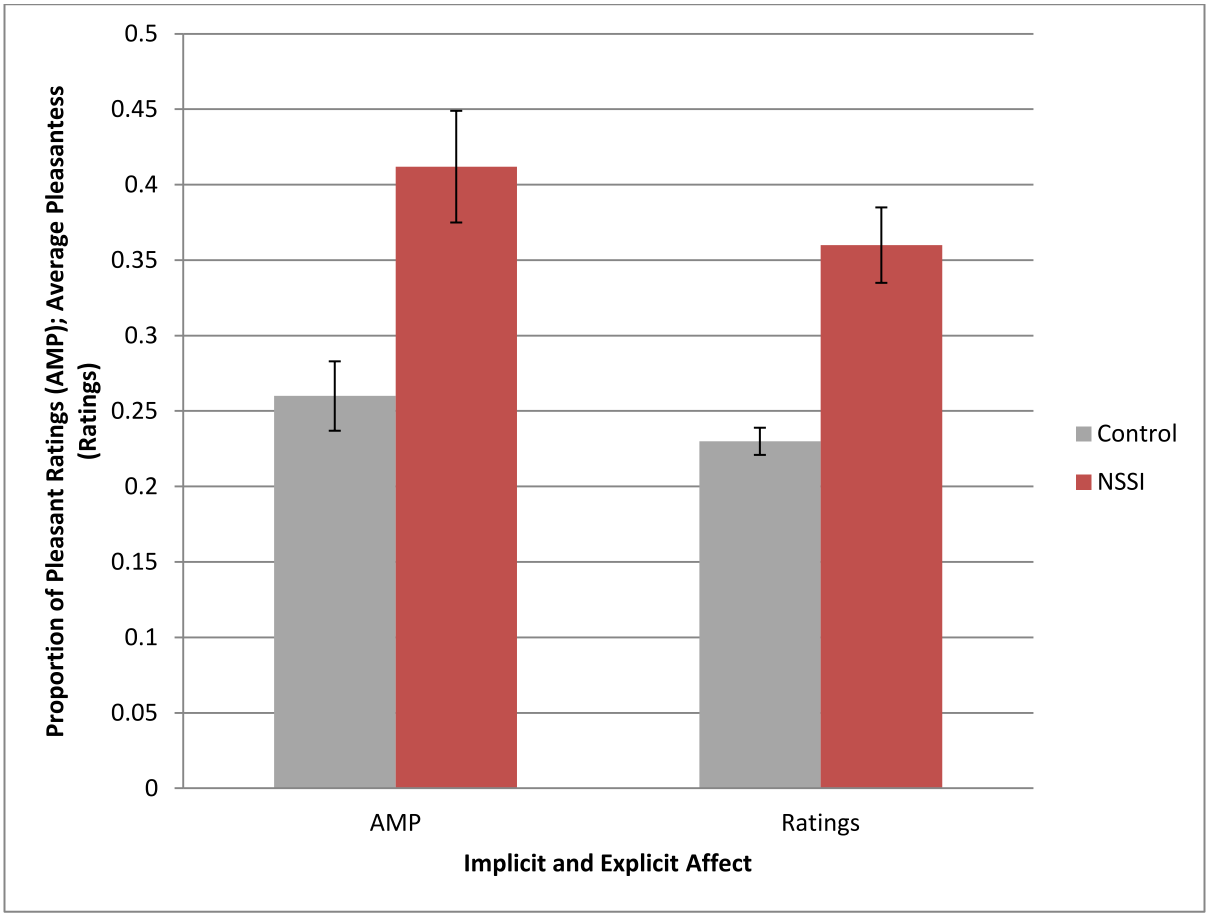
Note. Figure adapted from Nock and Banaji (2007) in *American Journal of Psychiatry*.

Figure 4. *Mean DERS (emotion dysregulation) and ERS (emotion reactivity) scores by group.*



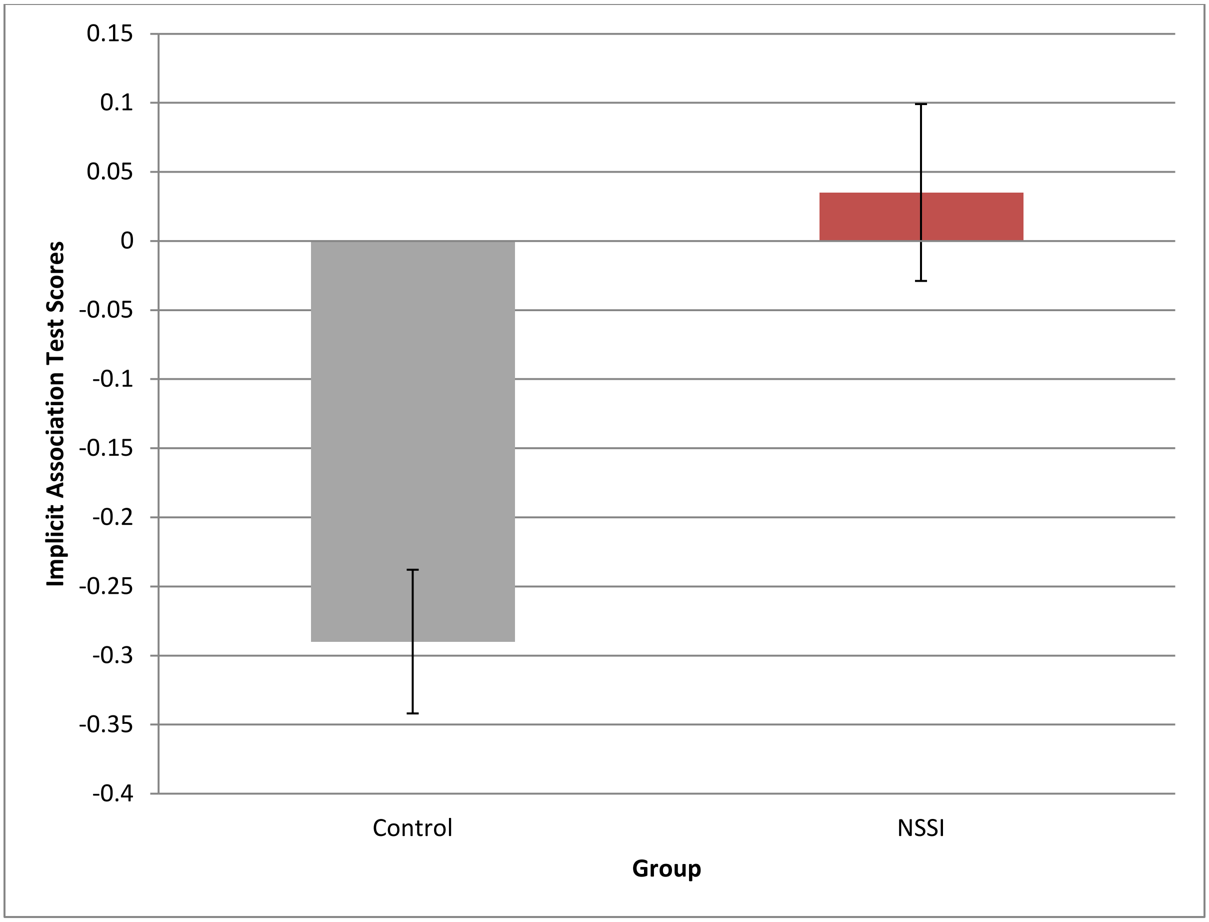
Note. Errors bars represent +/- 1SE. The NSSI group displayed higher mean DERS (M = 99.47; SD = 19.35) and ERS (M = 56.40; SD = 17.61) scores compared to the DERS (M = 81.07; SD = 13.01) and ERS (M = 38.29; SD = 13.49) scores of the control group.

Figure 5. *AMP scores (performance-based implicit affect) and explicit affective ratings of NSSI images by group.*



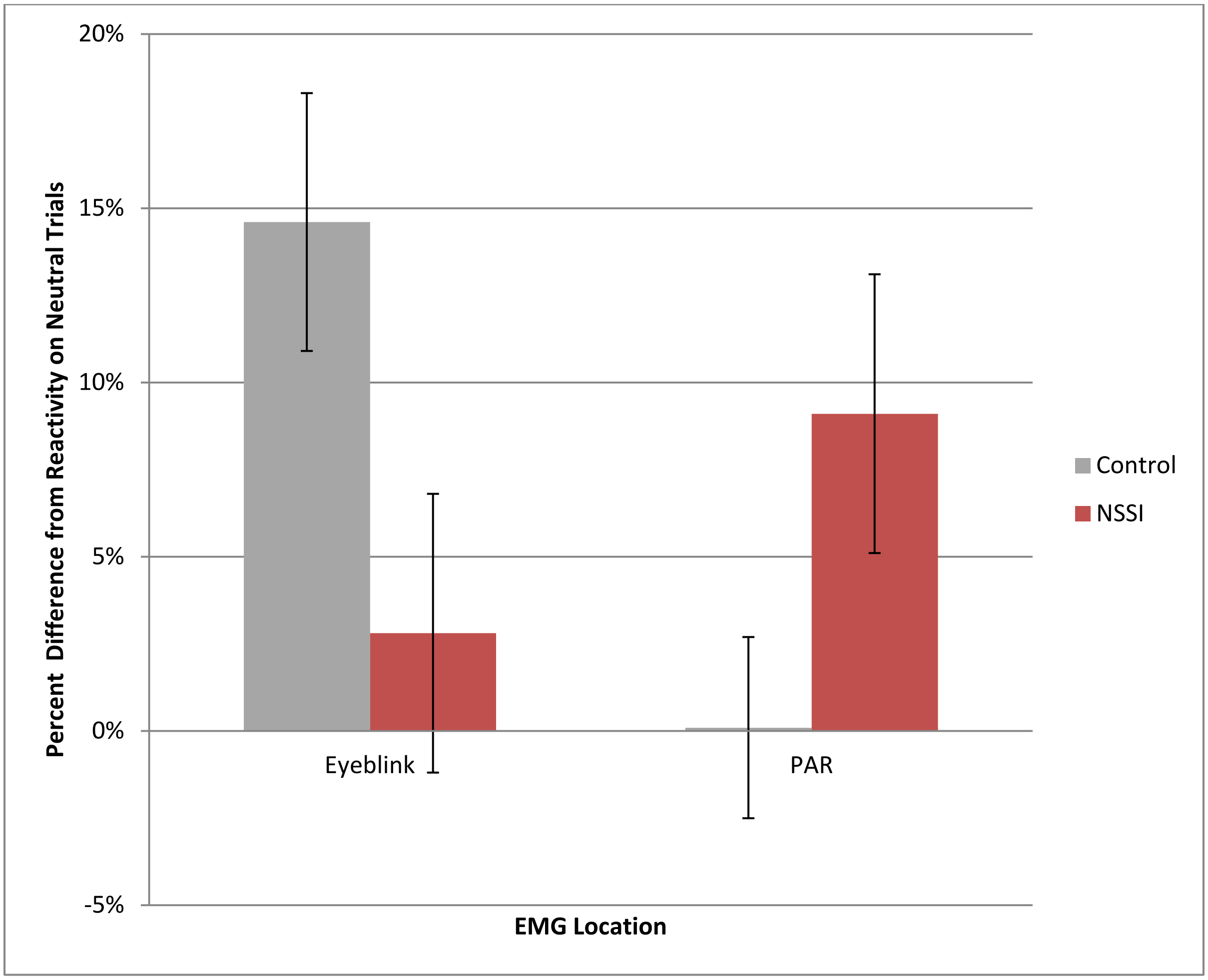
Note. Error bars represent +/- 1SE. The NSSI group displayed higher AMP (M = .41; SD = .28) and explicit affective ratings (M = .36; SD = .20) scores compared to the AMP (M = .26; SD = .21) and explicit affective ratings (M = .23; SD = .09).

Figure 6. *IAT scores (implicit associations between the self and self-cutting) by group.*



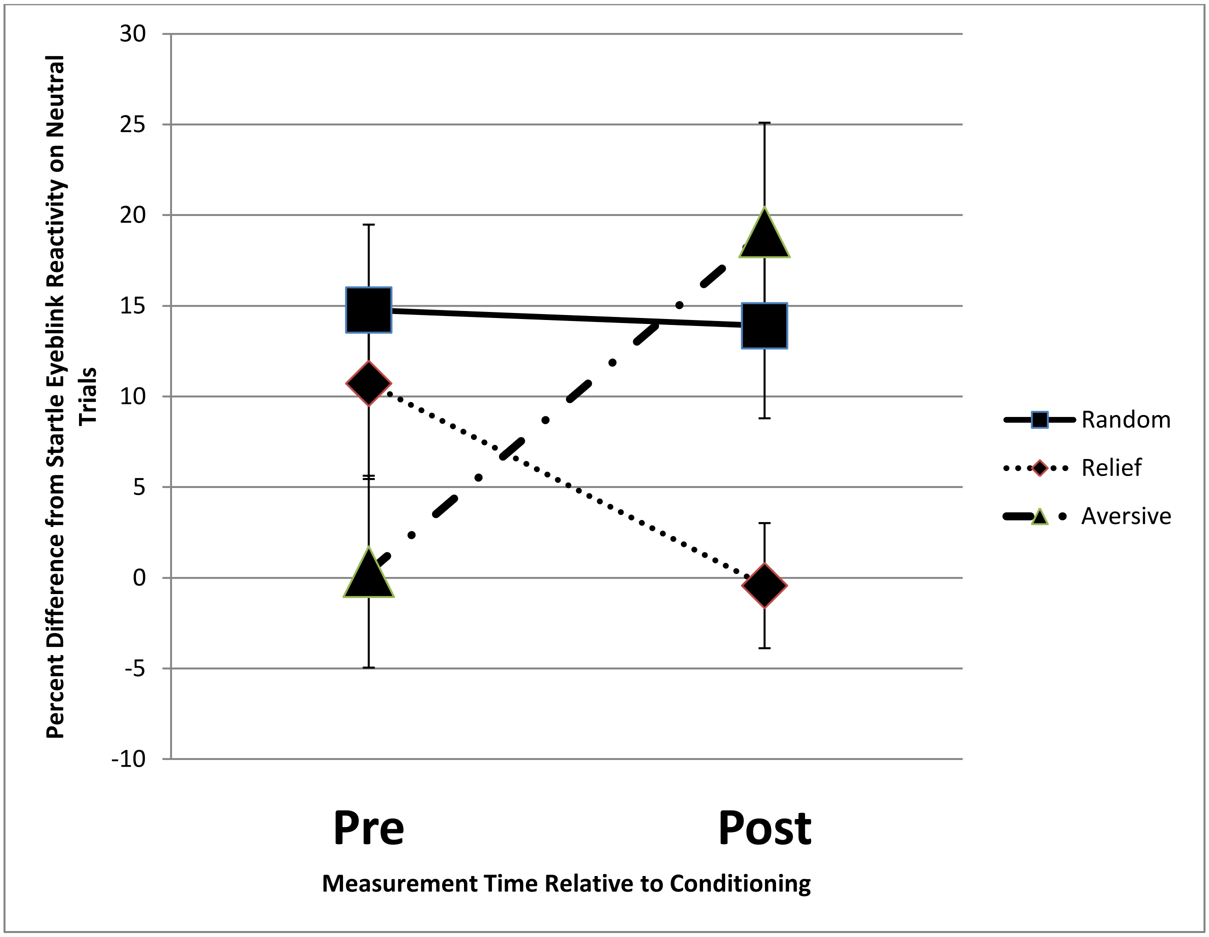
Note. Errors bars represent +/- 1SE. The implicit identification with self-cutting was more positive for the NSSI group (M = .04; SD = .48) than for the control group (M = -29; SD = .48).

Figure 7. *Startle eyeblink (defensive motivation) and postauricular reactivity (appetitive motivation) by group.*



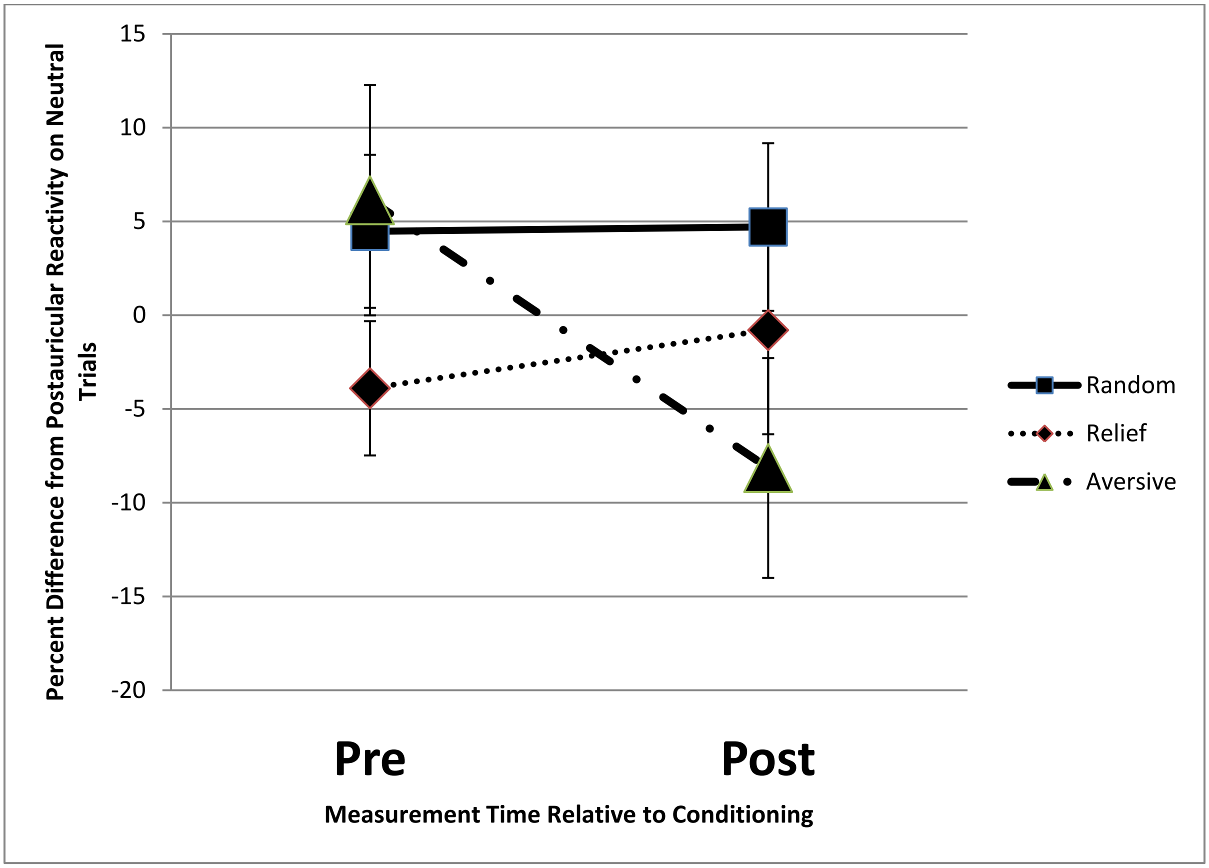
Note. Error bars represent +/- 1SE. In the context of self-cutting images, startle eyeblink reactivity was lower for the NSSI group (M = 2.83%; SD = 34.59%) compared to the control group (M = 14.61%; SD = 34.34%). Conversely, postauricular reactivity was higher for the NSSI group (M = 9.15%; SD = 37.90%) compared to the control group (M = 0.01%; SD = 24.28%).

Figure 8. *Startle eyeblink reactivity pre- and post-conditioning by conditioning type.*



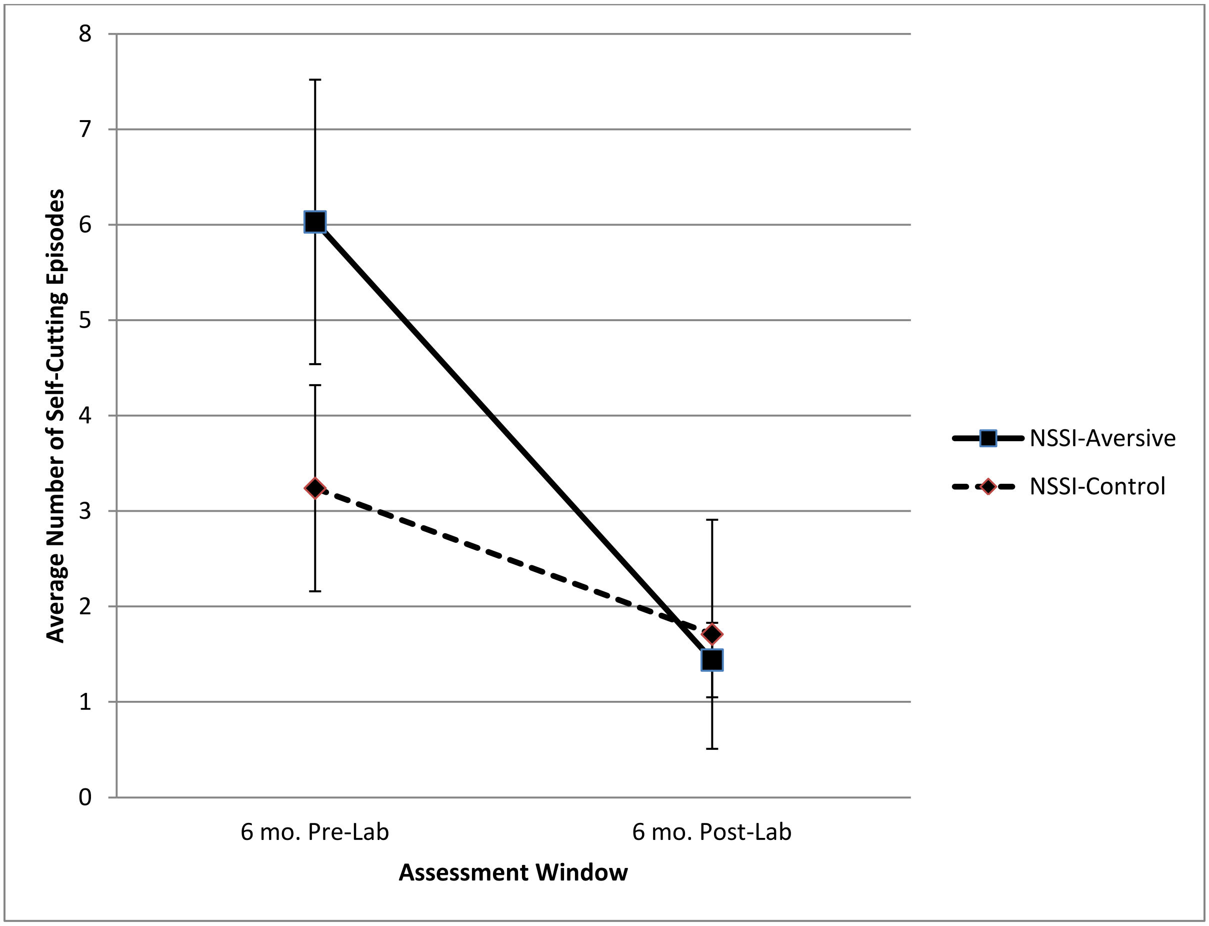
Note. Error bars represent +/- 1SE mean. In the context of self-cutting images, startle eyeblink reactivity of the NSSI-aversive (i.e., Aversive) group increased from pre- (M = 2.04%; SD = 31.56%) to post-conditioning (M = 21.49%; SD = 36.86%); the reactivity of the Control-offset (i.e., Relief) group diminished from pre- (M = 10.72%; SD = 31.64%) to post-conditioning (M = -0.43%; SD = 20.70%); and the reactivity of the NSSI-control and control-control (i.e., Random) groups remained steady from pre- (M = 14.78%; SD = 36.52%) to post-conditioning (M = 13.90%; SD = 39. 59%).

Figure 9. *Postauricular reactivity pre- and post-conditioning by conditioning type.*



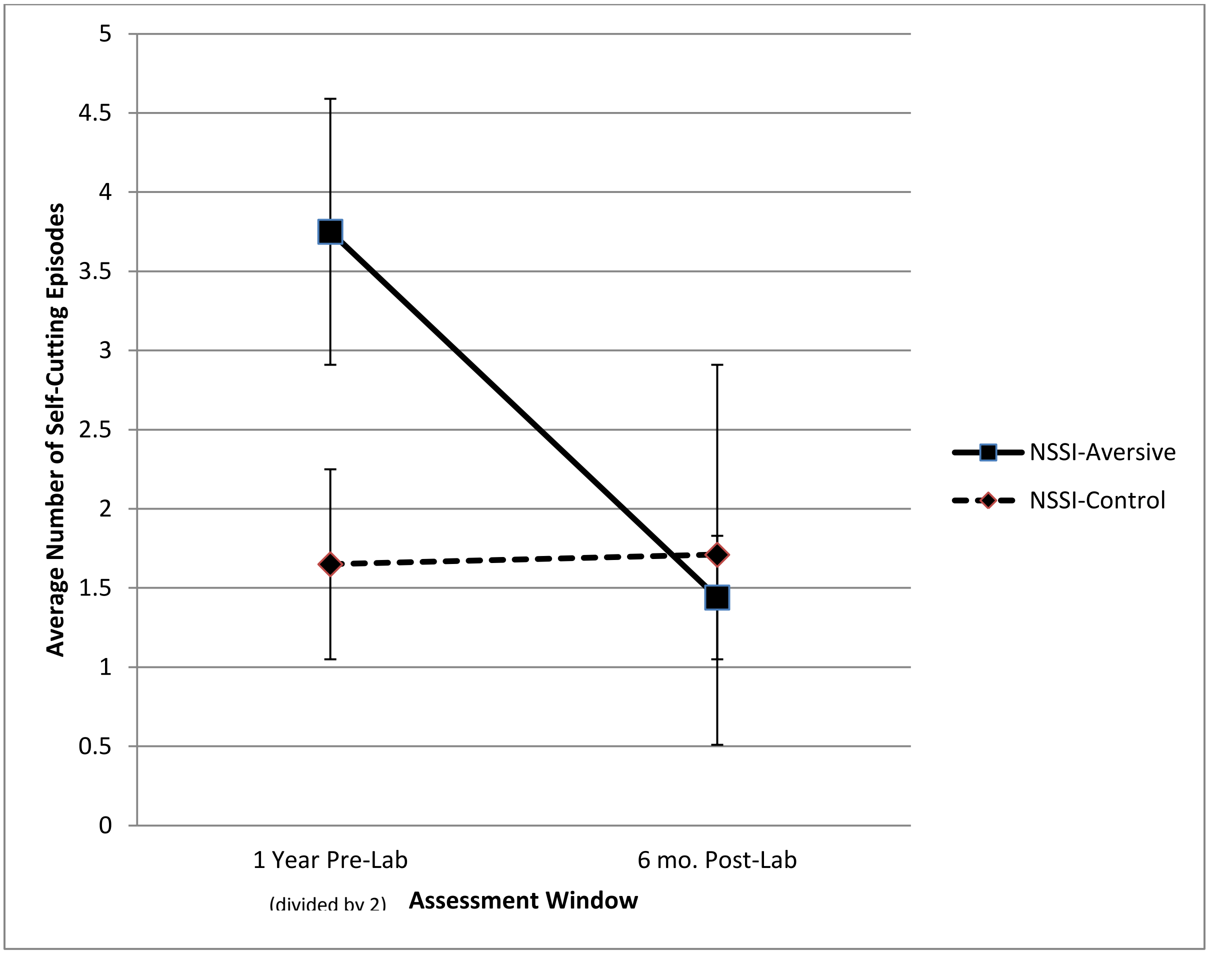
Note. Error bars represent +/- 1SE. Postauricular reactivity of the NSSI-aversive (i.e., Aversive) group diminished from pre- (M = 6.13%; SD = 33.60%) to post-conditioning (M = -8.14%; SD = 32.05%); the reactivity of the Control-offset (i.e., Relief) group increased slightly from pre- (M = -3.94%; SD = 22.04%) to post-conditioning (M = -0.90%; SD = 34.21%); and the reactivity of the NSSI-control and control-control (i.e., Random) groups remained steady from pre- (M = 4.47%; SD = 32.10%) to post-conditioning (M = 4.70%; SD = 33.44%).

Figure 10. *Self-cutting frequency (in six month windows) pre- and post-lab visit by conditioning type in the NSSI group.*



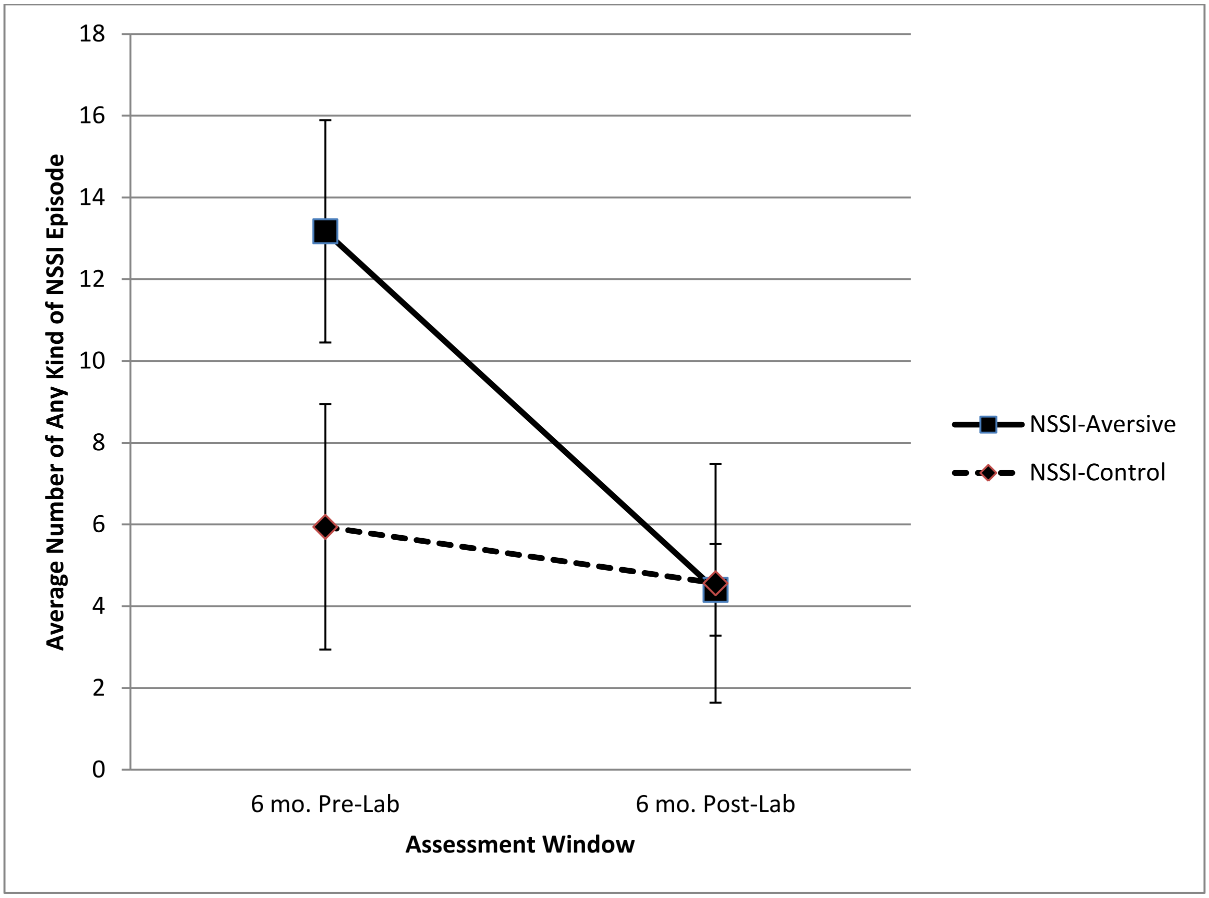
Note. Error bars represent +/- 1SE. The self-cutting rates of the NSSI-aversive group declined sharply from the 6 months before the lab visit (M = 6.03; SD = 8.45) to the 6 months after the lab visit (M = 1.44; SD = 2.26). Similarly, the NSSI-control group declined slightly from the 6 months before the lab visit (M = 3.24; SD = 4.76) to the 6 months after the lab visit (M = 1.71; SD = 4.95).

Figure 11. *Self-cutting frequency six months pre- and six months post-lab visit by conditioning type in the NSSI group.*



Note. Six months pre-visit frequency was calculated by dividing self-cutting frequency one-year prior to lab visit by two (to control for the greater length of the pre-visit assessment window). Error bars represent +/- 1SE. The self-cutting rates of the NSSI-aversive group declined sharply from pre-visit (M = 3.75; SD = 4.74) to post-visit (M = 1.44; SD = 2.26). Similarly, the NSSI-control group declined slightly from pre-visit (M = 1.65; SD = 2.47) to post-visit (M = 1.71; SD = 4.95).

Figure 12. *NSSI frequency (in six month windows) pre- and post-lab visit by conditioning type in the NSSI group.*



Note. Error bars represent +/- 1SE. The overall NSSI rates of the NSSI-aversive group declined sharply from the 6 months before the lab visit (M = 13.17; SD = 14.93) to the 6 months after the lab visit (M = 4.40; SD = 6.13). Similarly, the NSSI-control group declined slightly from the 6 months before the lab visit (M = 5.94; SD = 12.00) to the 6 months after the lab visit (M = 4.56; SD = 11.67).

REFERENCES

Andover, M.S., & Gibb, B.E. (2010). Non-suicidal self-injury, attempted suicide, and suicidal intent among psychiatric inpatients. *Psychiatry Research, 178,* 101-105.

Andreatta, M., Muhlberger, A., Yarali, A., Gerber, B., & Pauli, P. (2010). A rift between implicit and explicit conditioned valence in human pain relief learning. *Proceedings of the Royal Society B, 277*, 2411-2416.

Asarnow, J.R., Porta, G., Spirito, A., Emslie, G., Clarke, G.,… Brent, D.A. (2011). Suicide attempts and nonsuicidal self-injury in the treatment of resistant depression in adolescents: Findings from the TORDIA study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *50*, 772-781.

Barrett, L.F. (2012). Emotions are real. *Emotion, 12,* 413-429.

Barrett, L.F., & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology, 41*, 167-218.

Becerra, L. & Borsook, D. (2008). Signal valence in the nucleus accumbens to pain onset and offset. *European Journal of Pain, 12*, 866-869.

Benning, S. D., Patrick, C. J., Lang, A. R. (2004). Emotional modulation of the post-auricular reflex. *Psychophysiology, 41,* 426-432.

Blumenthal, T.D., Cuthbert, B.N., Filion, D.L., Hackley, S., Lipp O.V., & van Boxtel, A. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology, 42,* 1-15.

Bradley, M.M., Cuthbert, B.N., & Lang, P.J. (1999). Affect and the startle reflex. In M.E. Dawson, A.M. Schell, A.H. Bohmelt (Eds.), *Startle modification: Implications for neuroscience, cognitive science, and clinical science* (pp. 157-183). Cambridge, UK: Cambridge University Press.

Bradley, M.M., Codispoti, M., Cuthbert, B.N., & Lang, P.J. (2001). Emotion and Motivation I: Defensive and appetitive reactions in picture processing. *Emotion, 1,* 276-298.

Bresin, K., Gordon, K.H., Bender, T.W., Gordon, L.J., & Joiner, T.E. (2010). No pain, no change: Reductions in prior negative affect following physical pain. *Motivation and Emotion, 34,* 280-287.

Cacioppo, J.T., Marshall-Goodell, B.S., Tassinary, L.G., & Petty, R.E. (1992). Rudimentary determinants of attitudes: Classical conditioning is more effective when prior knowledge about the attitude stimulus is low than high. *Journal of Experimental Social Psychology, 28*, 207-233.

Carver, C.S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin, 135,* 183-204.

Darche, M. A. (1990). Psychological factors differentiating self-mutilating and non-self mutilating adolescent inpatient females. *Psychiatric Hospital, 21*, 31-35.

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Association learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127,* 853-869.

Derbyshire, S.W.G., & Osborn, J. (2008). Enhancement of offset analgesia during sequential testing. *European Journal of Pain, 12,* 980-989.

Derbyshire, S.W.G., & Osborn, J. (2009). Offset analgesia is mediated by activation in the region of the periaquedectual grey and rostral ventromedial medulla. *Neuroimage, 47,* 1002-1006.

Dichter, G.S., Benning, S.D., Holtzclaw, T.N., & Bodfish, J.W. (2010). Affective modulation of the startle eyeblink and postauricular reflexes in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *40*, 858-869.

DiClemente, R. J., Ponton, L. E., & Hartley, D. (1991). Prevalence and correlates of cutting behavior: Risk for HIV transmission. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 735-739.

Dijksterhuis, A. (2004). I like myself but I don't know why: Enhancing implicit self-esteem with subliminal evaluative conditioning. *Journal of Personality and Social Psychology, 86*, 345-355.

Franklin, J.C., Aaron, R.V., Arthur, M.S., Shorkey, S.P., & Prinstein, M.J. (2012). Nonsuicidal self-injury and diminished pain perception: The role of emotion dysregulation. *Comprehensive Psychiatry*. *53*, 691-700.

Franklin, J.C., Hessel, E.T., Aaron, R.V., Arthur, M.S., Heilbron, N., & Prinstein, M.J. (2010). The function of nonsuicidal self-injury: Support for cognitive-affective regulation and opponent processes from a novel psychophysiological paradigm. *Journal of Abnormal Psychology, 119*, 850-862.

Franklin, J.C., Hessel, E.T., & Prinstein, M.J. (2011). Clarifying the role of pain tolerance in suicidal capability. *Psychiatry Research*, *189,* 362-367.

Franklin, J.C., Lee, K.M., Hanna, E.K, & Prinstein, M.J. (in press). Feeling worse to feel better: Pain offset simultaneously stimulates positive affect and diminishes negative affect. *Psychological Science.*

Franklin, J.C., Puzia, M.E., Lee, K.M., Lee, G.E., Hanna, E.K., Spring. V.L., & Prinstein, M.J. (2013). The nature of pain offset relief in nonsuicidal self-injury: A laboratory study. *Clinical Psychological Science*. Published online February 26, 2013 as *doi:10.1177/2167702612474440*

Gable, P.A., & Harmon-Jones, E. (2009). Postauricular reflex responses to pictures varying in valence and arousal. *Psychophysiology, 46,* 487-490.

Gawronski, B. & Bodenhausen, G.V. (2006). Association and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*, 692-731.

Gawronski, B., & Bodenhausen, G.V. (2011). The associative-propositional evaluation model: Theory, evidence, and open questions. *Advances in Experimental Social Psychology, 44*, 59-124.

Glenn, C.R., & Klonsky, E.D. (2010). The role of seeing blood in non-suicidal self-injury. *Journal of Clinical Psychology, 66*, 466-473.

Glenn, C.R., & Klonsky, E.D. (2011). Prospective prediction of nonsuicidal self-injury: A 1-year longitudinal study in young adults. *Behavior Therapy*, *42,* 751-762.

Gratz, K.L. & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment, 26*, 41-54.

Greenwald, A.G., McGhee, D.E., & Schwartz, J.L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74,* 1464-1480.

Greenwald, A.G., Nosek, B.A., & Banaji, M.R. (2003). Understanding and using the Implicit Association Test, I: an improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197-216.

Greenwald, A.G., Poehlman, T.A., Uhlmann, E.L., & Banaji, M.R. (2009). Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 87*, 17-41.

Grill, J.D., & Coghill, R.C. (2002). Transient analgesia evoked by noxious stimulus offset. *Journal of Neurophysiology, 87,* 2205-2208.

Guan, K., Fox, K.R., & Prinstein, M.J. (2012). Nonsuicidal self-injury longitudinally predicts adolescent suicide ideation and attempts in a diverse community sample. *Journal of Consulting and Clinical Psychology, 80*, 842-849.

Guerry, J.D., & Prinstein, M.J. (2010). Longitudinal prediction of adolescent nonsuicidal self-injury: examination of a cognitive vulnerability-stress model. *Journal of Clinical Child and Adolescent Psychology, 39*, 77-89.

Hackley, S.A., Munoz, M.A., Hebert, K., Valle-Inclan, F., & Vila, J. (2009). Reciprocal modulation of eye-blink and pinna-flexion components of startle during reward anticipation. *Psychophysiology, 46*, 1154-1159.

Hankin, B.L., & Abela, J.R.Z. (2011). Nonsuicidal self-injury in adolescence: Prospective rates and risk factors in a two and a half year longitudinal study. *Psychiatry Research, 186*, 65-70.

Hazell, P.L., Martin G., McGill, K., Kay, T., Wood, A., et al. (2009). Group therapy for repated deliberate self-harm in adolescents: failure of replication of a randomized trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*, 662-670.

Hofmann, W., & Baumert, A. (2010). Immediate affect as a basis for intuitive moral judgment: An adaptation of the affect misattribution procedure. *Cognition and Emotion, 24*, 522-535.

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin, 136*, 390-421.

Hollands, G.J., Prestwich, A., Marteau, T.M. (2011). Using aversive images to enhance healthy food choices and implicit attitudes: An experimental test of evaluative conditioning. *Health Psychology, 30*, 195-203.

Hooley, J.M., Ho, D.T., Slater, J., & Lockshin, A. (2010). Pain perception and nonsuicidal self-injury: A laboratory investigation. *Personality Disorders: Theory, Research, and Treatment, 1*, 170-179.

Hooley, J.M., & St. Germain. (in preparation). Experimentally reducing self-criticism normalizes pain tolerance levels in individuals with a history of nonsuicidal self-injury.

Houben, K., Schoenmakers, T.M., & Wiers, R.W. (2010). I didn't feel like drinking but I don't know why: The effects of evaluative conditioning on alcohol-related attitudes, craving, and behavior. *Addictive Behaviors, 35*, 1161-1163.

Imhoff, R., & Banse, R. (2009). Ongoing victim suffering increases prejudice: The case of secondary anti-semitism. *Psychological Science, 20*, 1443-1447.

International Association for the Study of Pain Task Force on Taxonomy. (1994). *Classification of chronic pain syndromes and definition of pain terms* (2nd ed.) IASP Press.

Ito, T.A., & Cacioppo, J.T. (2007). Attitudes as mental and neural states of readiness: Using physiological measures to study implicit attitudes. In Wittenbrink & Schwarz (Eds.), *Implicit Measures of Attitudes*, 125-158.

Jacobson, C.M., & Gould, M. (2007). The epidemiology and phenomenology of non-suicidal self-injurious behavior among adelescents: A criticial review of the literature. *Archives of Suicide Research, 11*, 129-147.

Janis, I. B., Nock, M. K. (2009). Are self-injurers impulsive?: Results from two behavioral laboratory studies. *Psychiatry Research, 169,* 261-267.

Johnson, G.M., Valle-Inclan, F., Geary, D.C., & Hackley, S.A. (2012). The nursing hypothesis: An evolutionary account of emotional modulation of the postauricular reflex. *Psychophysiology, 49*, 178-185.

Joiner, T. (2005). *Why people die by suicide*. Harvard University Press: Cambridge, MA.

Joiner, T.E., Conwell, Y., Fitzpatrick, K.K., Witte, T.K., Schmidt, N.B., Berlim, M.T., Fleck, M.P., & Rudd, M.D. (2005). Four studies on how past and current suicidality relate even when "everything but the kitchen sink" is covaried. *Journal of Abnormal Psychology, 114*, 291-303.

Kazdin, A.E., Blasé, S.L. (2011). Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspectives on Psychological Science, 6*, 21-37.

Klonsky, E. D. (2007). The functions of deliberate self-injury: A review of the evidence. *Clinical Psychology Review, 27,* 226-239.

Klonsky, E.D. (2011). Non-suicidal self-injury in United States adults: prevalence, sociodemographics, topography and functions. *Psychological Medicine, 41*, 1981-1986.

Lane, S.T., Franklin, J.C., Curran, P. (in press). Clarifying the nature of startle habituation with latent growth curve modeling. *International Journal of Psychophysiology.*

Lang, P.J. (2010). Emotion and motivation: Toward consensus definitions and a common research purpose. *Emotion Review, 2,* 229-233.

Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (1990). Emotion, attention, and the startle reflex. *Psychological Review, 97,* 377-395.

Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Technical Report A-6. Gainesville, FL: University of Florida.

Leknes, S., Brooks, J.C.W., Wiech, K., & Tracey, I. (2008). Pain relief as an opponent process: a psychophysical investigation. *European Journal of Neuroscience, 28*, 794-801.

Levy, K.N., Yeomans, F.E., & Diamond, D. (2007). Psychodynamic treatments of self-injury. *Journal of Clinical Psychology, 63*, 1105-1120.

Lewis, S.P., Heath, N.L., Michal, N.J., & Duggan, J.M. (2012). Non-suicidal self-injury, youth, and the internet: What mental health professional need to know. *Child and Adolescent Psychiatry and Mental Health, 6*, 1-9.

Linehan, M.M., Comtois, K.A., Murray, A.M., Brown, M.Z., Gallop, R.J. et al. (2006). Two-year randomized controlled trial and follow-up of dialectical behavior therapy vs therapy by experts for suicidal behaviors and borderline personality disorder. *Archives of General Psychiatry, 63*, 757-766.

Loersch, C., & Payne, B.K. (2011). The situated interference model: An integrative account of the effects of primes on perception, behavior, and motivation. *Perspectives of Psychological Science, 6*, 234-252.

Lubman, D.I., Yucel, M., Kettle, J.C., Scaffidi, A.,… Allen, N.B. (2009). Responsiveness to drug cues and natural rewards in opiate addiction: Associations with later heroin use. *Archives of General Psychiatry, 66*, 205-212.

Lynch, T.R., & Cozza, C. (2009). Behavior therapy for nonsuicidal self-injury. In M.K. Nock (Ed.), *Understanding nonsuicidal self-injury: Origins, assessment, and treatment.* (pp. 211-250). Washington, DC, US: American Psychological Association.

Martijn, C., Vanderlinden, M., Roefs, A., Huijding, J., & Jansen, A. Increasing body satisfaction of body concerned women through evaluative conditioning using social stimuli. *Health Psychology, 29*, 514-520.

Newman, C.F. (2009). Cognitive therapy for nonsuicidal self-injury. In M.K. Nock (Ed.), *Understanding nonsuicidal self-injury: Origins, assessment, and treatment.* (pp. 301-319). Washington, DC, US: American Psychological Association.

Nisbett, R. E., Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231-259.

Nock, M.K. (2010). Self-injury. *Annual Review of Clinical Psychology, 6*, 339-363.

Nock, M.K., & Banaji, M.R. (2007). Assessment of self-injurious thoughts using a behavioral test. *American Journal of Psychiatry, 164*, 820-823.

Nock, M.K., Holmberg, E.B., Photos, V.I., & Michel, B.D. (2007). Self-injurious thoughts and behaviors interview: Development, reliability, and validity in an adolescent sample. *Psychological Assessment, 19*, 309-317.

Nock, M.K., Park, J.M., Finn, C.T., Deliberto, T.L., Dour, H.J., & Banaji, M.R. (2010). Measuring the suicidal mind: Implicit cognition predicts suicidal behavior. *Psychological Science, 21*, 511-517.

Nock, M.K., Wedig, M.M., Holmberg, E.B., & Hooley, J.M. (2008). The emotion reactivity scale: Development, evaluation, and relation to self-injurious thoughts and behaviors. *Behavior Therapy, 39*, 107-116.

Olson, M.A., & Fazio, R.H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin, 32*, 421-433.

Payne, B.K., Burkley, M.A., & Stokes, M.B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology, 94*, 16-31.

Payne, B.K., Cheng, C.M., Govorun, O. & Stewart, B.D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*, 277-293.

Payne, B.K., Govorun, O., Arbuckle. N.L. (2008). Automatic attitudes and alcohol: Does implicit liking predict drinking? *Cognition and Emotion, 22*, 238-271.

Payne, B.K. Krosnick, J.A., Pasek, J. Lelkes, Y., Akhtar, O., & Thompson, T. (2010). Implicit and explicit prejudice in the 2008 American presidential election. *Journal of Experimental Social Psychology, 46*, 367-374.

Payne, B.K., McClernon, F.J., & Dobbins, I.G. (2007). Automatic affective responses to smoking cues. *Experimental and Clinical Psychopharmacology, 15*, 400-409.

Prinstein, M.J., Heilbron, N., Guerry, J.D., Franklin, J.C., Rancourt, D., Simon, V., & Spirito, A. (2010). Peer influence and non-suicidal self-injury: Longitudinal results in adolescent community and clinically-referred samples. *Journal of Abnormal Child Psychology, 38*, 669-682.

Prkachin, K.M., & Craig, K.D. (1986). Social transmission of natural variations in pain behavior. *Behaviour Research and Therapy, 24*, 581-585.

Rathus, J.H., & Miller, A.L. (2002). Dialectical behavior therapy adapted for suicidal adolescents. *Suicide and Life-Threatening Behavior, 36*, 19-34.

Rehme, A.K., Frommann, I., Peters., S., Block, V., Bludau.,… Wagner, M. (2009). Startle cue-reactivity differentiates between light and heavy smokers. *Addiction, 104*, 1757-1764.

Sandt, A.R., Sloan, D.M., & Johnson, K.J. (2009). Measuring appetitive responding with the postauricular reflex. *Psychophysiology, 46*, 491-497.

Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J.,… Dunbar, G.C. (1998). The mini-international neuropsychiatric interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry, 59*, 22-33.

Solomon, R. L., (1980). The opponent-process theory of acquired motivation: The costs of pleasure and the benefits of pain. *American Psychologist, 35,* 691-712.

St. Germain, S.A., & Hooley, J.M. (in press). Aberrant pain perception in direct and indirect non-suicidal self-injury: An empirical test of Joiner's interpersonal theory. *Comprehensive Psychiatry*.

Tanimoto, H., Heisenberg, M., & Gerber, B. (2004). Event timing turns punishment to reward. *Nature, 430*, 983.

Tyrer, P., Thompson, S., Schmidt, U., Jones, V., Knapp, M. et al. (2003). Randomized controlled trial of brief cognitive behavior therapy versus treatment as usual in recurrent deliberate self-harm: the POPMACT study. *Psychological Medicine, 33*, 606-615.

Vaidyanathan, U., Patrick, C.J., & Cuthbert, B.N. (2009). Linking dimensional models of internalizing psychopathology to neurobiology systems: Affect-modulated startle as an indicator of fear and distress disorders and affiliated traits. *Psychological Bulletin, 135*, 909-942.

Whitlock, J., Eckenrode, J., Silverman, D. (2006). Self-injurious behaviors in a college population. *Pediatrics, 117,* 1939-1948.

Wood, A., Trainor, G., Rothwell, J., Moore, A. & Harrington, R. (2001). Randomized trial of group therapy for repeated deliberate self-harm in adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 1246-1253.

Zajonc, R.B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science, 10*, 224 – 228.

Zanna, M.P., Kiesler, C.A., & Pilkonis, P.A. (1970). Positive and negative attitudinal affect established by classical conditioning. *Journal of Personality and Social Psychology, 14*, 321-328.