Self-Criticism and Responses to Self-Critical Statements:
An investigation of Self-Criticism, Self-Submission, Self-Compassion, and Attack Resistance

by

Eli Cwinn

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Doctorate of Philosophy
in
Psychology

Guelph, Ontario, Canada
© Eli Cwinn, September, 2018
The overall aim of the current project was to examine self-criticism from a meta-theoretical interpersonal and evolutionary framework. To this end two studies were undertaken. Study 1 involved an experimental priming design to examine whether individuals in a compassion or resistance prime condition would be less self-critical than a non-primed comparison group following a failure task. 114 participants completed self-report measures, either a guided imagery task or a non-listening control, and a short spatial reasoning task after which they were given negative feedback. A one-way ANOVA was used to examine whether the imagery conditions had lower state self-criticism after negative feedback as compared to the non-listening control. While the results of Study 1 were null, they did support an alternative hypothesis; that there is a disconnect between interpersonal motives (compassion of self-to-other) and reflexive motives (self-to-self; e.g., self-compassion). This disconnect between the interpersonal domain and the
intrapersonal domain was examined in Study 2 which used a mixed qualitative-quantitative approach to clarify the domain of self-submission, self-compassion, and attack resistance. In particular, relations between interpersonal and intrapersonal submission were examined and found to be related but distinct. Further, results suggested that self-submission, self-compassion, and attack-resistance validly capture patterns of self-relating after self-criticism. Use of self-compassion or attack-resisting in addition to self-submission was associated with lower negative affect as compared to use of self-submission alone. Further, more awareness of one’s self-to-self interactions following self-criticism was also associated with lower levels of negative affect. Implications for therapy and future research are discussed.
Acknowledgments

The completion of this project would not have been possible without the support and guidance of my supervisor, Dr. Karl Hennig. He was a beacon in the storm, a compass in the wilderness, and a barometer at times of uncertainty. With his guidance I have acquired greater clarity of thought, writing and presenting skills, and have learned an immense amount of content that I will always carry with me. I am so very grateful for his unyielding support.

My PhD advisory committee, Drs. Melanie Parkin and Stephen Lewis were also invaluable and helped me improve this project time and again. Thank you.

I would also like to acknowledge the patience, love, and encouragement given to me from my wife, Uyen. She supported me during all late nights and early mornings without complaining of being passing ships in the night. She was a rock I could lean whenever I needed to. My parents and their partners have also been a great source of support and strength. Thank you for your faith in me and all of the practical ways you have helped me along in this journey.

Lastly, I would like to acknowledge the support of my RAs, volunteers, and the Psychology Department at the University of Guelph. I am proud to have cut my teeth in such a rigorous and supportive institution.
TABLE OF CONTENTS

Front Matter i-v
Abstract ii-iii
Acknowledgements iv
List of Tables vii

Chapter 1 - General Introduction 1
  Self-criticism as transdiagnostic 2
  Evolutionary models of behavior 3
  Gilbert’s Social Mentality Theory and self-criticism 5
  The relationship between interpersonal and intrapersonal domains 8
  Self-compassion and attack resistance ameliorate self-criticism 13
  Fostering self-compassion and attack-resistance through therapy, imagery work, and priming 17
  Current studies 21

Chapter 2 – Study 1: Experimental investigation of other-oriented compassion and attack-resistance on self-criticism 23
  Introduction 23
     Gilbert’s evolutionary model of self-criticism 23
     Self-Compassion and Attack-Resistance as responses to self-criticism 25
     Hypotheses of the current study 28
  Methods 29
     Participants 29
     Procedure 29
     Measures 32
  Results 33
     Preliminary Analyses 33
     H1: Compassion and defending groups will have lower self-criticism. 34
  Discussion 34

Chapter 3 – Study 2: Clarifying self-submission, self-compassion, and attack-resistance through a mixed qualitative-quantitative investigation 41
  Introduction 41
     Conceptualizing Self-Criticism and the Self that is Criticized 42
     Interpersonal and Intrapersonal Relations 43
     Psychological awareness is an adaptive context for self-criticism and responses to self-criticism 48
     Aims of the current study 50
  Methods 51
     Participants 51
     Procedure 51
     Measures 52
        Self-Report Measures 52
        Interview Measure 55
  Results 57
     Preliminary Analyses 57
     Aim 1: Replicate previous findings 58
     Aim 2: Develop a self-report measure of self-submission 59
Aim 3: Evaluate whether a direct measure of self-submission is more robust than a proxy measure
Aim 4: Develop and validate a semi-structured interview to examine responses to self-criticism
Aim 5: Examine which patterns of self-responding are most associated with negative affect
Aim 6: Examine psychological awareness as a predictor

General Discussion
Overview and summary of results
Implications for Therapy
Conclusion

References
Tables 107-111
Appendices 112-123

Appendix A – Imagery Transcript: Compassionate self-to-other
Appendix B – Imagery Transcript: Defending self-to-other
Appendix C – Selected items derived from the Raven’s Progressive Matrixes
Appendix D – Semi-structured interview script and instructions
Appendix E – Coding guide
Appendix F – Decision tree to guide psychological awareness scoring
List of Tables

Table 1 – Construct Definitions 108
Table 2 – Inter-correlations, means and standard deviations for variables used in Study 2 109
Table 3 – Content analysis of triggers for self-critical statements 110
Table 4 – Descriptive statistics of self-to-self responses 111
Table 5 – Hierarchical regressions predicting negative affect 112
CHAPTER 1: GENERAL INTRODUCTION

There is no person so severely punished as those who subject themselves to the whip of their own remorse - Seneca

Self-criticism is a transdiagnostic construct associated with a variety of adverse mental health outcomes and diagnostic categories, such as negative emotional biases (Casalin, Luyten, Besser, Wouters, & Vliegen, 2014), suicidality (Campos, Besser, Abreu, Parreira, & Blatt, 2014); decreased emotional responsiveness among offspring (Casalin, Tang, Vliegen, & Luyten, 2014); anger, low anger control, and anger towards self and others (Abi-Habib & Luyten, 2013); and shame (Kelly & Carter, 2013). Efforts to understand and ameliorate self-criticism have long been a focus of concern (e.g. Freud, 1917). Following self-critical statements individuals then need to respond to their inner critic. Recently, (e.g. Kelly, Zuroff, & Shapira, 2009; Whelton & Greenberg, 2005) attention has been paid to people’s responses to their own self-criticism as important factors that impact the pathogenic effects of self-criticism (e.g., responding compassionately to their inner-critic; "I failed, but everyone makes mistakes at times" V.S. submitting to their inner-critic “this proves it, I really am totally incompetent”).

Self-criticism is interesting because is aversive yet self-inflicted, (Gilbert Baldwin, Irons, Baccus, & Palmer, 2006) and involves a unique dynamic in which a person is both the domineering self-attacker and simultaneously the target submitting to the attacks (Freud, 1917). In Freud’s conceptualization of melancholia (depression), protest anger at the loss of an attachment figure is directed towards the now internalized object to which the individual identifies (the “self”). Subsequent object-relations theories have variously conceptualized self-criticism as an internalized parent-child relationship involving the “self”s” hatred and contempt towards an internal, identified-with “object” (Scharff & Tsigounis, 2003). What the various
models share is some conceptualization of the relationship between the interpersonal and intrapersonal domains, how an outside becomes a troubled inside.

In this project, three different responses to self-criticism will be examined as important to understanding and ameliorating self-criticism and its adverse mental health sequelae, they are: self-submission, self-compassion, and attack-resistance. In particular, the current set of studies focuses on clarifying the concept of submission in the interpersonal domain and its relation to the separate construct of self-submission in which one submits to oneself.

Self-Criticism as Transdiagnostic

Self-criticism is a broad concept that exists as a personality-like trait (Blatt, 2008); yet, a self-critical emotional valence can elicit self-critical behaviors at a state level (Kelly, Zuroff, & Leybman, 2016). Sidney Blatt developed the construct of self-criticism as a transdiagnostic vulnerability reflecting the propensity for some individuals to experience strong negative affect in response to agentic failure (Blatt, 2004; Blatt, 2008). This is contrasted against those individuals who are more sensitive to interpersonal injuries and are motivated more to “belong” than to “get ahead”. This characterological depiction of self-criticism, as well as others, (such as perfectionistic self-criticism, for example) has been identified as a context explaining why agentic stressors trigger psychopathology in some people and not others (Blatt, 2008).

The notion of self-critical statements differs from this broader conceptualization in that it represents a common form of self-relating that can occur in as a result of difficulties in either agentic or relatedness domains (Kannan & Levitt, 2013). Indeed, at the state level, self-criticism is a common and not inherently maladaptive (alternatively, corrective feedback; Kannan & Levitt, 2013). Self-criticism can serve an important self-monitoring function that helps people become aware of the need to modify their behavior and facilitate successful goal pursuit. Indeed, when one is making self-critical statements the lateral prefrontal cortex and anterior cingulate
cortex are activated (Longe et al, 2010). These brain regions are associated with detecting and correcting errors, planning, monitoring, and what are more broadly referred to as executive functioning (Longe et al, 2010). This may suggest that discrete self-critical statements may serve an important function. Nonetheless, self-criticism has also been identified as a transdiagnostic marker is implicated in the development, maintenance, and exacerbation of many mental health difficulties (e.g. Hooley & St. Germain, 2014; Kelly & Carter, 2013; Kopala-Sibley, Zuroff, Russell, & Moskowitz, 2014; Yamaguchi, Kim, & Akutsu, 2014).

An evolutionary (Kelly, Zuroff, & Shapira, 2009; Gilbert, 2010) and interpersonal (Wiggins, 1996) model of self-criticism provides a strong explanation for these varied findings. Namely, that the noxious effects of self-criticism do not only relate to self-criticism per se but also as a result of the ways in which we respond to self-critical statements (Kelly, Zuroff, & Shapira, 2009; Gilbert, 2010). In pursuit of better understanding self-criticism and methods of mitigating its noxious effects, this thesis applies an evolutionary model and a corresponding treatment model (Study 1) with added insights from interpersonal theory (Study 2) to examine self-criticism and responses to self-criticism.

**Evolutionary Models of Behavior**

Any specifically evolutionary model must begin with the notion of inclusive fitness (intergenerational transmission of genetic material) and the behaviors which promote or impede survival and reproduction for more than one generation (see Ferrier & Michod, 2011 for discussion). Dominance, submission, and compassion are three interpersonal behaviors that promote survival and inclusive fitness.

**Dominance and submissive behavior.** Social Rank Theory, a model of the evolutionary bases of submission and dominance, is a relatively old explanation of animal behavior (Louch & Higginbotham, 1967), yet it remains pertinent today for explaining the behavior and affect of
both human (Pinto-Gouveia, Ferreira, & Duarte, 2014; Gilbert, Allan, & Trent, 1995; Hawley, 1999; Sloman & Taylor, 2016; Troop & Baker, 2008) and animal populations (Atwood & Gese, 2007; Bauman, Toscano, Mason, & Lavenex, 2006; Grosenick Clement, and Fernald, 2007; Kruczek & Zatorska, 2008; Mainguy & Cote, 2007). Social Rank Theory posits that the ability to garner resources and support from others accounts for and predicts one’s place in a dominance hierarchy (Festinger, 1954; Gilbert, Price, & Allan, 1995). It is thought that the main purpose of a social dominance hierarchy is to act as an efficient system of resource distribution. Since each member of the species desires as much of a given resource as possible, when there are limited resources, challenge for these resources emerge. For example, two crocodiles square off, one makes an opening bid to dominate the other by vigorously hissing. The other refuses the bid in a counter-dominant response by hissing back (attack-resistance). Each crocodile escalates back and forth until one of the two submits their position as “beach master” in recognition that the more dominant crocodile represents a threat to survival (Garrick & Lang, 1977). However, the submissive crocodile needs a threat-detection system to indicate that the risks of further escalation outweigh the benefits. In humans, social rank has been associated with income, health (Wood, Boyce, Moore, & Brown, 2012), and conspicuous consumption (Walasek, Bhatia, & Brown, 2008), highlighting the parallels between rank behavior in humans and non-human animals (Gilbert, 2010).

**Compassion/care-giving behavior.** When there is no competition for resources or threats to survival, social animals engage in bonding and cooperative behavior (Reite & Field, 1985). This promotes group selection (Smith, 1964) and also underlies nurturance behaviors required to help infants mature to adulthood (Simpson & Belsky, 2008). Of particular importance is the provision of care and soothing to injured individuals (Gilbert, 2015). This compassionate behavior serves an important function because it not only preserves the ability for injured
members to continue to contribute to group survival but also because it helps foster group cohesion and a sense of social safeness among individuals within the group (Gilbert, 2015).

**Social Mentality Theory and Self-Criticism**

Corresponding to the behaviors observed above (dominance, submission, caregiving), Gilbert's Social Mentality Theory posits the existence of three corresponding systems that organize or motivate the corresponding behaviors of dominance, submission, and compassion. It is argued that humans (and non-humans) have evolved similar motivational systems that promote successful completion of various biosocial goals (Gilbert, 2000; Gilbert, 2015). These systems allow individuals to enter a motive-oriented state of mind where their attention, cognition, emotion, and actions are coordinated to help promote survival in different ways (Gilbert, 2010). Each state of mind coordinates the pursuit of different types of biosocial goals. Although there are a wide variety of motivational systems, evolutionary models of self-criticism (e.g., Social Mentality Theory) primarily speak about (1) a threat-detection system, (2) a caregiving system, and (3) a drive system. The dominance/ drive, threat-detection, and caregiving systems organize cognition, emotion, and behavior to create social interaction patterns that facilitate the achievement of different types of biosocial goals. As such, dominance, submission, and caregiving are behaviors that are elicited when a person is in a corresponding social mentality.

It is acknowledged that some people will adopt habitual ways of relating to the world and that certain mentalities will essentially become a default way of responding because those neural mechanisms will become potentiated (Gilbert, 2010). Nonetheless, the process is still conceptualized as involving the activation of a mentality and then corresponding behaviors are elicited to achieve mentality-relevant bio-social goals (Gilbert, 2010).

**Dominance/ Drive System.** Dominance-oriented behaviors are believed to be organized by a drive or competition state of mind (Cheng, Tracey, Foulsham, Kingstone, & Henrich, 2013;
Kelly, Zuroff, & Shapira, 2009). The drive system is believed to orient a person towards opportunities to gain both social and non-social resources. In a review of the literature, Fredrickson concluded that drive-based emotions "build an individual’s resources for survival. The means by which this build function was achieved was by a momentarily broadened scope of awareness, creating a form of consciousness within individuals that included a wider array of thoughts, actions and percepts than typical" (Fredrickson, 2013, p.15). Key drive system emotions include hope, excitement, interest, inspiration, (Fredrickson, 2013), superiority, strength, and power (Price, Sloman, Gardner, & Gilbert, 1994). Experiencing positive emotion during challenge tasks is associated with quicker cardiovascular recovery after a stressful task ends (Tugade & Fredrickson, 2004) suggesting that drive-based emotions play an important role in seeking opportunities and resources.

Within the context of interpersonal dynamics, this "driven" state of mind motivates bids for dominance (Van Kleef, De Dreu, & Manstead, 2010). Interpersonal theory suggests that those with high interpersonal dominance are more agentic and extraverted and that agency and extroversion are related to positive affect and emotions of excitement and arousal (Plutchik, 1997). In the context of interpersonal challenges, these emotional states motivate people to act superior and to self-aggrandize (Depue, 2006; Fu, 2013; Tracey, Cheng, Robins & Trzesniewski, 2009). By experiencing feelings of strength, power, and superiority individuals are confident to make bids for dominance (Price, Sloman, Gardner, & Gilbert, 1994) and, when successful, they experience more positive affect (Fredrikson, 2013). Further, when in drive-based states of mind people are less receptive to negative feedback (Van Kleef, De Dreu, & Manstead, 2010), which promotes ongoing goal pursuit even in the face of challenge.

**Submission/ Threat-Detection System.** Submission is understood as a behavior pattern organized by the threat-detection system. The threat-detection system orients individuals towards
possible harm and promotes behavior that facilitates the avoidance or resolution of these threats (Gilbert, 2010; MacBeth, Schwannauer, & Gumley, 2008). Subsumed within a threat-detection system is a psychobiological mechanism that helps the individual of a species evaluate the relative risk of a given challenge (Gilbert, Allan & Price, 1995). This mechanism is imperative because continued challenges for resources with a member of a much higher status, and thus skill, would likely result in frequent physical injury or death. Thus, a mechanism for de-escalating conflicts must be available to avoid conflicts that cannot be won. Price, Sloman, Gardner, and Gilbert (1992) suggest that an Involuntary Submissive Strategy is initiated when an individual cannot win a challenge (Sloman & Gilbert, 2000). The Involuntary Submissive Strategy is believed to work via negative affect, which de-motivates the individual to compete for the resource in question, thus leading to submission (Gilbert, Allan, & Trent, 1995). As such, a core feature of interpersonal submission is feelings of defeat and powerlessness, which promotes withdrawal and passivity, as well as anxiety and sensitivity to threat.

**Compassion/Caregiving System.** The compassion and caregiving system orients people towards affiliation, cooperation, and prosocial behavior (Gilbert, 2015). It has been argued that the compassion and caring system is an extension of the attachment system and involves the same neuro-physiological characteristics as attachment such as oxytocin (Carter, 2014) and vagal modulation (Porges, 2007). It is associated with feelings of warmth, contentment, and the absence of both threat and competitive drives (Lutz, Brefczynski-Lewis, Johnstone & Davidson, 2008; Depue & Moronne-Strupinksy, 2005). Although attachment and bonding are observed in a variety of social animals, human compassion is unique in that it involves highly sophisticated mentalizing (Liotti & Gilbert, 2011). Mentalizing is the process by which we develop a cognitive and experiential understanding of the motives and experiences of others (Liotti & Gilbert, 2011; Richter, Gilbert, & McEwan, 2009). This is important because it involves a *coming together of*
minds and a sense of common humanity (Neff & Dahm, 2014) and kin-like concern (Gilbert, 2015). Indeed, the root of compassion, \textit{compati} means ‘to suffer with’ (Strauss, Taylor, Gu, Kuyken, Baer, Jones, & Cavanagh, 2016). This allows humans to detect, understand, and resonate with the suffering of others and subsequently adopt a motive to help and alleviate their suffering (Liotti & Gilbert, 2011; Shahar, Henrich, Blatt, Ryan, & Little, 2003; Suddendorf & Whitten, 2001). A second component of compassion is the ability to tolerate the distress of others and the ability to engage in behaviors to alleviate that suffering. In a critical review of the measures and definitions of compassion, Strauss, Taylor, Gu, Kuyken, Baer, Jones, and Cavanagh, (2016) propose the following definition of compassion:

“a cognitive, affective, and behavioral process consisting of the following five elements:

1) Recognizing suffering; 2) Understanding the universality of suffering in human experience; 3) Feeling empathy for the person suffering and connecting with the distress (emotional resonance); 4) Tolerating uncomfortable feelings aroused in response to the suffering person (e.g. distress, anger, fear) so remaining open to and accepting of the person suffering; and 5) Motivation to act/acting to alleviate suffering” (p.19).

Interpersonally, compassionate states of mind are characterized by taking a somewhat passive role as one hears and validates the suffering of another, while also being high in caregiving and with a readiness or strength to act and help (Goetz, Keltner, & Simon-Thomas, 2010). Thus, the key is the ability to notice the distress in others and to respond with a motivation to alleviate that distress through action (Gilbert, 2015).

\textbf{The Relation between Interpersonal and Intrapersonal Domains}

People respond to imagined or symbolic stimuli in a similar manner as they would respond to the represented stimuli itself (Deacon, 2000; Gilbert, 2010; Gilbert, Baldwin, Irons, Baccus, & Hempel, 2006). A consequence of this is that people respond to \textit{intrapersonal stimuli}
in much the same way as they respond to interpersonal stimuli (Gilbert, Baldwin, Irons, Baccus, & Hempel, 2006). Despite the absence of sensory inputs, symbolic representations (e.g. self-to-self interactions) trigger bodily responses in much the same way as the represented stimuli themselves, so, for example, bullying or chastising oneself (as is the case with self-criticism) will result in a neurological and bodily response similar to that which would occur if one were being bullied or chastised by another person (Liotti & Gilbert 2011; Longe, Maratos, Gilbert, Volker, Rockliff, & Rippon, 2010).

A core assertion within Social Mentality Theory is the notion that motive oriented states of mind organize the whole person (thoughts, attention, physiology, emotion, action tendencies) based on the enacted motive and not the stimuli that evoked the motive (Gilbert, 2009; Gilbert, 2010). For example, if one is in a threat-detection mentality whilst driving and they are cut-off by another driver, they will be more likely to interpret the others driver’s actions as hostile or dangerous and may then respond to that other driver with aggression (e.g. swearing, shaking fist, honking, etc.). On the other hand, if one were in a safeness or caregiving mentality and was cut-off whilst driving, they may interpret the behavior of the other driver as a natural part of Toronto driving and respond by giving more space for the person to merge. This global organization of self is important because it means that the action tendencies are organized by the enacted motive and will generalize across stimuli.

This generalization across stimuli also applies to social stimuli, specifically, that there is generalization in action tendencies across interpersonal and intrapersonal stimuli (Gilbert & Proctor, 2006) and that this generalization occurs as a result of an activated social mentality which organizes the whole person (Gilbert, 2015). One can reference the Talmudic idea we do not see things as they are, we see things as we are (Simons & Epstein, 2018) to capture the notion that we interpret both interpersonal and intrapersonal stimuli based on our state of mind at
the time of the event. In the current study the phrase ‘transitive action of motive-oriented states of mind’ is used to capture the generalization of action tendencies between interpersonal and intrapersonal stimuli. This is important for the study of self-criticism because it suggests that different mentalities will alter the ways in which people experience and respond to self-critical statements.

As is the case interpersonally, the current thesis suggests that people can respond to their own self-critical statements by submitting to their inner-critic, by providing compassion to themselves, or by resisting the self-critical statement. As with the interpersonal domain, submission is assumed to be associated with more negative affect and a sense of giving up or withdrawal (potentially contributing to the noxious impacts of self-criticism) whereas soothing oneself or resisting the self-attack is assumed to be associated with less negative affect and a quicker return to positive affect (perhaps accounting for the finding that self-criticism can be adaptive for some people).

**Submission and self-criticism.** Investigators have suggested that individuals might submit to their self-criticism through a mechanism similar to the involuntary submissive strategy which evolved in an interpersonal context (Castilho, Pinto-Gouveia, Amaral, & Duarte, 2014; Santor, 2003). In much the same way as one can respond to others with submission (Price, Sloman, Gardner, & Gilbert 1994), people can also respond to their self-criticism with submission and/or to be self-critical is to already have submitted (Whelton & Greenberg, 2005). Much like interpersonal submission, researchers have found that self-submissiveness (submission to an inner-critic) is associated with displays of powerlessness and depressive symptoms (Whelton & Greenberg, 2005).

Whereas interpersonal submission serves a clear function of de-escalating potential conflict (Gilbert, Allan, & Trent, 1995), the benefits of self-submission are less clear (Gilbert,
Self-submission (submitting to one’s self-criticism) may be partially adaptive by helping an individual to acknowledging fault and correct maladaptive behaviors. Yet, if the mechanism for self-submission is similar to the Involuntary Submissive Strategy, then negative affect, powerlessness, and disengagement are likely involved in self-submission. If this is the case, self-submission would not likely result in behavior changes and motivation to pursue goals in a meaningful sense. While individuals may appease themselves by submitting to self-criticism and then reassuring themselves that they will do better next time, when in a threat-based state of mind and engaging in self-submission they are more likely to passively ruminate rather than engage in positive behavior change. Indeed, research suggests that when people experience stressful life events that they are powerless to change (a state of mind similar to a submission) they are more likely to experience rumination and this mediates the relation between stressors and depressive symptoms (Michl, McLaughlin, Shepard, & Nolan-Hoeksema, 2013).

Further supporting the claim that self-submission is maladaptive, researchers have found that submission to an inner-critic is associated with displays of powerlessness and depressive symptoms (Whelton & Greenberg, 2005). The majority of research on self-submission has used interpersonal submission as proxy variable or probabilistic indicator and has not strongly articulated the role of self-submission per se, but rather, has implied that is has a deleterious effect. Using this methodology, researchers have found that submission is associated with depressive symptoms (Kopala-Sibley, Rappaport, Lance, Sutton, Moskowitz, & Zuroff, 2013; Ongen, 2006), self-criticism (Gilbert, Baldwin, Irons, Baccus, & Palmer, 2006), and severity of auditory hallucinations people with Schizophrenia and self-critical thoughts in people with Major Depressive Disorder (Gilbert, Birchwood, Gilbert, et al, 2001). Researchers investigating cognitive therapy for depression have also argued that believing one’s self-criticism and being
unable to reframe or refute one’s self-criticism is depressogenic and associated with interpersonal submission (Rector, Bagby Segal, Joffe & Levitt, 2000).

In the current thesis self-submission is conceptualized as an active agreement with the self-critical statement combined the absence of self-protective responses. This is congruent with self-submission coded using two-chair dialogues where submission to an inner-critic was associated with an inability to refute the critic as well as non-verbal signs of shame when such attempts were made (signaling that the person does not feel entitled to disagree with the specific self-critical statements) (Whelton & Greenberg, 2005).

Despite a growing recognition that self-submission is a maladaptive form of self-to-self relating with few adaptive features (Whelton & Greenberg, 2005), research has yet to clarify the concept of self-submission and is relation to interpersonal submission. There are no self-report measures on self-submission and other than using interpersonal submission as an indirect indicator, research has only investigated self-criticism using a two-chair dialogue. In addition, there is an absence of research that has examined self-submission concurrently with other forms of self-to-self interactions to evaluate the intrapersonal dynamics following self-criticism. This is important because the research cited above suggests that self-submission may be an important factor in explaining why self-criticism can be harmful for some people and not others (Longe et al, 2015). Clarifying the construct of self-submission advances the field by adding a more robust ability to capture a potentially important aspect of pathogenic self-criticism. Furthermore, given that self-criticism is difficult to change once it becomes a personality orientation (Bergner, 2013; Zuroff, Moskowitz, & Coté, 1999) identifying the role of self-submission in relation to other forms of self-to-self responses may point to an important means of ameliorating the pathogenic effects of self-criticism because it focuses on the response to the self-criticism.
Self-Compassion and Attack-Resistance Ameliorate Self-Criticism

**Self-compassion ameliorates self-criticism.** Self-compassion has received considerable attention in the literature over the past 14 years due to its relation to a variety of positive outcomes such as subjective wellbeing (Hollis-Walker & Colosimo, 2011; Wei, Liao, Ku, & Shaffer, 2011), happiness and self-esteem (Mongrain, Chin, & Shapira, 2011), and a sense of mattering (Raque-Bodgan, Ericson, Jackson, Martin, & Bryan, 2011), among others. At the same time self-compassion is associated with better psychological adjustment after parental separation (Sbarra, Smith, & Mehl, 2012), lower levels of interpersonal problems (Schanche, et al, 2011), and fewer symptoms of post-traumatic stress after trauma (Dahm, et al, 2015).

Self-compassion has become especially important in the context of self-criticism. Interventions directed at increasing self-compassion are associated with lower levels of self-criticism (Gilbert & Procter, 2006; Judge, Cleghorn, McEwan, & Gilbert, 2012), shame (Judge, Cleghorn, McEwan, & Gilbert, 2012), and disorders characterized by shame and self-criticism like eating disorders (Gale, Gilbert, Read, & Goss, 2012). Although congruent with one-another, Neff (2016) and Gilbert (2015) have differing ways of defining and measuring self-compassion.

Neff defines self-compassion as “compassion turned inward and refers to how we relate to ourselves in instances of perceived failure, inadequacy, or personal suffering” (2016, p. 265). Her measurement of self-compassion is broad, multi-faceted, and rooted in elements of Buddhist thought (kindness vs. criticism, common humanity vs. isolation, mindfulness vs. over-identification). Gilbert's conceptualization of self-compassion is rooted in Social Mentality Theory and the caregiving system which encapsulates the attachment system. Gilbert’s definition focuses on the motive of diffusing distress (rather than the descriptive attributes as in Neff). While the act of self-compassion may involve the specific qualities described by Neff (2016), the current thesis is guided by Gilbert’s model which places a greater emphasis on the motive of
self-compassion as compared to describing its features. Self-compassion is herein based upon the
notion of a secure based script involving actions aimed at diffusing distress. In the interpersonal
domain, “If I encounter an obstacle and/or become distressed, I can approach a relationship
partner for help; he or she is likely to be available and supportive; I will experience relief and
comfort as a result of proximity to this person; I can then return to other activities” (Mikulincer,
Shaver, Sapir-Lavid, & Avihou-Kanza, 2009, p. 616). The sensitive and responsive capacity to
diffuse of one’s own distress is a foundational component of self-compassion in the current
paper. Self-compassion therefore involves being attuned to and noticing one’s own distress and
seeing and responding to it with caregiving and soothing behaviors. This conceptualization is
used in the current paper because it centers on the motive and provides a functional account of
the transitive action of motive-oriented states of mind might be leveraged to interrupt pathogenic
self-criticism (Judge, Cleghorn, McEwan, & Gilbert, 2012).

Gilbert’s caregiving-soothing system (rooted in the attachment system) remains global,
vis-a-vis the motive to provide comfort and soothing to others and/or the self (Gilbert, Clark,
Hempel, Miles, & Irons, 2004). Gilbert conceives of compassion as a global capacity, a motive-
oriented state of mind that, when activated, promotes compassionate behavior towards others as
well as oneself. Conceptually central to the current studies, when individuals are in highly
distressed states of mind, reflecting activation of the submission/ threat-detection system, the
compassion/ caregiving system is less accessible (Gilbert, 2010). The capacity for metacognition
and be empathically attuned to and understanding of one’s inner state of distress is off-line in
distressed states of mind. Stated otherwise and central to Study 1, to activate the compassion/
care-giving system, associated with feelings of warmth and safeness, simultaneously activates
one's capacity for self-reflection, self-compassion and compassion towards others.
One function of self-compassion that is particularly important to self-criticism is that it helps people acknowledge faults, be empathically attuned to oneself, without becoming overwhelmed by negative affect (Leary, Tate, Adams, Allen, & Hancock, 2007). By facilitating self-soothing (Liotti & Gilbert, 2011), self-compassion may allow people to accept their faults and inadequacies as true while maintaining a positive view of themselves and positive affect. Additionally, self-compassionate responses are not based on success or prestige, making them ideal for soothing distress after self-criticism. Increasingly, self-compassion is being viewed as a powerful means of treating difficulties related to self-criticism (Neff, Rude, & Kirkpatrick, 2007a,b). In summary, self-compassion can be phrased as being mindfully present in a non-judgmental manner or being metacognitively self-present via the activation of the compassion/caregiving system.

**Attack-resistance ameliorates self-criticism.** In Gilbert's early writing, the experience of submission was focused on interpersonal submission and self-criticism as a reflection of perceived low social status (e.g., Allan & Gilbert, 1997). It was not until later that Gilbert's writing reflected the notion of standing up to one's inner critic (e.g., Gilbert & Proctor, 2006; Mayhew & Gilbert, 2008).

A fuller intrapersonal (or self-to-self) relating enters in a study where Kelly, Zuroff, and Shapira (2009) had participants imagining and talking back to an "inner critic" (attack resistance). Kelly, Zuroff, & Shapira (2009) developed a short self-directed intervention to foster attack-resistance as a means of reducing participants' self-submission. After a two-week self-directed intervention, participants who practiced attack-resistance had lower depression symptoms and shame than controls. Other studies have found that individuals who are able to resist their self-attacks have lower levels of self-criticism as measured by the Depressive
Experiences Questionnaire and that those with lower levels of self-criticism are less likely to accept the content of their self-critical statements (Whelton & Greenberg, 2005).

Within the field of self-criticism, attack-resistance is a relatively new and understudied form of self-to-self relating. In a dialogical interaction following self-criticism, attack-resistance involves taking a stance where one fights back against the inner-critic by presenting him/her with arguments in a way that is brave, strong, and resilient (Kelly, Zuroff, & Shapira, 2009). At its core, attack-resistance is viewed within the context of dominance and submission and can be adopting a dominant role in response to an inner self-critic during the dialogical interaction (Kelly, Zuroff, & Shapira, 2009).

In an effort to sharpen the definition of attack-resistance, the current thesis used the vertical dominance-submission axis of interpersonal theory (Wiggins, 1996) and therefore places self-submission and attack-resistance as two ends of the same bipolar continuum. That is, a person is not typically thought of as being both dominant and submissive at the same time. While this model makes sense in the interpersonal sphere where dominance by definition involves someone being dominated (or submitting), it may be less clearly applicable to self-to-self relating. The model of intrapersonal dominance and submission only makes sense if self-submission and attack-resistance are understood as discrete states of mind (or part selves) that one can alternate between, or perhaps are active simultaneously in a self-conflictual way.

Supporting the notion of alternating global states of mind, experimental evidence suggests that when individuals are asked to write about their own strengths they are protected against fear related to being evaluated as indicated by both self-report and cardiovascular data (Fredrickson et al, 2000). That is, when the dominance/drive system is activated, the submission/threat-detection system is off-line. Public speaking is a frequently feared task because it makes one vulnerable to judgement and scrutiny by others (Gilbert, 2000; Gilbert,
2001). When people are in a drive-based state of mind they are more likely to evaluate public speaking as a challenge rather than a threat (Tugade & Fredrickson, 2004) suggesting that drive-based states of mind may desensitize people to harm while sensitizing them to potential gains.

Within the context of self-to-self relating following self-criticism, a drive-based state of mind may promote attack-resistance. In response to self-criticism (a bid for dominance from a domineering internal critic) one responds by counter-dominating and protecting the self (attack-resistance). As such, attack-resistance is believed to involve a focus on one’s strengths and competencies, a refutation of the content of criticism, and a focus on remounting the goal pursuit related to the initial failure. Although understudied, attack-resistance is a promising means ameliorating the noxious effects of self-criticism.

**Fostering Self-Compassion and Attack-Resistance through Therapy, Imagery Work, and Priming**

Compassionate Focused Therapy (henceforth, CFT) (an intervention based on Gilbert’s evolutionary model of self-criticism) for shame-based difficulties often (though not prescriptively) begins by providing psychoeducation on how compassion mindsets will apply equally regardless of the target (self or other) and proceeds to focus on training the ability to activate a model of compassionate self-to-other as a first phase of intervention (Gilbert, 2009). In one study, participants completed a 12-week CFT group. As with many other CFT programs, participants begin by practicing the ability to activate a model of compassionate self-to-other (Gilbert, 2010). Later the participants trained their ability to create an internal representation of their ideal caregiver and to imagine how this internal caregiver would react to the inner-critic (activating compassionate other-to-self). Participants were provided with psychoeducation explaining how compassionate mind states will transfer across stimuli (self-other) prior to the training in imagery (highlighting the global action of motive-oriented states of mind). The 12-
week course was associated with significant decreases in self-criticism and shame (Gilbert & Procter, 2006). It is important to note that these interventions help clients activate both compassionate self-to-other and compassionate other-to-self. Intervention studies demonstrate that CFT is associated with global improvement in psychosis (Braehler et al 2013) and recovery after psychosis (Laithwaite et al 2009), depression (Gilbert & Procter, 2006; Judge, Cleghorn, McEwan, & Gilbert, 2012; Laithwait et al, 2009; Lucre & Corten, 2013; Shapira & Mongrain, 2010), anxiety (Gilbert & Proctor, 2006; Judge, Cleghorn, McEwan, & Gilbert, 2012), PTSD (Beaumont et al, 2012), and stress (Lucre & Corten, 2013; Judge, Cleghorn, McEwan, & Gilbert, 2012), among others. This seems to support the therapeutic benefit of training the ability to activate both compassionate self-to-other and compassionate other-to-self.

As indicated, attack-resistance is a relatively new construct in the self-criticism literature but it is based on principles from Cognitive Behavior Therapy (CBT) in which clients separate themselves from their self-critical messages (in CBT – identifying automatic thoughts) and talk back against them (reframing, identifying cognitive distortions, cognitive restructuring), albeit not as personifications as an "inner critic." Research in this tradition has found that CBT is effective for disorders characterized by high self-criticism such as social phobia (Cox, Walker, Enns, & Karpinski, 2002), eating disorders (Fairburn, Cooper, & Shafran, 2003), and social anxiety disorder (Iancu, Bodner, Ben-Zion, 2015). Furthermore, self-criticism predicts differential response to treatment for depression (Blatt, Zuroff, Bondi, Sanislow, & Pilkonis, 1998; Marshall, Zuroff, McBride, & Bagby, 2008). Therapeutic interventions that target self-compassion and attack-resistance appear to be effective in treating both self-criticism and related disorders where self-criticism is a major feature. Within therapeutic work, the use of imagery exercises is a distinct mode of intervention.
**Imagery intervention research.** Evidence from imagery exercises supports the argument that one can harness the action of global motive-oriented states of mind to foster therapeutic change. Where imagery exercises differ is in whether the clients (or research participant) are invited to imagine themselves as the subject or the object of the compassion. That is, whether the participant is the subject receiving compassion from an imagined ideal other (other-to-self) or whether the participant is themselves imagined as the ideal compassionate person offering compassion towards another (self-to-other). A fundamental assumption in Gilbert's mentalities model is that activation of the compassion/caregiving system should result in a capacity for both other- and self-compassion.

In terms of fostering the ability to activate compassionate other-to-self, a pilot study of CFT for traumatic-brain injury focused on helping the client develop a sense of a compassionate-other who could replace the inner-critic and provide compassionate self-correction as opposed to self-criticism. They found significant decreases in depression and anxiety post-treatment as well as lower levels of anger towards herself (Ashworth, Gracey, and Gilbert, 2011). Compassionate other-to-self imagery also alters heart rate variability and salivary cortisol in a manner congruent with emotional regulation (Rockliff, Gilbert, McEwan, Lightman & Glover, 2008).

Evidence also supports the benefits of training people to activate compassionate self-to-other. For example, Loving kindness meditation is an imagery exercise where one imagines that they are providing compassion to others. In a review of the literature on Loving kindness meditation, researchers found that this intervention decreased anxiety, depression, and stress than a passive control as well as a tendency to interpret neutral stimuli more positively (Galante, Galante, Bekkers & Gallacher, 2014).

There have been no studies examining resisting other-to-self or resisting self-to-other imagery on mental health outcomes. Nonetheless, Kelly, Zuroff, and Shapira (2009)
demonstrated that practicing attack-resistance in response to an internal critic resulted in a decrease in depressive symptoms and shame compared to a waitlist control. Further, Gilbert and colleagues found that in imagery exercises, people with low attack-resistance struggled to talk back against their self-criticism and reported feelings of powerlessness in response to the self-criticism (2006). Low attack-resistance was associated with depressive symptoms in this sample.

**Priming research.** Priming compassionate states of mind where the participant gives compassion to others (self-to-other) also results in a broader global action. In terms of activating compassionate other-to-self, priming research has demonstrated that when participants are asked to imagine that they are being cared for (other-to-self), they act more compassionately towards others as demonstrated through a decrease in out-group bias (Mikulincer & Shaver, 2001; Mikulincer, Shaver, Gilrath, & Nitzberg, 2005). Similarly, when asked to imagine that they are being cared for (other-to-self), participants are more likely to agree to switch places with depictions of other people suffering and they report a greater willingness to help people in distress as compared to neutral controls. Activating compassionate other-to-self through priming also increases self-compassion (Kelly, Zuroff, Leybman, & Gilbert, 2012).

Another study asked participants to read either a story about compassion giving, a funny story, or a neutral story, those in a compassion giving condition report higher levels of compassionate emotions than the other groups (Mikulincer, Gillrath, Halevy, Avihou, Avidan & Eshkoli, 2001). Further, priming compassionate self-to-other though a single love and kindness meditation session increases one’s sense of social connectedness and interdependence (Hutcherson, Seppala, & Gross, 2008). To this author’s knowledge, no research has evaluated
whether activating compassionate self-to-other changes the way people interact with themselves at the state level.

Due to the relatively recent introduction of attack-resistance as construct of import, no research has directly investigated the transitive action when priming either other-resisting or self-resisting. Nonetheless, research from different yet related experiments suggests that a transitive action is likely to occur. For example, research has found that semantic primes involving pride (a drive system emotion tied to attack-resistance) results in changes to body posture (Oosterwijk, Rotteveel, Fischer, & Hess, 2009). Further, priming feelings of pride increases inhibitory control compared to priming feelings of happiness (Katzir, Meiran, & Kessler, 2010). This suggests that pride may be related to goal directed behavior, a hypothesized feature of attack-resistance. The same is true in terms of reducing indulgence for both health and consumeristic behaviors (Wilcox, Kramer, & Sen, 2011). To this author’s knowledge, no research has examined whether activating resisting self-to-other alters the way people interact with themselves.

By activating compassionate and resisting self and other, interventions have harnessed the transitive action of motive-oriented states of mind for therapeutic gain. The effectiveness of these motive-oriented states of mind has been demonstrated by asking participants to practice an imagery exercise over several days. Priming research provides further support of the transitive action of motive-oriented states of mind. Nonetheless, research has yet to examine whether priming compassionate and resisting-self changes the ways in which people relate to themselves.

**Current Studies**

Guiding the current studies, an interpersonal and evolutionary model of behavior was used to re-conceptualize self-criticism. Focusing on the internalization of interpersonal submission, compassion, and dominance the model suggests that self-submission, self-compassion, and attack-resistance are important constructs in understanding and ameliorating the
pathogenic role of self-criticism (Kelly, Zuroff, & Shapira, 2009). Of particular importance is the role of motive-oriented states of mind in promoting a transitive action between interpersonal and intrapersonal interactions (Gilbert, 2010; Gilbert, 2015).

Two studies were undertaken. In Study 1 participants were asked to complete a guided imagery exercise designed to elicit either compassionate self-to-other or resisting self-to-other states of mind. Subsequently they were exposed to negative feedback and were then asked about their levels of self-criticism in that moment. This study was designed to examine whether activating motive-oriented states of mind focused on giving aid provided a benefit in terms of giving oneself aid (as indicated by lower self-criticism). In study 2 participants completed self-report measures on negative affectivity, interpersonal submission, self-submission, and self-compassion and to complete a semi-structured interview about their self-criticism and responses to self-critical statements. The goal of this second study is to examine self-to-self interactions following self-criticism with a focus on clarifying the domains of self-submission, self-compassion, and attack-resistance and how self-submission may be related to interpersonal submission.
CHAPTER 2: Study 1-
An Experimental Investigation of Self-Compassion and Attack-Resistance on Self-Criticism

The current study is guided by an evolutionary model of self-criticism, in particular Gilbert's Social Mentality Theory (2009). The model has been used to inform a number of promising treatment approaches for self-criticism (Leaviss & Uttley, 2015). Self-criticism is a common response to failure, often viewed as self-punishing and motivating increased efforts towards success. However, where self-criticism is excessive or not followed by more adaptive responses, maladaptive outcomes arise (Kopala-Sibley, Zuroff, Russell, & Moskowitz, 2014).

Recent research has focused on the examination of individuals' counter-responses to their own self-critical remarks following failure; in particular, a rapidly growing body of research has examined *self-compassion* as an ameliorating response (Neff, 2003; Neff, Rude, & Kirkpatrick, 2007a,b). The general purpose of the current study is to examine whether priming compassionate and resistance through a guided imagery exercise results in lower levels of self-criticism after a failure induction.

**Gilbert's Evolutionary Model of Self-Criticism**

The current study was guided an evolutionary model of self-criticism, and Gilbert's Social Mentalities Theory in particular (e.g., 2010). Any “evolutionary” model of self-criticism must begin with the foundational premise of a competition for survival (fitness). Social Mentality Theory posits that human behavior can be understood as an interplay between three motivational systems believed to orient attention, cognition, emotion, arousal and behavior to promote pursuit of different types of biosocial goals (Gilbert, 2009): resource acquisition and domination within the social hierarchy (hereafter, the *dominance/drive system*), avoiding harm within situations of conflict with similarly dominance seeking con-specifics (hereafter, the *submission/threat-system*), and...


detection system), and compassionate behavior (hereafter, the compassion/caregiving system).

Well-being (or psychopathology) results from problems among the three systems.

**Dominance/ Drive System.** Members within a species compete for limited resources, territory, mates, and social alliances within a broader social dominance hierarchy (Allan & Gilbert, 1995). The motivation to seek dominance is important because higher positions on a hierarchy increase access to resources and therefore influence inclusive fit (intergenerational transmission of genetic material). Within Gilbert's model the (dopamine related) drive/dominance-seeking system is resource focused.

**Submission & Threat-Detection System.** Humans and non-human animals possess a threat-detection system that signals to an individual that one is better off submitting than risking life and limb in further escalation (Gilbert, Allan, & Trent, 1995). The internal correlated emotion is that of anxiety to threat. Highly correlated with anxiety, the depressive state arises when the individual experiences a final loss of the goal-directed target (Fournier, Moskowitz, & Zuroff, 2002). Withdrawal is a behavioral criterion of depression. In a contest, the individual who submits can lose territory or a chance to mate, thereby reducing their likelihood of transmitting genetic material but increasing their chance at longevity and therefore future opportunities to regain their rank (Garrick & Lang, 1977).

**Compassion and Caregiving System.** A final system of importance is the compassion system, which promotes bonding, cooperation, attachment, and safeness (Gilbert, 2015). The most salient feature of this ethological model is the global nature of the motivational system which allows for a "transitive" action of the motive-oriented state of mind to organize behavior by motive rather than stimulus. Central here is the notion of "transitive action" or "transitive states of mind" where the term is used to convey both grammatically transitive (e.g., "He dominates me"); other-to-self) and intransitive (e.g., "I have compassion for him"); self-to-other)
actions, as well reflexive actions involving an intransitive action towards oneself (e.g., "I have compassion for myself"). The assumption of Gilbert's global systems/mentalities is that the activation of any given system (e.g., compassion system) crosses all or is "transitive" across all domains, interpersonal and intrapersonal. Interventions involving the priming of compassion for others is expected to foster compassion towards oneself.

**Self-Compassion and Attack-Resistance as Responses to Self-Criticism**

Within Social Mentality Theory, well-being results from an ease of activation and balance among the three interacting systems (Gilbert, 2010). For example, individuals with a low threshold for activating the submission/threat-detection system readily perceive themselves of low social standing and inferior to others, inducing behavior aimed at appeasing and submitting to others to avoid conflict. Resulting feelings of inferiority lead to depressive symptoms (Gilbert, Allan, & Trent, 1995). Similarly, excessively seeking status and competitiveness, actions of the dominance/drive system may result in depressive symptoms in the face of failure (Dunkley, Zuroff, & Blankstein, 2003). Interventions have used imagery (priming) exercises to activate either compassion/caregiving or dominance/drive systems to effect change (Gilbert & Proctor, 2006).

**Self-Compassion.** Self-compassion has been identified as an important construct in understanding ways of ameliorating self-criticism, understood as “compassion turned inward and refers to how we relate to ourselves in instances of perceived failure, inadequacy, or personal suffering” (Neff, 2016, p.265). Self-compassion reflects the self-reflexive (i.e., turned inward) actions of the compassion/caregiving system. Arising from this literature Compassion Focused Therapy (CFT) is a 3rd wave Cognitive Behavioral Therapy that utilizes a range of imaginal exercises (among other techniques) to facilitate therapeutic change. CFT is associated with global improvement in psychosis (Braehler et al 2013) and recovery after psychosis (Laithwaite
et al 2009), depression (Beaumont et al, 2012; Gilbert & Proctor, 2006; Judge, Cleghorn, McEwan, & Gilbert, 2012; Laithwait et al, 2009; Lucre & Corten, 2013; Shapira & Mongrain, 2010), anxiety (Gilbert & Proctor, 2006; Judge, Cleghorn, McEwan, & Gilbert, 2012), and stress (Lucre & Corten, 2013; Judge, Cleghorn, McEwan, & Gilbert, 2012), among others.

Numerous studies have demonstrated that self-reported self-compassion is associated with low levels of self-criticism and related mental health sequelae (Neff, 2003; Neff, Rude, & Kirkpatrick, 2007a,b), suggesting that people who can access a compassionate and self-soothing state of mind are protected from the pathogenic effects of self-criticism. Longer term self-compassion programs are associated with lower body dissatisfaction (Albertson, Neff, & Dill-Shackleford, 2015), higher levels of wellbeing (Neff, 2004; Neff & Germer, 2013), and less shame and self-criticism (Gilbert & Proctor, 2006). Accessing compassionate states of mind help people tolerate distress (Leary, Tate, Adams, Allen, & Hancock, 2007) and facilitate self-soothing (Liotti & Gilbert, 2011).

Priming self-compassion activates the compassion/caregiving system and has been shown to be associated with positive outcomes. Researchers have found that high trait self-compassion may associate with greater ease and intensity of being able to generate state self-compassion through a guided imagery task, which, in turn, may associate with lower levels of depressive symptoms (Gilbert, Baldwin, Irons, Baccus, & Palmer, 2006). Activating self-compassion is associated with lower self-criticism during experimental manipulations that elicit self-criticism (Leary, Tate, Adams, Allen, & Hancock, 2007). By training individuals to develop the ability to activate compassionate states of mind, intervention research has attempted to increase self-compassion (Galante, Galante, Bekkers & Gallacher, 2014). Even short-term imagery interventions exert a therapeutic benefit. Self-compassion imagery has been shown to
alter heart rate variability and salivary cortisol in a manner congruent with emotional regulation (Rockliff, Gilbert, Mcewan, Lightman & Glover, 2008).

Loving kindness meditation is an imagery exercise in which people imagine themselves providing compassion to others. In a review of the literature comparing the effect of loving kindness meditation to control groups, loving kindness meditation was found to be associated with decreases in anxiety, depression, and stress compared to passive controls (Galante, Galante, Bekkers & Gallacher, 2014). Using a pre-post design, Klimecki, Leiberg, Lamm and Singer (2012) found that one day of guided compassionate self-to-other imagery increased positive affect and empathy when watching videos of others in distress. The study used fMRI, self-report, and compassionate memory tasks to measure the change in the ways participants react to the portrayals of distress. Investigators found that on all methods of measurement, participants practicing compassionate self-to-other imagery increased neural, behavioral, and experiential empathy and positive affect.

**Attack-Resistance.** While not labeled as such, "attack resistance" is a central component of cognitive behavioral and other forms of therapy in emphasizing the refutation of dysfunctional, self-critical thoughts (Gilbert, 2000). Research has demonstrated that being able to refute and reframe one’s self-criticism protects one from depression (Rector, Bagby Segal, Joffe & Levitt, 2000). Some clients, however, have found little change despite recognizing that their self-critical beliefs are unfounded. Beliefs regarding self-blame over incidents (e.g., sexual assault), for example, are often found to be highly resistant to challenge (Harkins, Howard, Barnett, Wakeling, & Miles, 2015). Yet, the use of imagery exercises to activate attack-
resistance (associated with the dominance/drive system) is a relatively new approach in the self-criticism literature (e.g. Kelly, Zuroff, & Shapira, 2009).

Studies in self-criticism have demonstrated that attack-resistance (the diametric opposite of self-submission; alternatively, counter-dominance) is negatively associated with depressive symptomology (Whelton & Greenberg, 2005). When comparing the effects of a short self-directed intervention Kelly, Zuroff, and Shapira (2009) found that activating attack-resistance was associated with lower levels of depressive symptomology and shame than a waitlist control.

**Hypothesis of the Current Study**

CFT arose out of the finding that cognitive reframing and challenging the veracity of people’s thoughts was not effective in persons with high self-criticism because they did not believe the reframe (Gilbert, 2010). Gilbert (2000) noticed that cognitive reframing works best when there is congruence with the reframe and the emotional tone of that reframe. Using imagery (and other techniques) CFT helps clients enter a caregiving mentality so that they are able to reframe their cognitions from a stance of warmth, caring, and non-judgement rather than from a threat-based response (Gilbert, 2000; Gilbert 2010). At the same time, CFT helps clients shift their default stance in the world away from threat-detection mentalities so that when they do experience setbacks or failures, they are more likely to respond with self-compassion or attack resistance (as a result of the globally organizing effect and transitive action of motive oriented states of mind) (Gilbert, 2010; Laithwaite, et al, 2009).

While studies examining priming self-compassion have been undertaken, the experimental manipulation of attack resistance viewed as strengthening the drive system (“imagine you are a strong, confident... persistent person”) has been relatively understudied. The general aim of the current study was to examine whether priming compassion and resistance
resulted in lower levels of self-criticism after failure induction. The hypotheses of the study were:

**H1.** Participants in both compassion and resistance guided imagery prime conditions will report lower levels of state self-criticism after a failure induction task compared to the non-primed condition.

**H2.** Compassion and resistance primes would result in similar treatment effects, although the current between-subjects design will be unable to examine their relative independent contributions.

**Methods**

**Participants**

Participants were 114 predominantly White (78.1%) female (93.0%) undergraduate students with a mean age of 20.17 years ($SD = 1.5$ yrs.) who were recruited from 1st year psychology courses from a university in Southwestern Ontario. Ethics clearance was granted from the University of Guelph Ethics Review Board and participants were given course credit for volunteering their time to take part in the study. Participants completed all components of the study on-line.

**Procedure**

Participants were randomly assigned online into either a compassion, resistance, or control condition. The experimental study involved three components: a guided imagery listening exercise, followed by a failure induction procedure, and self-report survey. The control condition excluded the guided imagery induction. After receiving negative feedback participants were directed to complete a state self-criticism measure and a manipulation check. To the author’s knowledge there have not been any previously validated self-criticism induction procedures. Self-criticism is believed to occur after one fails to meet one’s expectations or standards. As such, it was assumed that if participants exerted effort into solving a series of problems and were then told that they performed poorly, the majority of participants would experience some level of state self-criticism.
Participants were initially told that the purpose of the study was to examine the effects of imagery exercises on spatial reasoning, when the actual purpose was to examine the effects of imagery on state self-criticism. In concluding the experiment participants were debriefed and informed of the deception and that the performance feedback was false. Participants had the option of their data being omitted from the study without penalty. No participants opted to omit their data. The two guided imagery exercises are described next.

**Compassionate Imagery Condition.** Participants in the compassionate condition listened to an audio-recorded guided imagery exercise based on instructions from Gilbert (2010, p.160-161, p177-179). This exercise was chosen because it is recommended as an introductory exercise to help foster cognitive and emotional qualities that allow one to activate a compassionate state of mind and is frequently recommended as means of beginning to train compassion and self-compassion (Gilbert, 2010). As outlined in Gilbert (2010), participants were first instructed to alter their breathing and facial expressions to establish the embodied experience of soothing and compassion. They were then instructed to imagine that they were a ‘perfectly compassionate person’ and focus on three specific qualities of compassion: wisdom, sensitivity to the suffering of others and the distress tolerance to sit with the suffering, and strength. Next, the participant was instructed to bring to mind someone else who they care about and who is suffering. They were asked to imagine that as a ‘perfectly compassionate person’ they could help the person find peace and happiness. At the end, they were asked to repeat the phrase, ‘may you find peace, may you be happy’ with each out-breath until a bell chimed.

**Resistance Imagery Condition.** The resisting self-to-other imagery was adapted from instructions on imagery training within the Compassionate Mind Training tradition (e.g. Gilbert, 2010) and the description of an attack-resistance intervention outlined in Kelly, Zuroff and Shapira (2009). As described in Kelly, Zuroff, and Shapira (2009) the imagery exercise was
designed to foster a sense of perseverance, clear thinking, and strength. Because resisting can come from a threat-protection stance or an agentic and helping stance, the intervention used in the present study followed a similar script as that outlined in Compassionate Mind Training to induce the helping form of resisting. This allows for a clearer evaluation of the transitive action of resisting states of mind (activated interpersonally) on self-criticism.

As outlined by Gilbert (2010), participants were first instructed to engage in soothing rhythm breathing and to manipulate their posture to begin to establish the embodied experience of strength and calm. They were then instructed to focus on specific qualities of resisting: perseverance, powerful reasoning and clarity of thought, and great strength. They were then instructed to bring to mind someone else who they care about and who is suffering. They were asked to imagine that as a ‘strong and loving person’ they could help the person solve their problems and overcome their challenges. At the end, they were asked to repeat the phrase, ‘together I will help you solve your problem. I will help you. We will succeed’ with each out-breath until a bell chimed.

The above imagery exercises were pilot tested on 15 volunteers of similar demographics to the participant population. Pilot tested participants were asked to rate the degree to which it altered their mood and provide 3-5 adjectives describing their new mood. Pilot testing suggested that the imagery exercise induced the targeted self-compassionate or resistance state of mind. The full imagery transcript can be found in Appendices A and B. Upon completion of the guided imagery exercise, participants engaged in a failure induction task.

**Failure Induction Task.** Following previous studies, the Raven’s Progressive Matrices were used to induce feelings of failure in the current study (Belanger, Lafrienier, Vallerand, & Kruglanski, 2013; Stoeber, Hutchfield, & Wood, 2008; see Appendix C for the 11 items used). Given the link between feelings of failure and experiences of self-criticism, it was believed that
this failure induction procedure would also result in increased state self-criticism. The task required participants to complete a series of increasingly difficult spatial reasoning puzzles, after which participants' performance was evaluated and the participant was given false feedback indicating that they performed poorly. In standardized testing, the protocol contains sixty multiple-choice items. In order to prevent the priming effects from dissipating over the course of 60 cognitively demanding questions, 11 items were taken from across the range of the test so that easy, moderate, and difficult items were represented. After completing the questions, participants were told to press a button to “submit” their answers to the research assistant who pretended to mark their performance after which all participants were told “Compared to other students, you have scored in the bottom 20%. Sorry”.

**Measures**

**State Self-Criticism.** This measure was adapted from the Forms of Self-Criticism and Self-Reassurance Scale (Gilbert *et al.*, 2004). Whereas the original scale asks about typical ways of self-relating, the state version was adapted to ask about how participants were feeling towards themselves *in this moment*. The scale consists of 15 items. Participants rate their level of agreement on a 5-point Likert scale ranging from 0 (*Not at all*) to 4 (*Extremely*). A sample item reads, “Right now…I am disappointed with myself”. Scale reliability in the current the scale was very good (*α* = .94).

**Manipulation check.** Participants were asked to answer two questions about their participation as rated on a 5-point Likert scale ranging from 0 (*Not at all*) to 4 (*Definitely*): "I found that the imagery resulted in a change of mood" and "I felt upset after hearing that I did badly on the spatial reasoning test."
Results

The general aim of the current study was to examine whether priming compassion and attack-resistance through interpersonally oriented imagery exercises would result in lower levels of self-criticism after completing a failure-induction task.

Preliminary Analyses

Prior to analyses statistical assumptions relevant to correlations, linear regressions, and ANOVA were conducted. An initial visual inspection of histograms, Q-Q plots, and the matrix of scatterplots was undertaken to assess normality (i.e., skewness), potential outliers, and linearity. The presence of overly skewed data was assessed using the ‘skewness < 2 X std error’ rule (skewness $R = |0.03|$ to $|1.12|$; kurtosis $R = |0.78|$ to $|1.0|$; Cramer & Howitt, 2004). State self-criticism required a transformation (log10), after which it met criteria of normality. Potential outliers were identified using the ‘outlier labelling’ rule with a multiplier of 2 (i.e., 2 X the difference between the upper and lower quartiles (Hoaglin, Iglewicz, & Tukey, 1987). No outliers were identified. Scatterplots did not indicate any nonlinear relations. Examination of residual scores ‘skewness < 2 X std error’ rule indicated that the residuals was normally distributed. Levene’s test of homogeneity of variance was not significant ($F(2,111) = .22$, $p = .81$). As a result, all assumptions required for ANOVA were met.

Manipulation check - change of mood. The majority of participants reported that the imagery exercise "resulted in a change of mood": 63.1% reported the intensity of this change as 2 scale points or higher on the 5-point scale (2 = ‘Somewhat’, 3 = ‘Quite a bit’, 4 = ‘Definitely’). Participants also reported that they "felt bad" after receiving negative feedback: 88.1% reporting ‘a little bit’ (= 1) of distress or higher and 25.5% reporting ‘quite a bit’ (= 3) of distress or higher.
**Manipulation check - self-criticism induction.** A one-sample t-test was conducted to further examine the efficacy of the failure induction task. If the failure induction task were unsuccessful participants’ level of state self-criticism would be expected to be 0. To be conservative and to reflect the general levels of self-criticism within the population, a criterion of 1 (‘a little bit’ self-critical) was chosen as the comparison score. Results indicated a significant difference in state self-criticism compared to 1 ($t = 12.28$, $p < .01$) suggesting that participants experienced more state self-criticism than would be expected by chance alone and that the intensity of this self-criticism was more than just ‘a little bit’ self-critical. These results indicated that both the priming and failure manipulations were successful.

**H1: Compassion and defending groups will have lower self-criticism.**

Consistent with theory (Gilbert & Irons, 2006), it was hypothesized that individuals within the compassion and resistance imagery conditions would experience less self-criticism following failure compared to the control condition. A one-way ANOVA was conducted across the two experimental groups and the control group with state self-criticism as the dependent variable. Contrary to expectations, results of the ANOVA were not significant, $F(2, 111) = .55$, $p=.58$. This is likely unrelated to the low observed power (.14) as the effect size was very small ($\eta^2 = .01$) and the means in state self-criticism between the Compassion, Defending, and Control groups were similar: $M_{\text{compassion}} = 27.45(SD=10.29)$, $M_{\text{resisting}} = 28.84(SD=11.60)$, $M_{\text{control}} = 27.32(SD=10.32)$. In summary, no differences in state self-criticism were found between the three groups. Hypotheses 2 which sought to investigate the comparative treatment effect of the compassion and resistance primes could therefore not be undertaken.

**Discussion**

Gilbert’s Social Mentality Theory (2010) suggests that at the state level individuals enter into motive-oriented states of mind that organize the whole person and promote action.
tendencies that support the attainment of biosocial goals related to that system. Central to the current study is the notion that these mentalities (or motivational systems) are stimulus-independent, meaning that when in a particular state of mind, people will respond to themselves and others on the basis of the activated mentality rather than the nature of the stimulus as either interpersonal or intrapersonal.

The general aim of the current study was to examine whether priming compassion and resistance (using an interpersonal script) would result in lower levels of self-criticism after failing on a task. Although the priming procedures resulted in self-reported changes in mood and the negative feedback procedure elicited state self-criticism, differences were not found between any of the groups. The guided imagery priming exercise did not result in lower levels of self-criticism.

There are several potential reasons finding similar means in state self-criticism across the three groups. Of note, priming studies frequently fail to produce expected effects. There is an increasingly expanding body of literature suggesting that there is a strong publication bias in priming research with non-successful trials being suppressed (Francis, 2012). Nonetheless, assuming that the current results do actually diverge from previous findings, there are four salient accounts for the null results: the effect of the prime did not last, the prime was not strong enough, the outcome measure may not have been sufficiently sensitive, and perhaps more interesting, that using a compassion-giving script (self-to-other) does not reflexively transfer into self-compassion (self-to-self).

The first reason for the null findings might be that the effects of the prime dissipated during the spatial reasoning questions. The completion of the spatial reasoning questions may have undone the effects of the imagery prime because it requires participants to switch from the motive activated during the imagery task (care for others) to a spatial problem-solving motive
(being competent). If the motives and emotions that were primed during the imagery faded over the course of the spatial reasoning questions, it would not be surprising that no significant differences were found among the conditions. The spatial reasoning questions require participants to adopt an assertive motive of problem solving.

A second explanation might be that the failure induction procedure was stronger than the effect of the prime. Pilots of the imagery exercises indicate that they did produce a change in mood in the expected direction. Similarly, self-report during the study indicated that participants experienced a shift in mood after the imagery exercise. Nonetheless, this primed state of mind may have been weaker than the self-critical state of mind which was triggered by the failure induction procedure.

Alternatively, the outcome measure may not have been sufficiently sensitive to detect the variance between conditions. The measure of state self-criticism was adapted from a trait self-criticism measure designed to capture different forms and functions of self-criticism (Gilbert, Clarke, Hempel, Miles & Irons 2004). As a result, the questions may not have accurately captured critical self-to-self interactions in the moments following the failure induction task.

A fourth explanation for these findings highlights the importance of the target of the motive (self or other). Interpersonal theory suggests that people’s behavior differs depending on the stimulus as responding to others, making overtures towards others, or engaging in self-to-self interactions (Benjamin, 2003). From an interpersonal model priming a self-to-other action pattern compassion, for example, may not simultaneously activate a corresponding self-to-self action pattern (i.e. compassion towards the self). From an interpersonal model, these are different interpersonal foci that exist somewhat independent of one-another (Bjerke et al 2016); individuals don't always offer themselves the same compassion and sympathy they might offer to others. Attachment informed research and therapy has focused more on priming and imaginal
exercises involving a compassionate other that cares for the self (other-to-self focus; e.g., Mikulincer, Shaver, Gillath, & Nitzberg, 2005).

**The target of the prime may be more relevant than the activated social mentality.**

As noted above, past research has demonstrated that there is a transitive action of motive oriented states of where compassion can be increased to self and others simply by asking participants to imagine receiving compassion. Indeed, activating compassionate other-to-self (receiving compassion from others) results in higher levels of compassion to others (Mikulincer & Shaver, 2001; Mikulincer, Shaver, Gilrath, & Nitzberg, 2005) and activating compassionate other-to-self also increases self-compassion (Kelly, Zuroff, Leybman, & Gilbert, 2012). At the same time activating compassionate self-to-other is related to higher level of compassion towards others (Mikulincer et al 2001; Hutvherson, Seppala, & Gross, 2008).

Due to the relatively recent introduction of attack-resistance as a construct of import, no research has directly investigated the transitive action when priming either other resisting or self-resisting. Nonetheless, research from different yet related experiments suggests that a transitive action is likely to occur (Katzir, Meiran, & Kessler, 2010; Oosterwijk, Rotteveel, Fischer, & Hess, 2009; Wilcox, Kramer, & Sen, 2011).

Unique to the present study is the aspect of examining compassion and resistance giving (self-to-other) on one’s own self-relating (self-to-self). The results of the current study did not find that priming self-to-other compassion and resisting resulted in lower state self-criticism. The null results may indicate that the focus of the prime may be more relevant than the social mentality which was activated. Thus, the priming procedure may have induced a self-to-other role but once engaging in a self-to-self interaction the adopted role would be determined by one’s typical ways of self-relating (i.e., the target) regardless of the interpersonal motive that was primed. This can be understood from the relative positions of the felt sense of self as subject
When people experience receiving compassion or resistance, they adopt a self-referential position of being injured and deserving of help and so this position becomes predominant when state self-compassion is accessed (see below for discussion on self-referential position within dialogical self-to-self interactions). In other words, when people are primed using imagery in which they receive compassion (as is the procedure in past studies) their mental posture is that of being open to help and support. Self-compassion and attack-resistance also requires the person to have a mental posture of being open to help and support. Thus, primes in which one imagines receiving compassion may be effective because felt sense of self is in a position of receiving help. In contrast, when one is primed into giving compassion or resistance (as was the case in the present study), the felt sense of self has attention oriented towards others and away from the self. As such, one’s felt sense of self needs to change positions from giving compassion outwardly and focusing outwardly to giving and receiving self-compassion and focusing inwardly. The target of the behavior becomes more important than the motive because the target organizes the positions of the self as an object whereas motives organize the positions of the self as a subject (see below for a discussion on dialogical self-to-self interactions).

Limitations & future research. The current study was limited by the homogeneity of the sample, the use of only one order for the priming and failure tasks, and the use of a single outcome measure. Participants were primarily Caucasian, female, and in their early 20’s. This limits the generalizability of the results. Participants were recruited from a first-year psychology class. As part of the class participants were required to either volunteer for experiments or complete short writing assignments. As a result, many participants had already volunteered in a number of other studies and may not have been as motivated to fully immerse themselves in the imagery exercises or answer the questions with a high level of self-reflection (Schmeichel & Baumeister, 2004; Webster, Richter, & Kruglanski, 1996). Future research should use a more
heterogeneous sample who know less about priming procedures and false feedback on performance. Furthermore, the non-clinical nature of the sample may have also impacted results in that the levels of self-criticism were relatively low at baseline. It is possible that the primes may be more robust in individuals struggling with self-criticism.

The study was also limited by using only one sequence for the procedure. The current study used the procedure: (1) prime, (2) reasoning task, (3) failure induction, (4) outcome. It did not stagger participants between having the prime before and after the reasoning and failure feedback conditions. As such the study is unable to rule out a dissipation of the motive as a cause for the results. The study was also limited by the use of a single outcome measure.

It is possible that other more sensitive measures may have been able to detect differences in state self-criticism that the current measure did not. A shorter measure asking about self-critical experiences related to performance on the test may be more appropriate.

Future research should attempt to replicate these findings to confirm that the non-significant findings are not the result of a Type II error. Future replication attempts may benefit from altering the procedure slightly. Rather than (1) imagery, (2) spatial reasoning (3) negative feedback, (4) state self-criticism, future trials should ask participants to complete the spatial reasoning task prior to the imagery so that the effects of the prime are better preserved at the time of negative feedback. The current study used procedure it did because of concerns related to expectancy effects. If participants believed that their imagery condition should protect them from self-criticism there would be a risk of artificially inflating the results. If participants were asked to imagine themselves being compassionate or resisting and were immediately told they performed poorly, they may ‘see through’ the procedure. It would also be difficult to explain why they are completing the spatial reasoning task. In other words, there was a concern the alternative procedure would produce demand characteristics. Nonetheless, the procedure that was
used in the current study jeopardized the risk that the effects of the prime would persist until the outcome measure. The results of the prime may have been washed out by the spatial reasoning task resulting in false conclusions regarding the effects of interpersonal compassion and attack-resistance on state self-criticism.

Future research should attempt to clarify whether the transitive action of motive-oriented states of mind is constrained by the target of the behavior. This might help clinical applications of imagery exercises by better attuning the ameliorative effects of imagery to a client’s needs. The clarifying of the transitive action VS target of behavior could be accomplished by examining whether compassionate self-to-other primes result in lower levels of criticism-towards-others and whether a self-compassion task reduces self-criticism after failure. A third condition could examine the effects across targets. If the other-towards-other effects and self-towards-self effects are consistent and robust but the self-towards-other effects are non-significant it would suggest that transitive action of motive-oriented states of mind are constrained by the target.

Future research should use methodology similar to that used in Study 2 (the interview study) to further examine the dialogical process that unfolds after failure. Further examination of the intrapersonal dynamics following failure will help clarify how trait self-compassion might impact self-to-self relating at times of difficulty.
Chapter 3: Study 2 -

Clarifying self-submission, self-compassion, and attack-resistance through a mixed qualitative-quantitative investigation

Self-criticism is a transdiagnostic construct associated with numerous adverse psychological outcomes (Gilbert, Clarke, Hempel, Miles, & Irons, 2004; Shahar, 2001; Whelton & Greenberg, 2009). From the view of Gilbert's (2012) evolutionary model, global systems (dominance/drive, compassion/caregiving, and submission/threat-detection systems) form the conceptual components by which self-criticism is understood. For example, individuals with a low threshold for activating the submission/threat-detection system are likely to more frequently view themselves as low in social rank, inferior to others, which leads to self-criticism and depressive symptoms. Or an excessive, over-active dominance/drive system results in self-criticism and depression following failure to reach unrealistic goals (Kannan & Levitt, 2013).

The current study extends Gilbert's conceptual lens by including insights from interpersonal theory, which in contrast begins with the assumption of three different types of "focus": self-to-other, other-to-self, and self-to-self relations (e.g., Benjamin, 2003). Rather than viewing them as being organized by global social mentalities which are stimulus independent (Gilbert, 2015), this interpersonal model views different types of self-other relations as fundamentally distinct (Benjamin, 2003). For example, while perfectionism expected of others (self-to-other) may be related to self-perfectionism (self-to-self), their empirical relation has generally been only moderate (Benjamin, Rothweiler, & Critchfield, 2006). Similarly, that others may have high expectations of ‘me’ (other-to-self), it does not necessarily mean that I have high expectations of myself (self-to-self). Indeed, feelings of parental pressure contributes to both defiance (Vansteenkiste, Soenens, Petegem, & Duriez, 2014) and academic underachievement (Nagpal & Sinha, 2016).
Until recently Gilbert neglected to directly investigate a more dialogical conception of submission. Rather than assume that interpersonal submission (other-to-self) represented a global interpersonal trait which automatically implied self-criticism, the notion of “self-submission” entails one part-self submit to an inner critic part-self (Kelly, Zuroff, & Shapira, 2009). Currently research guided by Gilbert's Social Mentality Theory (2010) has used an interpersonal measure of submission as a proxy for ((or probabilistic indicator of) intrapersonal self-submission, or rather, assuming that interpersonal submission resulted in a sense of felt inferiority (i.e., self-criticism).

The current study examines whether a tendency to respond to interpersonal challenges with submission is related to a tendency to respond to self-critical statements by submitting or acquiescing to the self-attack. Since previous research has not explicitly examined self-submission (except two studies which coded content from a two-chair dialogue), the current study attempts to measure self-submission in a self-report questionnaire as well as a semi-structured interview. If successfully developed, these instruments may provide a more efficient means of measuring self-submission and its relation to other types of self-to-self interactions following failure. Implications for models of therapy (e.g., CFT; Gilbert, 2009, 2010) that have arisen from the self-compassion literature will be discussed.

Conceptualizing Self-Criticism and the Self that is Criticized

As noted in Chapter 1, self-criticism is interesting because it involves self-inflicted distress (being chastised) and it involves a self-to-self relationship wherein one is both a domineering attacker and (at least in that first moment of self-criticism) a submissive receiver of the attack (Freud, 1917). Thereafter, individuals can respond to the self-criticism using a number of different interpersonal strategies (further submission, compassion, attack resistance) again all within the context of the Self (Kelly, Zuroff, & Shapira, 2009).
Self-criticism, as explained in object-relations theories, has variously conceptualized self-criticism as an internalized parent-child relationship involving the “self’s” hatred and contempt towards an internal, identified-with “object” (Scharff & Tsigounis, 2003). The form of conceptual discourse varies, whether these subsequent internal relationships are spoken of as dynamic relations among “objects” or part-selves, or of different self-reflexive attitudes towards one and the same self. Interpersonally oriented investigators have begun to extend Gilbert's model by examining self-criticism as a self-to-self interaction, wherein the focus is less on the more distal interpersonal submission and more on the proximal submission to an inner critic (Kelly, Zuroff, & Shapira, 2009). That is, a conception of selfhood is examined in which internal part-selves interact and dialog with one another.

Interpersonal and Intrapersonal Relations

Informed by both Gilbert's evolutionary model (Gilbert, 2010) and interpersonal theory (e.g., Benjamin, 2003; Wiggins, 1996), the current study seeks to examine types of self-to-self interactions following self-criticism and their relation to interpersonal submission. As noted above, self-submission (further agreement with the self-criticism without any self-protective responses), self-compassion (regardless of the degree of agreeing with the self-criticism, providing care and self-soothing to the emotional impact of the critical message/content that drives the critical message), and attack resistance (disagreeing with the critical message/building oneself up) are examined as three potential responses to self-critical statements.

Self-Submission. As an analog, self-submission is viewed as interpersonal submission acted out internally in response to a domineering inner critic. Much like interpersonal submission, negative affect motivates a withdrawal from the contest (with one’s inner-critic) and feelings of defeat which promote withdrawal behaviors. Many studies examining self-submission have used interpersonal submission as a proxy variable (e.g. Gilbert & Proctor, 2006; Kopala-
Sibley, Rappaport, Lance, Sutton, Moskowitz, & Zuroff, 2013; Ongen, 2006). The use of an interpersonal measure as a probabilistic indicator belies a distinction in focus that should not be over-looked. Interpersonally there can be bids for dominance and there are many occasions where submission to the more powerful other is an appropriate response (e.g., when a superior makes a reasonable request that you do not want to do). Self-submission is similarly understood as taking a submissive role in response to a domineering inner-critic. This seems to be fundamentally different than submitting to a powerful other because in the context of self-criticism the individual is both the critic and the target.

The use of an interpersonal measure of submission would seem to be validated by findings showing that people who are high in interpersonal submission have more difficulty standing up to their inner-critic (Whelton & Greenberg, 2005) and that they feel powerless against their inner-critic (Gilbert et al., 2006). Using an interpersonal proxy may also make sense from a pragmatic perspective. When speaking about self-criticism and responses to self-criticism some scholars speak about a monological model of self-relating (e.g. Leary & Baumeister, 2000, “I think that I am inadequate”) while others use a more dialogical model of self (e.g. Kelly, Zuroff, & Shapira, 2009, “I tell myself I am inadequate”). Given the evolutionary model which informs the current study, it is expected that most people experience their self-criticism dialogically. The evolutionary model of self-submission involves a back-and-forth dialogue between an inner-critic and an inner-target. This back-and-forth could occur self-reflexively where one shifts between different “I” positions in an interchange or from a 3rd-person perspective, where one observes different part-selves interacting. One benefit of using an interpersonal proxy for self-submission is that it avoids potential conflicts (or confounds) associated with the choice of adopting either dialogical or self-reflexive language in writing self-report measurement items: “I believe I am bad” (monological) versus “I often fail to stand up for
myself when I am critical of myself” (self-reflexive) vs. “I often fail to stand up to my inner-critic” (3rd-person).

In contrast to viewing interpersonal submission as a potentially valid proxy for self-submission, research within the interpersonal and object-relations traditions suggest that self-to-self behavior manifests differently than other-to-self behavior (Benjamin 2003). Furthermore, research on self-critical personalities suggests that self-criticism motivates people to preserve their rank and status, making them less likely to submit to their inner-critic (Santor, 2003). These varied findings cast suspicion on the use of interpersonal submission as a proxy for self-submission, but more than that, they highlight a lack of conceptual clarity with respect to the nature of self-submission.

The current study endeavors to add conceptual clarity the forms of self-to-self relating following self-criticism by using a semi-structured interview to capture different forms of self-to-self responding. In the current study, self-submission is viewed as a response based in detecting threats where emotions and motives are oriented towards withdrawal and giving-in. The withdrawal is evidenced by agreement with an inner-critic where that agreement is characterized by defeat, powerlessness, and other threat-based emotions. Whereas interpersonal submission has clear evolutionary advantages, self-submission alone is not viewed as an adaptive response (see Table 1 for construct definition). The current study hopes to explore participants’ subjective experience of their self-criticism and to capture the dialogical self-to-self interactions which follow self-critical statements.

**Self-Compassion.** Two prominent models of self-compassion have emerged. Neff’s model (2003) places a focus on describing the features of self-compassion which she conceptualizes as involving kindness towards oneself, mindful awareness of one’s thoughts and feelings, and a belief that one’s shortcomings and mistakes are a normal part of the human
condition. Research within this tradition has demonstrated that self-compassion is inversely related to self-criticism (Gilbert & Procter, 2006; Judge, Cleghorn, McEwan, & Gilbert, 2012) and that increasing self-compassion reduces self-criticism (Gilbert & Procter, 2006) as well as other mental health difficulties (Neff & Germer, 2013).

A second model of self-compassion emerges from Gilbert's conception of self-to-self responding (2000) and is grounded within the motive of giving care (Gilbert, 2015). Within this model, self-compassion is viewed as an internalized care-giving script where one mentalizes about their own suffering and then acts in a way that acknowledges and soothes the distress (Liotti & Gilbert, 2011). This internalization model suggests that providing compassion to the self is largely undifferentiated from providing compassion to others (Gilbert & Procter, 2006). It is important to note that Neff’s (2016) and Gilbert’s (2015) models of self-compassion are compatible but that Neff focuses on the descriptive attributes of what self-compassion looks like, whereas Gilbert focuses on the motive thereby allowing more variants of self-compassionate behaviors so long as they are directed towards soothing and diffusing distress.

Following and extending on Gilbert’s definition, the current study views self-compassion similar to the attachment concept of a secure based script involving actions aimed at “diffusing distress,” although typically assessed in the attachment literature as the diffusion of other’s distress (Gilbert, 2015). The sensitive and responsive capacity to diffuse of one’s own distress will be coded in the interview component of the study as self-compassion. Of note, some authors suggest that self-criticism and self-compassion form a bi-polar continuum such that someone high in self-compassion is necessarily low in self-criticism (e.g. Neff 2003; Neff, 2016). Others argue that self-criticism and self-compassion are independent dimensions (Lopez et al, 2015) on which one can be high in both self-criticism and self-compassion. In line with interpersonal theory (Wiggins, 1998) and the results of Study 1, the current study adopts a model with self-
criticism and self-compassion being independent dimensions (Bjerke, Solbakken, Friis, & Monsen, 2016) where people can shift between being self-critical and then self-compassionate.

**Attack-Resistance.** More recently examined, attack-resistance is beginning to receive some research attention as an adaptive way of responding to self-criticism. It was first formally introduced in a brief self-directed intervention study in 2009 (Kelly, Zuroff, & Shapira). Attack-resistance is viewed as the ability to separate oneself from his/her self-criticism and stand up against it in a dominant and powerful manner. This model of self-to-self responding is based on the dominance-submission axis of the interpersonal circumplex (Wiggins, 1996). Those with higher levels of dominance are believed to have greater abilities to stand-up to their self-criticism rather than being beaten down by it (Kelly, Zuroff, & Shapira, 2009).

The ability to resist one’s self-critical messages is associated with lower levels of interpersonal submission and less harsh self-critical messages (Whelton & Greenberg, 2005). Self-critics with higher interpersonal submission are less able to separate themselves from the content of their self-criticism and were less able to resist their self-attacks (Gilbert et al 2006). Deficits in attack-resistance have been shown to be associated with adverse mental health outcomes. For example, individuals who submit and feel beaten down by their self-criticism rather than fighting back against it are also more likely to meet criteria for depression (Greenberg et al 1990). Research from cognitive therapies suggests that challenging one’s self-critical thoughts can be an important aspect of alleviating self-criticism and its relation to depression (Rector, Bagby, Segal, Joffe & Levitt, 2000). Symptom severity is also higher in individuals who experience themselves as powerless against their self-critical thoughts (Gilbert et al, 2001). These findings suggest that the ability to resist self-attacks may indeed be related to an internalized interpersonal style and that a greater ability to resist self-critical messages may be adaptive.
The construct of attack-resistance has not been well articulated in the literature and as a result the construct lacks conceptual clarity. To date the studies reviewed above forced a perspective of dominance-submission without looking at the horizontal access of the interpersonal circumplex (communion-detachment). As a result, they did not account for self-compassion as an alternative means of self-responding. Additionally, these studies treated self-submission as having complementarity with attack-resistance (e.g. high self-submission means low attack-resistance), however, if an evolutionary and dialogical model of self-to-self relating is accurate then an individual might engage in high levels of self-submission and then respond with high levels of attack-resistance.

In the current study, attack-resistance was understood as a form of responding where one resists their own self-criticism, from a position of dominance and drive, by directly contradicting their self-criticism, by focusing on their strengths, and/or by resuming goal directed behavior (see Chapter 1 for full discussion, Table 1 for construct definition). The current study will endeavor to add conceptual clarity by asking participants about their responses to self-critical statements. A semi-structured interview will be used to examine whether self-submission, self-compassion, and attack resistance do indeed capture the majority of responses to self-critical statements.

**Psychological awareness is an adaptive context for self-criticism and responses to self-criticism**

The intrapersonal model of the mind outlined above, does not necessarily imply that the individual is aware of their internal self-talk or exchanges between the various part-selves. Self-awareness has been an important construct in Western thought since Ancient Greece as indicated by the Delphic Maxim “know thyself” (Jopling, 2000) and in clinical psychology specifically, since Freud “where the id was, there shall the ego be” (1932). While psychologists have most
commonly spoken conceptually in terms of reflexive language, in which the self-as-subject regulates the self-as-object (e.g., I am critical of myself), clinicians have long seen the clinical value of “externalizing” problematic parts of oneself such as an “inner critical voice”.

Psychological awareness is understood as a metacognitive ability involving one’s awareness of their self-talk and the different ways in which they experience their mind (reflexively or dialogically) (Taylor, 1995).

In the present study, psychological awareness is defined as the degree to which a person can speak about their mental responses following self-criticism in a detailed and descriptive manner. Psychological awareness encompasses more than mindfulness because it allows for people to be highly fused to their mental content whereas mindfulness requires non-judgmentally observing one’s mental experiences. Further, whereas mindfulness is highly related to self-compassion and may be required to cultivate compassion (Tirch, 2010) and is even included as a component in some models of self-compassion (Neff, 2003), psychological awareness can be related to any form of self-to-self responding. For example, someone might ruminate in self-criticism and be deeply fused to the content of their self-criticism. If this individual were able to explain their agreement and submission fluently they would have high psychological awareness but low mindfulness (because their awareness emerges from frequent rumination and fusion rather than non-judgmental noticing). Further, interpersonal submission (a proxy for self-submission) is related to self-criticism and depressive symptoms (Ghaed & Gallo, 2006; Pearson, Watkins, & Mullan, 2010). It would not be surprising that individuals who spend considerable amounts stuck in self-critical rumination would develop a high level of awareness of their experiences of self-criticism and responses to self-criticism, especially if they look back on their rumination after the self-to-self interaction.
On the other hand, individuals who have robust and complex self-to-self interactions following self-criticism may have a high degree of psychological awareness because the stimuli related to these complex interactions are more salient (i.e. it is harder to be unaware of a complex self-to-self interaction than it is to be unaware of passive submission). Additionally, having awareness of the content of one’s thoughts and one’s relationship to oneself requires a high level of mentalizing ability which is associated with self-compassion (Liotti & Gilbert, 2011). Further, higher levels of psychological awareness is likely associated with a willingness or courage to turn towards one’s self-criticism and its associated aversive emotions which is viewed as an important quality of self-compassion (Gilbert, 2009) and is also a necessary step in separating oneself from their self-criticism in order to refute it (Rector, Bagby Segal, Joffe & Levitt, 2000). As such, it is likely that higher levels of psychological awareness will be associated with both self-compassion and attack-resistance.

In the present study, psychological awareness was conceptualized as a positive feature of one’s self-critical process because it is expected to be associated with adaptive forms of responding following self-criticism. To this author’s knowledge no research has examined people’s level of psychological awareness and how this might be related to self-to-self interactions following self-criticism. Further research is needed to clarify the construct of psychological awareness and examine its relation to different forms of self-to-self interactions following self-criticism.

**Aims of the Current Study**

Guided by an evolutionary model of self-criticism with insights from interpersonal theory, the current study is intended to clarify constructs central to the self-criticism literature: self-submission (to self-criticism), self-compassion, and attack-resistance, which are viewed as forms of self-to-self relating following self-criticism. In addition to using self-report
questionnaires, a semi-structured interview is used to capture the dynamic self-to-self interchanges that follow self-criticism. More specifically, the current study will:

**Aim 1:** Replicate previous findings in the literature that self-reported self-criticism, -compassion, and –submission predict negative affect and provide a detailed examination of the content of participants’ self-criticisms based upon the semi-structured interview.

**Aim 2:** Adapt a measure of interpersonal submission to reflect self-submission, reflecting problems with resisting or standing up to one’s inner critic, and explore whether it correlates with self-criticism and self-compassion in a similar way as interpersonal submission.

**Aim 3:** Ascertain whether a measure of self-submission would be more strongly correlated with negative affect interpersonal submission.

**Aim 4:** Examine whether participant responses to self-critical statements can be coded for self-criticism, self-submission, self-compassion, and attack-resistance.

**Aim 5:** Capture the more dynamic aspects of intrapersonal responses to self-criticism, and examine which patterns of self-responding are most associated with negative affect.

**Aim 6:** Test the prediction that awareness is associated with the use of self-compassion and resistance in response to self-criticism.

**Methods**

**Participants**

Participants were 57 undergraduate students with a mean age of 20.49 years ($SD = 2.85$) and who were predominantly female (77.19%) and Caucasian (91%). Participants were recruited from 1st and 3rd year psychology courses from a university in Southwestern Ontario. Ethics clearance was granted from the University of Guelph Ethics Review Board.

**Procedure**

Participants came into a meeting room where the limits of confidentiality and elements of informed consent were discussed. If participants agreed to participate they began by completing a pre-study test of mood (PANAS State). Subsequently they completed a battery of computerized self-report questionnaires in private with a research assistant nearby. When they were done the questionnaires, they completed a short semi-structured interview about their self-criticism and responses to self-criticism (see below). Participants then watched a short video of baby animals.
as a mood repair and completed a post-test of state affect (PANAS State). There were no significant differences in pre-post mood suggesting that the procedure was not harmful ($t(56) = .85, ns$).

**Measures**

**Self-report measures.**

*Self-criticism.* Self-criticism was measured using the Forms of Self-Criticism and Self-Reassurance Scale (Gilbert *et al.*, 2004). The self-criticism subscale comprises 14 items. Participants rate their level of agreement on a five-point scale ranging from 0 (not at all like me) to 4 (Extremely like me). A sample item reads “I find it difficult to control my anger and frustration at myself”. The scale has demonstrated adequate validity and reliability in several studies (e.g. Gilbert *et al.*, 2004; Gilbert *et al.*, 2006; Longe *et al.*, 2010). This scale measures different forms of self-criticism as separate constructs; however, in the current study these items were amalgamated to form a single measure of self-criticism. In the current sample the internal consistency of the scale was adequate ($\alpha = .91$; item-total correlations $\geq .40$).

*Self-compassion.* Two measures of self-compassion were used in the present study. Self-Compassion (1) was measured using the Self-Compassion Scale (Short Form) (Neff, 2011). The Self-Compassion Scale Short form comprises 12 items measuring both self-compassion and non-compassion. Recent factor analytic studies have suggested that self-compassion and non-compassion are two separate factors with the non-compassion (reverse) items measuring a construct similar to self-criticism. In the current study, only the positive self-compassion items were used to avoid inflating relations with self-criticism and the negative outcomes variables.$^1$

---

$^1$ By way of exploration, we examined the patterns of correlations between the self-compassion, non-compassion, and total score with other variables in the study. An examination of zero-order correlations where the reverse self-compassion items were added did increase (conflate) correlations between self-compassion and both self-criticism and negative affect. As such, the six self-compassion items were used to measure Self-Compassion (1).
Participants rated how often they behave in a particular manner on a five-point Likert scale ranging from 1 (almost never) to 5 (Almost Always), a sample item reads “I try to see my failings as part of the human condition”. The Self-Compassion Scale Short Form has demonstrated adequate reliability and validity in a number of studies (e.g. Raes, Pomier, Neff, & Van Gucht, 2011; Smeets, Neff, Alberts, & Peters, 2014). In the present study the scale demonstrated adequate internal consistency ($\alpha = .86$, inter-item correlations $\geq .4$).

A second measure of self-compassion was used. Self-Compassion (2) was measured using the Reassured-Self subscale from the Forms of Self-Criticism and Self-Reassurance Scale (Gilbert et al., 2004). The subscale comprises nine items. Participants rate their level of agreement on a five-point scale ranging from 0 (not at all like me) to 4 (Extremely like me). A sample item reads “I am gentle and supportive with myself”. The scale has demonstrated adequate validity and reliability in a several studies (e.g. Gilbert et al., 2004; Gilbert et al., 2006; Longe et al., 2010). In the present study, the scale demonstrated good internal consistency ($\alpha = .86$, inter-item correlations $\geq .4$ other than one item whose IIC = .38).

**Interpersonal Submission.** Two scales were used to measure interpersonal submission. Interpersonal Submission (1) was measured using the Submissive Behavior Scale (Allan & Gilbert, 1997). This is a 16-item scale asking participants to report the degree to which a series of statements are true for them on a five-point Likert scale ranging from 0 (never) to 4 (always). A sample item reads “I let others criticize me or put me down without defending myself”. The scale has demonstrated adequate validity and reliability in a several studies (e.g. Connan, Troop, Landau, Campbell, & Treasure, 2007; Sturman, Rose, McKeighan, Burch, & Evanico, 2015). In the present study, the scale demonstrated adequate internal consistency ($\alpha = .80$; 3 items demonstrated item-total correlations $\leq .39$ but their removal would significantly increase $\alpha$).
Interpersonal Submission (2) was measured using Inventory of Interpersonal Problems-Youth Version (IIP-Y; Hennig & Browne, 2017). The measure assesses participants’ interpersonal problems and is based on the original 64-item adult Inventory of Interpersonal Problems-Circumplex (IIP-C; Horowitz, Alden, Wigens, et al, 2000). Items measuring interpersonal problems with over-submissive behavior were assessed using the Submission Octant subscale. Participants however, completed the full eight circumplex subscales of the IIP-Y. Each subscale contains 6 items rated a 5-point Likert scale ranging from 0 (not at all) to 4 (extremely). A sample submission item reads “I find it hard to be firm when I need to be”). Owing to the circular structure of the IIP-Y (like other circumplex measures), items are first ipsatized to remove a general, interpersonal distress factor (see fuller description below in the Results section). That is, prior to ipsatization, individuals who report problems on one subscale, tend also to report problems in the opposite forms of problem (e.g., dominance and submission or on opposite locations along the circle). Ipsatization removes the general distress factor. In the present study, the submission sub-scale demonstrated good internal consistency (α = .82, item-total correlations all ≥ .40).

**Self-Submission/attack-resistance.** Self-submission was measured using the Self-Submission Scale, a new measure designed for the current study. Items were based on the submissiveness octant of the Inventory of Interpersonal Problems-Circumplex (IIP-C). Items were adapted to reflect difficulties standing up to one’s self-criticism (example item reads “I find it difficult to stand up to my inner-critic”. In this way, the measure assumes a dialogical model of the self where one shifts between being domineering and critical towards the self, followed by potential intrapersonal (self-to-self) responses. The measure captures the degree to which participants experience problems with being overly submissive towards their self-criticism.
Psychometric details of the scale are provided in the Results section. Of note, the psychometric properties of the Self-Submission Scale were all excellent (α = .90, item-total correlations ≥ .40).

**Negative Affect – State.** State negative affect is used as an outcome variable in the present study. It was measured using the Negative Affect Subscale of the PANAS (Watson et al., 1988). Participants are asked to complete 10 items using a Likert scale ranging from 1 (not at all) to 5 (extremely). A sample item reads “indicate to what extent you are feeling this way right now… upset”. The scale has demonstrated adequate validity and reliability in a several studies (e.g. Brown & Ryan, 2003; Gross & John, 2003). In the present study, the scale demonstrated adequate internal consistency (α = .78; 3 items demonstrated item-total correlations ≤ .39 but their removal would significantly increase α). Of note, this is the same measure used at pre-test to examine whether the study caused distress in participants.

**Negative Affect – Trait.** Trait negative affect was measured using the Neuroticism subscale of the Big Five Inventory of personality (John, Naumann & Soto, 2008). The subscale asks participants to complete eight items asking about their self-perceptions in a variety of situations on a 5-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). The scale has demonstrated adequate validity and reliability in a several studies (e.g. John, Naumann & Soto, 2008; Srivastava, Gosling, & Potter, 2003). In the present study, the scale demonstrated adequate internal consistency (α = .89; all item-total correlations ≥.4)

**Interview and Constructs**

**Self-Critical Process Interview (SCPI; Cwinn & Hennig, 2017).** The Self-Critical Process Interview is a semi-structured interview that asks participants about their self-critical statements and responses to self-critical statements. They were first asked to think of times when they remember being self-critical and to focus on those memories. They were then asked the specific self-critical statements (self-talk) they made to themselves. They were subsequently
asked “after you are self-critical, what happens next in your mind?” To elaborate on the constructs of interest they were then asked whether they ever submit to their inner-critic, soothe themselves while acknowledging the criticism, or resist the self-criticism while building themselves up. The questions are structured such that they begin open-ended and, if they are unable to provide a response, further prompts are given to provide more guidance. The author and trained undergraduate research assistants who conducted and transcribed the interviews. Two coders familiar with the evolutionary model of self-criticism and self-compassion rated 31 (54%) interviews to evaluate inter-rater reliability. Inter-rater reliabilities are reported in the Results section, to the question of whether self-critical dialogues could be reliably coded using a metatheoretical interpersonal and evolutionary model of self-criticism.

Transcripts were micro-coded to capture each sentence segment that represented a full thought relating to self-submission, self-compassion, or attack-resistance. Additionally, interviews were macro-coded using dimensional ratings to capture the degree to which participants employed the different forms of self-to-self relating (self-submission, self-compassion, and attack-resistance) using a 10-point Likert scale ranging from 0 (not present) to 9 (very much present) (see Table 1 for construct definitions). Finally, interviews were coded using a categorical approach to capture which form of self-to-self responding is most salient for each participant. The midpoint of the dimensional rating (>=7) was to be used as a cut-score to sort participants into different categories. Participants who scored over 7 in multiple categories were to be sorted into a Mixed category. See Appendix D for an outline of the interview, and Appendix E for coding instructions. More details on the psychometric properties and coded material is provided in the Results section.

*Psychological Awareness.* Participant’s level of psychological awareness was evaluated on the basis of the number of prompts they required to adequately describe their mental
experiences following self-criticism. Coders applied a decision tree based on the number of prompts to code participants’ level of awareness. This decision tree can be seen in Appendix F.

- 5 = no prompts needed, detailed discussion of their internal experiences provided;
- 4 = no prompts needed, however responses lacked detail;
- 3 = 1 prompt needed;
- 2 = 2 prompts needed;
- 1 = ≥ 3 prompts needed;
- 0 = little description of internal processes provided.

A 5 was given if participants did not need any prompts and were able to give a detailed account of their internal experiences following self-criticism; a 4 was given if they did not need any prompts but their account of responses after self-criticism was lacking in detail and required follow-up questions; a 3 was given if participants needed at least 1 prompt; a 2 was given if the participant required 2 prompts they receive a score of 2; if they required 3 or more prompts they were scored as 1; and, if they were unable to provide an account of how they respond to their inner critic they were given a 0. Regarding reliability, two raters familiar with the evolutionary model of self-criticism and the literature on self-compassion each coded 31 participant transcripts (54% overlap). Intra-Class correlations using a two-way random effects model (Hallgren, 2012) was very high ($r = .95$, $p<.001$).

**Results**

**Preliminary Analyses**

Prior to analyses, statistical assumptions relevant to correlations and linear regressions were conducted. An initial visual inspection of histograms, Q-Q plots, and the matrix of scatterplots was undertaken to assess normality, potential outliers, and linearity. The presence of overly skewed/ kurtotic data was assessed using the ‘$< 2 \times$ standard error’ rule (results: skewness $R = |.79|$ to $.00$; kurtosis $R = |.77|$ to $.05$; Cramer & Howitt, 2004). The following variables required transformations as indicated in parentheses: self-compassion (2) (log10); interpersonal submission (1) (log10), unipsatized interpersonal submission (2) (log10) and state
negative affect (log10). Potential outliers were identified using the ‘outlier labeling’ rule with a multiplier of 2 (i.e., $2 \times$ the difference between the upper and lower quartiles; Hoaglin, Iglewicz, & Tukey, 1987). Analyses were conducted with and without the outliers and reported below. Scatterplots did not indicate any nonlinear relations.

Descriptively, examining relations between negative affect and the predictors, levels of trait negative affect (Big 5 neuroticism) were examined. In the current sample, 50.9% reported “slightly agreeing” or “agreeing” and 14.0% strongly agreed that they experienced elevated negative affect (e.g., “worry,” “depression”; see Table 2 for $M/SD$s). No gender differences were found on either state or trait negative affect as the independent variables, $F$s(55) = .31 & .02, $p$s = .58 & .88, respectively.

Aim 1: Replicate previous findings that self-criticism, -compassion, and –submission predict negative affect

Replicating previous studies, bivariate correlations reveal that self-criticism and submission were positively associated with trait negative affect, whereas the two measures of self-compassion were negatively associated with trait negative affect (see Table 2).

A content analysis was conducted from transcriptions of the semi-structured interview (for further details see below), examining the reported events that precipitated participants’ self-criticism. The content analysis was conducted using the procedure recommended by Renner and Taylor-Powell (2003), resulting in the creation of 9 categories plus an addition 10th “other/idiosyncratic” category. Over the duration of the interview, on average participants endorsed 2.46 ($SD = 0.80$) different events that triggered self-criticism. The most common events triggering self-criticism were: failing to meet academic expectations (70.2%), poor study habits (35.1%), relationship dissolution (28.1%), and acting in a socially awkward manner (24.6%; see Table 3 for complete results). There were 14 participants that reported events coded in the
“other/ idiosyncratic” category: 9 pertained to life style values (e.g., not having sex outside a committed relationship, smoking marijuana less), 3 were related to carelessness (spilling things, difficulty focusing, spitting while talking); and 2 were related to difficulties with emotion regulation (e.g., anger management difficulties).

**Aim 2: Explore whether an adapted a measure of intrapersonal submission, which asks about problems with resisting or standing up to one’s inner critic, correlates with self-criticism and self-compassion in a similar way as interpersonal submission**

To date the literature has used an indirect or distal measure of interpersonal submission as a proxy for self-submission (i.e., intrapersonal submission). The newly developed, final 9-item measure of self-submission demonstrated excellent reliability (2 items were dropped for failing to meet the item-total correlations > .40 criteria, final $\alpha = .91$). By way of examining preliminary indices of construct validity, correlations between self-criticism, self-submission, and self-compassion were as expected: positive correlations were found between self-submission and self-criticism, and negative relations were found between self-submission and self-compassion. (see Table 2). The relation between self-submission and interpersonal submission is examined in Aim 3.

**Aim 3: Evaluate the measure of self-submission as a stronger predictor of negative affect than an interpersonal proxy**

Following from Aim 2, it was expected that the newly developed direct, intrapersonal measure of self-submission (Aim 2) would be both weakly correlated to the interpersonal domain, and secondly, would be a more strongly correlated to negative affect compared to the proxy measure. First, and challenging the premise that the intrapersonal domain represents a simple internalization of the interpersonal domain, neither self-reported nor interview-coded measures of self-submission were related to interpersonal submission (see Table 2).
Second, and seemingly contrary to the current expectations, both measures of interpersonal submission were more strongly correlated with self-criticism and negative affect than was the direct measure of self-submission (that is, submission to one’s inner critic). Note however, that in the interpersonal literature, circumplex scale items are first ipsatized (individual item scores are standardized using the individual’s mean and SD prior to computing subscale scores). Ipsatization removes the general interpersonal distress (neuroticism) variance, that is, the tendency for individuals to report interpersonal problems of various and even opposite forms (e.g., “I dominate others too much” vs. “I submit to others too much”; see Hennig & Brown, 2017 for scoring details). After ipsatization (removal of the shared neuroticism variance), interpersonal submission no longer predicted negative outcomes ($r$ fell from .51 to .15). Replication of the effect was approximated by partialling out the interpersonal distress factor of the IIP-Y from Gilbert and Allan (1994) interpersonal measure, the correlation fell from .55 to .27 (see Table 2). The unipsatized interpersonal measure (IIP-Y) and Gilbert’s measure otherwise have very similar predictive patterns: negatively associated with “positive” measures (self-compassion) and positively associated with “negative” measures (self-criticism). Given the dominance-submission model of self-criticism, it should be highlighted that self-submission was a stronger (and significant) correlate of self-criticism compared to both of the interpersonal measures (see the respective ipsatized and partialled values in parentheses; $r$s of .20 & .13 versus .29 for self-submission).

**Aim 4: Use a semi-structured created for this study to examine whether participant responses to self-critical statements can be coded for self-criticism, self-submission, self-compassion, and attack-resistance.**

A semi-structured interview was developed to capture the more qualitative, dynamic self-to-self responding to one’s self-criticism. Guided by an evolutionary model and a review of the
literature, a coding scheme was developed (see Appendix E). After describing an incident that prompted self-criticism, the interviewer asks the participant “[Then] what words go through your head?.... What goes through your mind after that?... What else would you say?” The interviewer’s aim was to obtain the most detailed account of participants’ experiences of self-to-self interactions following self-criticism, with some participants requiring a greater number of additional prompts (see discussion below re. the coding of “psychological awareness”). The interviews ranged from just over 12 min. to almost 22 min. \((M = 20.4 \text{ min}., SD = 6.0 \text{ min.})\).

The initial approach to coding was to begin with a micro-coding strategy, defining a codable unit as a unit of speech that began with one of the four coded Table 1 categories (self-compassion, self-submission, etc.) terminated by initiating speech reflective of another codable category or a non-codable category. Results indicated several codes being assigned to each transcript with numerous shifts back and forth between codes. That is, participants’ responses were very dynamic. By way of illustration, in the self-talk sequence "resisting, non-codable content, resisting, submissive, resisting, compassionate, submissive, resisting, non-codable content, resisting, submissive" (ID#27), the number of exchanges reflects no straightforward or simple pattern. With no apparent subset of underlying patterns, (e.g., self-criticism was always followed by either self-compassion or resistance), a global dimensional coding scheme was adopted instead. Global ratings were assigned from 0-9 on self-criticism, self-submission, self-compassion, and attack-resistance. Raters coded participants on each of the dimensions. Self-submission, which was conceptualized as agreement with the self-critical statement without any self-protective responses. Self-compassion was coded on the basis of the provision of care or soothing to oneself especially in terms of making statements related to forgiveness, kindness, and loving concern. Self-compassion was also coded when there was an acknowledgement of the error in a non-judgemental and grounded manner. Attack-resistance was coded on the basis of
refuting the content of the self-critical statement, by blaming the harm on others, or by remarking on one’s strengths.

In terms of reliability, two raters familiar with the evolutionary model of self-criticism and the literature on self-compassion each coded 31 participant transcripts (54% overlap). Coders provided dimensional ratings for each of the 4 global scales (e.g., self-compassion). Good reliability was found for each of the four dimensions: percent coder agreement (±1 scale point) ranged from 67.7% (self-compassion) to 90.3% (self-criticism). While agreement was somewhat lower for self-compassion, there were never more than 3 participants who differed greater than ± 2 points on any of the coded scales.

While strong relations across methods (self-report vs. interview-coded measures) were not expected, “validity” correlations were: .48 (self-criticism), .14 (self-submission; relations with interpersonal submission were also low), and .29 & .53 (self-compassion 1 & self-compassion 2, respectively; see Table 2).

Coders were also asked to rate whether the participant responded in a way that reflected a discussion within themselves (e.g. I tell myself) or whether they reported their experience in a manner suggesting a coherent self-narrative. In the initial review of this question the coders clarified whether an assumed listening-subject should be considered as a dialogical interaction and the raters decided that it would be. Coders rated monological responses as 0 and dialogical responses as 1. There was 96.77% agreement (30/31).

**Self-submission.** Participants who were rated as having high self-submission tended to have one or both of two types of responses. Either they actively agree with their self-critic and know it to be true or they had little follow-up self-dialogue after the self-critical statement. When prompted and cued to different types of responses (e.g. “do you ever say things to yourself afterwards to make yourself feel better?”; “do you ever fight back against your inner-critic and
find reasons to justify why you messed up?”) participants who scored high in self-submission tended to deny such responses and would eventually say that they acquiesce to the critic. Unfortunately, the interviews did not follow-up to ask whether they experienced this submission as submission or as a unified sense of self. Given that the coding scheme was constructed on an a priori definition of self-submission that involved the features of actively agreeing or acquiescing with the inner-critic it is not surprising that these types of responses were common in participants rated as high in self-submission.

**Self-Compassion.** In line with the construct definition and coding scheme, participants who were rated as high in self-compassion tended to acknowledge the harm or fault and would follow-up with statements of self-care. Interestingly, several participants who scored high in self-compassion reported that they would first seek support from a loved one and would later say soothe themselves. At the level of codable units, there was some variability in whether statements involving a sentiment “it is not your fault” were highlighted as self-compassion or attack-resistance. Nonetheless, the global quantitative ratings still had high inter-rater reliability regardless of the coding of the specific codable unit. Other sentiments which were expressed in participants rated as high in self-compassion included ideas like everyone makes mistakes, you are only human, in a week from now you won’t care as much, and it’s alright/ it’s OK.

**Attack-Resistance.** Participants who were rated as high in attack-resistance most often blamed circumstances or others, such as, saying that a test did not actually test what it was supposed to, blaming roommates for interfering with studying, or saying that their lifestyle interfered with their ability to work towards the goal (with the lifestyle being framed as outside their control). Less frequently participants rated high in attack-resistance also spoke about remounting a goal pursuit by saying things like I will do better next time; I will speak to the professor to try and improve my grade, I will practice more. Interestingly, many participants
who remount the goal made note of a delay between their self-criticism and the remounting of the goal. Lastly, participants who were rated as having high attack-resistance would also minimize the cost of the failure/setback that triggered their self-criticism by speaking about their strengths saying things like *I am a good student; I am strong; I am very likeable* even when these statements seemed unrelated to the context of the self-criticism. Again, the presence of these themes is expected because the construct definition and coding scheme were designed to capture these themes.

**Aim 5: Capture the more dynamic aspects of intrapersonal responses to self-criticism and examine which patterns of self-responding are most associated with negative affect**

**The dynamic aspects of intrapersonal responding: self-to-self interactions are dialogical.** All participants described their responses to self-critical statements using dialogical terms. As noted above, there was disagreement about one participant who spoke monologically. After discussing the case, both coders eventually decided that it was indeed dialogical and that there was an assumed audience. This high rate of dialogical responding is not surprising because the interview led people to describe dialogical responses “after being self-critical, what happens next in your mind” which may have elicited dialogical responses.

**Person-centered analyses with the interview-coded data.** It is not surprising given that participants were explicitly asked to discuss instance(s) in which they were self-critical, that slightly over 50% of the participants scored high (≥ 7) on both the self-criticism and self-submission dimensions (see Table 4). Only half of the self-critical individuals were also high on submission, that is, a person can be self-critical without necessarily submitting to their criticisms. Only approximately one quarter and one third of participants scored high on compassion and resistance, respectively.
The largest theme emerging from the interviews was that of being self-critical and either largely agreeing with their inner-critic (submission), or having weak ameliorative responses to self-criticism (compassionate or attack-resistant). Only six participants (10.5%) were high on both self-compassion and resistance. As a correlation, the number of ameliorative response categories that were in the high category (0 = none, 1, 2) were correlated with self-report measures: negative affect (r = -.37, p = .005), self-criticism (r = -.34, p = .01), self-compassion_1 (r = -.41, p = .001), self-compassion_2 (r = .31, p = .02), but did not predict the newly developed measure of self-submission (r = .07, p = .64).

**Variable Centered Hierarchical Regression Analyses.** Self-compassion and attack-resistance were examined as potential moderators of the relationship between self-criticism and negative affect. Hierarchical regressions were conducted, first for the self-report and then the interview data. The homoscedasticity assumption for multiple regressions was confirmed using the Koenker test (appropriate for smaller samples) with trait negative affect as the dependent variable for both the self-report predictors, $X^2(3) = 2.99, p = .39$, and for the interview data, $X^2(4) = 2.99, p = .56$. Negative affect was entered as the dependent variable and self-criticism entered at Step 1.

Examining the *self-report predictors*, results indicated that after accounting for 46% of the variance in predicting negative affect from self-criticism entered at Step 1, self-compassion and resistance entered at Step 2 accounted for a further significant 6% (see Table 4). Of greatest interest was the ‘Criticism X (Submission X Compassion)’ interaction entered at Step 3. The ameliorative responses (low submission and high compassion) were especially effective at reducing negative affect among those high in self-criticism.

For the *interview-coded data*, self-criticism entered at Step 1 account for 16% of the variance in predicting negative affect, and the two ameliorative responses entered at Step 2
contributed a further 14%. The ‘Criticism X (Submission X Compassion)’ interaction entered at Step 3 was not significant, leaving an independent main effect only for Compassion (see Table 4).

**Aim 6: Psychological-Awareness as a Predictor of Negative Affect**

A global dimension of psychological awareness was coded for each participant on the basis of the number of prompts they required in order to describe their self-to-self interactions in a fulsome manner. Higher scores (1- 5) were given to interviews where few prompts were required to enlist participant discussion and a “detailed awareness of the dialogical sequence” was demonstrated.

Results indicate that 43.9% of participants had a high level of awareness regarding their self-critical process (4 or 5). 24.6% had a moderate level of awareness, requiring only one prompt. A large minority of respondents had low levels of awareness, requiring two or more prompts (19.8%) and one participant was unable to provide an account of their manner of responding to their inner-critic.

Participants with higher ratings of psychological awareness was correlated with interview coded self-compassion (r = .37, p <.05) and resistance, (r = .37, p<.01). In relation to the self-report data, psychological awareness also corelated to self-reported compassion 1 (r = .27, p = .045) (Neff’s measure), but not compassion_2 (r = .01, p = .96) (Gilbert’s measure) or self-submission/ attack-resistance (r = .06, p = .66) (the newly adapted measure).

**Discussion**

Guided by evolutionary models of self-criticism (Wiggins, 1998; Gilbert, 2010) the current study examined self-criticism and self-to-self responses to self-criticism using a novel mixed qualitative-quantitative design. The general aim of the study was to clarify the constructs of self-submission, self-compassion, and attack-resistance. By using data from both self-report
and semi-structured interview, this study was better able to capture the ways in which people respond to self-critical statements.

**Replicating Previous Findings**

Results of bivariate correlations were consistent with those observed in past research with self-criticism and interpersonal submission being positively correlated with negative affect (Dunkley, Zuroff, Blankstein, 2003) whereas the two measures of self-compassion were negatively associated with trait negative affect (Hollis-Walker & Colosimo, 2011). The current study also identified a failure to meet academic expectations, poor study habits, relationship dissolution, and acting in a socially awkward manner as the most common triggers for self-criticism. Interestingly, these could be organized along the domains of agency and communion (Wiggins, 1998), highlighting the relation between the interpersonal world and how we interact with ourselves (Kelly, Zuroff, & Leybman, & Gilbert, 2012; Kopala-Sibley, Rappaport, Sutton, Moskowitz, & Zuroff, 2013).

**Examining an adapted measure (self-submission) and interpersonal submission in relation to self-criticism and negative affect**

Aims 2 and 3 examined an adapted measure of interpersonal submission to reflect self-submission, that is, difficulty in standing up to one’s own self-critic. Previous research has suggested that self-submission is a pathogenic form of self-relating (Bergner, 2013). Although not explicitly framed as such, studies examining this have used measures of interpersonal submission as a proxy variable to capture self-submission (e.g. Castihlo, Pint-Gouveia, Amaral, & Duarte, 2014; Kelly, Zuroff, & Shapira, 2009; Ongen, 2006). The current measure was adapted from the submissiveness octant of the Inventory of Interpersonal Problems (Horowitz, Alden, Wiggins, et al, 2000). The phrasing of the items was dialogical and asked participants to rate their difficulties in standing up to an internal critic. While the measure of self-submission
demonstrated good internal consistency, it did not correlate with either measure interpersonal submission. Interview measures of self-submission were also unrelated to interpersonal submission.

One reason for this may be related to cognitive aspects of participant responding. With respect to the self-report measure, if people experience their self-criticism as part of a coherent self (i.e., monologically) then from their phenomenological perspective they would not experience self-submission as submissiveness but rather as an ego-syntonic reaction to their thoughts about themselves as flawed. For example, if you really do think you are lazy when you say “I am so lazy” and agree with it you may not interpret this as submission. From this perspective, more monologically oriented individuals are less likely to endorse difficulties standing up to an inner-critic because they their sense of self is more coherent and undifferentiated. As a result, even highly submissive individuals (interpersonally) may rate themselves as having low levels of self-submission because agreement with self-criticism is not experienced as submission per se but rather, as a part of a coherent self-construal.

Other cognitive aspects may have similarly confounded the interview measure of self-submission. In order to answer the questions accurately, participants would have had to have sufficient self-awareness to not only recognize the ways in which they respond to themselves, but also the ability to mentalize about their experiences in such a way where they take on the position of a subject who responds to an internal-critic that is somehow separate from the self. As such, answering the questions required a high level of psychological awareness on behalf of the respondent. Given that 56.1% of participants required at least 1 prompt to answer the question in the interview, it is possible that the newly constructed measure is measuring something more like attitudes towards self-submission rather than self-submission itself. Supporting this explanation, participants who required no prompts demonstrated a strong
positive but non-significant correlation (due to sample size) between self-submission and interpersonal submission (1) and interpersonal submission (2) (rs = .34 and .29 respectively). Using a variety of different ways of depicting self-submission may have placed a lower burden on participants and may have therefore resulted in a more valid measure.

A second explanation for the lack of correspondence between self-submission and interpersonal submission is that an individual’s intrapersonal style does not mirror their interpersonal style. As noted above, those working from interpersonal and object relations perspectives suggest that our interpersonal and intrapersonal relations exist on three facets: self-to-other, other-to-self, and self-to-self (Benjamin, 2003). An individual’s behavioral tendencies are believed to be organized by the target of the behavior rather than the motive and state of mind and are therefore expected to differ across the three facets. Moreover, research in this field suggests that people tend to adopt one of two orientations that are almost orthogonal to one another: self-relatedness (vertical axis of the interpersonal circumplex) and dependency (horizontal axis of the interpersonal circumplex) (Kopala-Sibley, Zuroff, Hermanto, & Joyal-Desmarais, 2015). Those with high self-relatedness may have low interpersonal submission but high ego-syntonic self-submission. People with high self-relatedness are believed to be motivated by agentic needs to strive and achieve and their self-worth is based on success (Kopala-Sibley & Zuroff, 2014). These individuals would be less likely to naturally adopt a submissive role with others and would therefore report low levels of interpersonal submission. At the same time, at its extreme this model of self-and-other can have an implicit assumption that “I am not good enough as I am, therefore I need to achieve to demonstrate my worth” which is inherently subjugating. Thus, when engaging in self-to-self interactions following self-criticism it would not be surprising that these individuals would respond with high self-submission (though they would not experience it as submission but as a model of self). In contrast, those
high in dependency may have higher interpersonal submission and low self-submission. Having high dependency can be understood as being motivated to form bonds and get along. Their self-worth and sense of safeness is based in belonging rather than achieving. Those oriented towards dependency are more likely to be submissive interpersonally but will also have robust interpersonal scripts related to caretaking and nonjudgment (Kopala-Sibley & Zuroff, 2014). In self-to-self interactions, then, such individuals may be more likely to respond with compassion in addition to submission and would therefore rate themselves as lower on measures of self-submission. In concert these different interpersonal and intrapersonal orientations may cancel each other out leading to a non-significant correlation between interpersonal submission and self-submission.

**Examining whether participant responses in a semi-structured interview can be captured by self-submission, self-compassion, and attack-resistance.**

As noted above, every self-critical statement involves a degree of submission because the felt sense needs to ‘take in’ negative feedback from the world which is being delivered via the self-critic. After this initial submission, individuals can respond in a number of ways. Self-submission, as it is understood in the current paper, involves an affirmation or agreement with the self-critic. In the model guiding this study, submission involves a level of withdrawal and defeat. If over-used, this pattern of self-relating may actually be inversely related to learning and growing from mistakes because the withdrawal and passivity associated with submission might interfere with the active processes required to remount a goal pursuit.

Self-compassion is interesting because it also involves a degree of “taking in” negative feedback but it is also coupled with the provision of care and soothing (Neff, 2016; Gilbert, 2015). Although many people responded with self-compassion, very few responded with high levels of self-compassion and low levels of other responses (14%). This suggests that self-
compassionate responses likely emerge within a complex unfolding of self-to-self interactions, highlighting the importance of bolstering self-compassion rather than attempting to reduce other forms of responses.

Although features of mindfulness, non-judgement, and common humanity may characterize self-compassionate responses (Neff, 2003), the current study successfully used the desire to alleviate one’s own suffering as central to the construct of self-compassion (Gilbert, 2015). Further, the provision of care towards oneself may happen within a broadly unfolding self-to-self interaction over a significant period of time. What seems to matter most is the position of the self-subject that wins out after a dialogical interaction (see below for discussion on dialogical self-to-self interactions).

Attack-resistance can be contrasted against self-submission and self-compassion in that it fundamentally involves a disagreement with the critical message whereas the other two forms of self-responding do not. Past research on attack-resistance has highlighted a component of separating oneself from the critical message and talking back against it (Kelly, Zuroff, & Shapira, 2009). In essence, a highly dialogical model has been used to describe attack-resistance with an implicit assumption that attack-resisting individuals are sufficiently aware of their self-critical process to talk back to their self-criticism (Rector, Bagby Segal, Joffe, & Levitt, 2000). Consistent with this, a significant and moderate relation between attack-resistance and psychological awareness was found. Also, consistent with theoretical depictions of attack-resistance (e.g. Kelly, Zuroff, & Shapira, 2009), this form of self-responding tends to refute the content of the critical message by placing external blame (e.g. the exam did not measure what it was supposed to) or by focusing on one’s strengths as a means of minimizing the importance of the event that triggered self-criticism. A remounting of the original goal was also a salient feature of attack-resistance. It should be noted that these attributes were specifically sought out as
features of attack-resistance (as per the construct definition and coding scheme). Nonetheless these features align with the agency dimension of the interpersonal circumplex (Wiggins, 1998) and recent research has demonstrated that having agentic means of self-relating was positively associated with interpersonal dominance (Bjerke, Solbakken, Friis, & Monsen, 2016) suggesting that attack-resistance may indeed reflect a form of agentic self-to-self relating. In line with the notion of an agentic form of self-to-self relating, many participants rated high in attack-resistance spoke about remounting a goal. Again, this finding is due to the a priori criteria used to capture attack-resistance. At the same time this feature is consistent with a Drive System response articulated in Social Mentality Theory (Gilbert, 2010).

Taken together, the findings of the current study provide preliminary support for an internalized interpersonal world where people can respond to their own self-criticism by self-submitting (low agency, low communion), by being self-compassionate (moderate agency, high communion), or by resisting the attack which sometimes occur with tempered hostility directed outward (high agency, low communion) or with a focus on one’s strengths and a remounting of the goal (high agency, moderate communion). This preliminary model is consistent with recent research which examines styles of self-to-self relating and types of interpersonal problems (Bjerke, Solbakken, Friss & Monsen, 2016) and they also suggest that one’s intrapersonal style is not a simple mirror of their interpersonal style.

**Dialogical self-to-self interactions with self-compassion and/or attack-resistance are adaptive**

One component of Aim 5 was to examine the dynamic aspects of self-to-self responding following self-criticism. Due to the nature of the interview, participants described their responses dialogically. Broadly speaking there is a contrast between monological and dialogical theories of mind. Monological theories assume a singular undifferentiated self that processes sensory
information as inputs, uses language to categorize and discern patterns of categories, and produces behaviours as outputs; whereas dialogical theories or pan-psyche theories of mind suggest a complex network of self-to-self interactions and part-selves (Martin, 2011).

Researchers within dialogical schools of thought focus on ‘utterances’. An utterance “is used to highlight, precisely, the dialogical unit of discourse” (Haye & Larrain, 2011, p.37). Implicit within this notion is that any single utterance is interdependent with other utterances (Holquist, 2003). From a dialogical perspective, an utterance only has meaning in response to a network of other utterances or ideas.

A dialogical model mind does not simply stress that there is a series of interdependent ideas but that dialogical models of self-interaction involves a “change of speaking subjects” (Haye & Larrain, 2011, p.40). This is an important distinction because it highlights that the ‘listener self’ transforms into the ‘speaking self’. This has significant implications for reflexive models of self where the individual is both the subject and object of an utterance.

It is interesting to note the relation between the felt sense of self and one’s part-selves that interact with one-another. When one is being self-critical the felt sense of self takes a submissive position in relation to an inner-critic who is chastising the felt sense of self. Subsequent to this initial submissive position, people can take a stance of ‘affirming submission” which is when they agree with, and submit-to, their internal critic, thereby allying their felt sense of self with the internal critic. This positioning of self-as-subject is in agreement with a self-critical utterance. If this positioning wins out among a multiplicity of part selves that could have responded, it may result in establishing core beliefs about oneself as flawed and having shame-based characterological affect. In other words, when a felt sense of self affirms the self-critical message repeatedly, the person may adopt an identity and characterological affect rooted in self-criticism. This may explain the relation between trait and state self-criticism (Zuroff, Sadikaj,
Kelly, & Leybman, 2016). As an alternative to submitting to self-criticism, people can ally with the criticized self-object and respond with a compassionate or attack-resisting part-self that provides either care or protection after the chastising comment. These self-compassionate and attack-resisting behaviors occur subsequent to the initial submission involved in self-criticism. By providing care or protection to the self that is being chastised, people build a behavioral repertoire based on the assumption that they (a felt sense of self) are worthy of care and protection. As a result, these more complex self-to-self interactions may help build core beliefs about the self as being worthwhile. Both self-compassion and attack-resistance also involve some level of affirming the self as worth of love and kindness (self-compassion) or as competent or able to overcome the challenging situation (attack-resistance).

Individuals with high levels of self-compassion and attack-resistance may also have high levels of self-criticism and self-submission. What seems to matter more than the presence of self-criticism is the addition of compassionate and attack-resisting responses. Moreover, if a self-compassionate or attack-resisting response wins out within an internal discourse, this may foster identities and a sense of self that is resilient to psychopathology. In a way one’s self-to-self interactions followings self-criticism may act as building blocks for personality-like traits that are organized in terms of identity and stable patterns of relating to the world.

**Psychological awareness is adaptive**

This study was the first of its kind to examine psychological awareness in the context of self-criticism and responses to self-criticism. Psychological awareness was found to be positively related to self-compassion and attack-resistance suggesting that people who are more aware of their self-to-self interactions also have a more adaptive manner of responding to self-criticism. At the same time, psychological awareness was not related to self-criticism suggesting that people can have high levels of psychological awareness regardless of how intense and frequent
their self-criticism. This is important because it suggests that bolstering psychological awareness may be important in developing effective interventions to ameliorate self-criticism and its deleterious effects.

**Limitations & Future Directions**

The current study was limited by a small sample size. The nature of the interview as semi-structured, coupled with participant’s diversity of ways of speaking about their self-critical process resulted in too complex a data-set to capture prototypical patterns of self-to-self interacting. With a larger sample size it may have been possible to discern prototypical patterns of self-to-self interacting. Additionally, a more qualitative design, focusing on thematic analysis and allowing for direct quotations of participants, would significantly enhance the richness of the ways in which people experience and respond to self-critical statements.

Our study is also limited by using a homogeneous sample. It is unclear how these findings will generalize to other ethnic and age groups. There is good reason to believe that cultural values might impact one’s experiences of self-criticism and their reactions to it. For example, self-criticism is not as common as self-enhancement in Eastern cultures (Chang & Asakawa, 2003) suggesting that individuals from collectivistic societies may have different relationships to their internal critics than those in the West. Similarly, research suggests that the relationships with ourselves and orientation towards self-criticism develops and matures throughout development (Kopala-Sibley & Zuroff, 2014) suggesting that these results may not generalize to all ages. The data was not gathered from a clinical sample. Given that self-criticism can be adaptive, it is possible that the results do not accurately reflect patterns of self-to-self relating among those with high self-criticism and mental health difficulties associated with self-criticism.
The current study was also limited in that the interview was based on biographical memories of self-criticism rather than recent specific events that triggered self-criticism. The semi-structured interview was based on people’s memories of specific times when they are self-critical and their responses to that self-criticism. As such, the interview captured participant’s experience at the level of constructed self-narratives rather than the mental behaviors themselves. Having repeated interviews about self-critical events that occurred within that week may better capture the actual mental behaviors following self-criticism. Alternatively, capturing the self-critical process after inducing failure may also help researchers access the self-critical process as it unfolds.

This study suggests that interpersonal submission is not highly predictive of self-submission. Future research can further examine this and clarify this link. By further developing measures of self-submission, the discipline may be able to better capture participant’s self-to-self responses following self-criticism and use this information to inform an understanding of how self-criticism contributes to mental health difficulties. Future measures of self-submission should ask about self-submission behaviors themselves rather than as difficulties with standing up to an inner-critic. This will help capture ego-syntonic experiences of self-submission which are likely more frequent in those with reflexive orientations towards themselves. If a more qualitative approach were used to describe the features of self-submission, these data could be used to inform a measure of self-submission.

The current study also suggests that self-to-self interactions following self-criticism are complex and dynamic. Future research should examine self-to-self interactions in clinical samples to examine whether similar patterns of findings are observed. This study suggests that psychological awareness is contributes to adaptive forms of self-to-self responding. Future research should continue to examine psychological awareness as a variable of interest.
Intervention studies may also benefit from examining whether changes in psychological awareness moderate the benefits of bolstering self-criticism and problem-resisting.
Chapter 4: General Discussion

Overview and Summary of Results

In line with the growing recognition that evolutionary models of self-criticism are informative and that self-submission, self-compassion, and the more recent addition of attack-resistance are important constructs in understanding and ameliorating the pathogenic role of self-criticism (Kelly, Zuroff, & Shapira, 2009; Gilbert, 2010), the current project endeavored to help clarify the domain of self-criticism and self-to-self responses following self-criticism.

Gilbert's evolutionary model (2012) proposed the association of psychopathology with the functioning of three global "mentalities" or systems which link affect regulation with interpersonal functioning. Unworthiness, shame, and humiliation result from a state or perception of one's inferior social rank or status (i.e., upward social comparisons; Allan & Gilbert, 1997) and a threat-detection based view of the self and others (Cheng, Foulsham, Kingstone, & Henrich, 2013; Fournier, Moskowitz, & Zuroff, 2002; Gilbert & Proctor, 2006). When engaged in conflict with a higher status individual where there is low chance of success, an Involuntary Submissive Strategy demotivates the individual from conflict thereby avoiding unnecessary loss (Allan & Gilbert, 1997). Social Mentality Theory asserts that people’s behaviors are determined by whichever social mentality is enacted at the time of the event (Gilbert, 2010) and so if people are in a threat-detection mentality they will be more likely to submit rather than retaliate regardless of whether the stimulus is interpersonal or intrapersonal. In this way capacities which evolved interpersonally are experienced in self-to-self relations (Gilbert, 2015). On the other hand, interpersonal capacities for caregiving and seeking dominance can also impact the ways in which we relate to ourselves at times of stress (Gilbert, 2015). The current thesis examined the intersection between the interpersonal and the intrapersonal in the context of self-criticism.
**Study 1.** Study 1 examined whether compassionate and attack-resisting mentalities primed with interpersonal stimuli reduced self-criticism. Results did not support a protective effect for participants primed into compassionate and attack-resisting states of mind. It is recognized that methodological considerations likely interfered with the protective nature of these primes (see above for discussion on methodological considerations); nonetheless, it also highlighted the disconnect between the interpersonal and the intrapersonal (Benjamin, 2003).

Indeed, when one is providing care or defense for another, their felt sense of self is directed outwards and may reflect a somewhat *selfless* position. For example, Catarino, Gilbert, McEwan, & Baiao (2014) found that when people are engaged in genuine compassion (as compared to submissive compassion) there was a lower focus on sense of self. This is also consistent with a mentalizing model of compassion (Liotti & Gilbert, 2011) where compassion ultimately involves experientially based theory of mind involving both sympathy and empathy (Gilbert, 2015). From this perspective, a prime which encourages an other-oriented focus would not be expected to increase caregiving to oneself.

Although this interpretation contrasts with a Social Mentality Theory of motive-oriented states of mind, it is consistent with interpersonal research (e.g. Kopala-Sibley & Zuroff, 2014) and may reflect a more nuanced view of the ways in which social mentalities which evolved interpersonally are carried out in self-to-self interactions.

**Study 2.** Study 2 extended the findings from Study 1 by examining the self-to-self interactions following self-critical statements. The overarching aim of Study 2 was to examine self-to-self interactions from the perspective of interpersonal behaviors. As such participant responses to a semi-structured interview were coded along the lines of self-submission, self-compassion, and attack-resistance; thereby reflecting the interpersonal tendencies to submit, provide care and compassion, and counter-dominate/ resist overtures of dominance from others.
These self-to-self behaviors were contrasted against self-report measures of interpersonal submission.

Results indicate that self-submission is distinct from interpersonal submission and that those who submit interpersonally do not necessarily submit to their own inner critic. At the same time, an adapted measure of self-submission (difficulty standing up to an inner-critic) correlated to self-criticism and negative affect to similar magnitudes as ipsatized measures of interpersonal submission. This suggests that self-submission is distinct from interpersonal submission and that it is a potentially important construct in understanding the pathogenic effects of self-criticism.

A priori categories of self-submission, self-compassion, and attack-resistance were used to code responses to the semi-structured interview. Results suggest that these categories validly capture the majority of participants’ descriptions of their self-to-self interactions following self-criticism. Furthermore, responses to self-critical statements tended to be dialogical in nature, involving a shifting between part-selves. While the phrasing of the interview likely elicited dialogical descriptions of a self-critical process, it is striking that all participants spoke about their self-criticism and responses to self-criticism as interacting part-selves. While the data are not sufficiently robust to speak to the nature of the self as monological or dialogical, they do suggest that using a dialogical model of self may be helpful for people to articulate their varied experiences following self-criticism.

The results of Study 2 also shed light on forms of self-to-self responding that are potentially adaptive. Specifically, regardless of one’s level of interpersonal submission and self-reported self-submission, self-compassion and attack-resistance moderated the relation between self-criticism and negative affect. Having higher levels of awareness of one’s self-critical process was also related to lower levels of negative affect.
Implications for Therapy

The results of the current study suggest that priming motive-oriented states of mind does not necessarily impact modes of self-relating. Furthermore, the results suggest that the self-critical process is dialogical and that it can be complex. The results from Study 2 suggest that having higher levels of complexity is adaptive, as is higher levels of awareness of the self-critical process. In particular, having higher levels of self-compassion and attack-resistance appears to be helpful regardless of one’s level of self-submission. Taken together, these findings may have implications for therapy directed at ameliorating the effects of self-criticism.

Firstly, the results add support to Gilbert’s evolutionary model of self-criticism (2010) which suggests that it is not self-criticism *per se* that is problematic but that responses to self-criticism can be pathogenic. This perspective is firmly grounded in a dialogical model of self, or at least, a pan-psychic model of self (e.g. Minsky, 1986) with part-selves being serially activated (Gilbert, 2010). A monological model of self might suggest that pathological self-criticism involves a self-construal where the self is flawed or inadequate. Treatment programs based on this type of model might seek to find the triggers for self-criticism, note the self-critical statements and either reframe or refute these statements through evidence (Gilbert, 2000). The findings suggest that this approach could be improved upon by helping clients foster different modes of self that would interact with the self-criticism. Related to this is the importance of increasing psychological awareness of people’s self-critical process. Rather than ending with an awareness of triggers and content of self-critical messages, interventions should help foster a deeper awareness of the interchanges among different part selves as they interact in a complex and sequential manner. Building awareness of this process provides the opportunity to introduce other part selves that can emerge and interact in complex ways. The results suggest that there may be benefits for individuals who experience a caring or protective part-self as predominant
and powerful. In support of this claim, past research has found that people experience their inner-critics as powerful and their self-subject as unable to stand up to it are more likely to experience depression (Gilbert & Procter, 2006).

Our findings may also cast doubt on the effectiveness of using general motive-oriented mental states to transfer into adaptive patterns of self-relating. Nonetheless, these results suggest that people enter different positions within dialogical self-interactions such that they can respond to a criticizing self-object in adaptive ways. Currently, treatment programs place significant emphasis on imagery (Gilbert & Proctor, 2006) and cognitive restructuring (Bergner 2013). The results of the current dissertation suggest that experiential interventions may add to these interventions in an important manner. For example, a greater emphasis might be placed on method acting approaches where individuals practice different modes of self-relating (e.g. acting out the attack-resisting part-self). These can then be practiced towards inner-critics and inner-targets in two-chair dialogues or other dramatic representations that position the client in the role of a helpful self-subject (Whelton & Greenberg, 2005). It is worth noting that these activities are already recommended (Gilbert, 2010), however, these findings suggest that they deserve greater importance and centrality in treatment programs. Additionally, whereas psychoeducation in some treatment programs highlight the motive-oriented state of mind, greater generalizability of skills might occur if the psychoeducation focused on entering the different roles within self-to-self interactions. This might be accomplished by providing concrete explanations about part selves and explicit practice in entering self-compassionate or attack-resisting modes as helpful activities following self-criticism (e.g. Kelly, Zuroff, & Shapira, 2009).

Conclusion

The overall goal of this thesis was to examine self-criticism from a meta-theoretical interpersonal and evolutionary perspective. Study 1 examined an assertion specific to Gilbert’s
evolutionary model of self-criticism, namely, that when an individual is in a motive-oriented state of mind (say, compassion/care giving) they will interact with themselves and others in a manner consistent with that system. The results did not support this hypothesis. As a follow-up investigation, Study 2 examined self-to-self interactions following self-criticism.

The results of study 2 found that these interactions primarily involve self-submission, self-compassion, and attack-resistance. Further, and in support of interpersonal and evolutionary models of self-criticism, these results suggest that self-submission may be pathogenic and that self-compassion is ameliorative. Results of Study 2 also indicate that intrapersonal dynamics are complex. People seem to engage in self-referential dialogical exchanges after self-criticism. More complex exchanges involving a variety of part-selves appears to be adaptive and the complexity and types of roles people adopt in the context of self-to-self relating is not clearly linked to their interpersonal style.

This project suggests that responses to self-criticism matters deeply. The thesis opened with a quote from Seneca which reads “There is no person so severely punished as those who subject themselves to the whip of their own remorse”. Results of the current thesis suggest that suffering comes not only from “the whip of our own remorse”, but from our indifference to our injured Self and our submission to the part of us that holds the whip.
Chapter 5: References


https://doi.org/10/1016/j.paid.2008.04.021


Table 1. Table of operational definitions guiding the development of the coding scheme for the ASCPI.

<table>
<thead>
<tr>
<th>Definition from Literature</th>
<th>Relation to the Conceptual Model</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-Criticism</strong></td>
<td>A Threat System response whereby attention is drawn towards personal inadequacies, emotions involve shame or anger, and action tendencies involve either chastising oneself or ruminating about failure. Self-criticism is adaptive in that it can help an individual detect errors in their manner of pursuing a goal. It is maladaptive when it is self-denigrating, global, a pattern of self-responding, and when it is not followed by any supportive self-evaluation.</td>
<td>*I get frustrated with myself and think I will fail *I am lazy *I am disgusting *I am not following my diet *you could have done better *I keep going over my mistake in my head</td>
</tr>
<tr>
<td><strong>Self-Compassion</strong></td>
<td>A Safeness System response whereby attention is drawn towards ones’ own suffering, emotions involve compassion and sympathy, and action tendencies involve a desire to alleviate suffering. Self-compassion is thought to diffuse the sense of threat, thereby allowing the individual to acknowledge their short comings without feeling overcome by threat-based emotions.</td>
<td>*Things will get better with time *Everyone makes mistakes *You are only human *it won’t feel so badly in the future</td>
</tr>
<tr>
<td><strong>Resisting Self</strong></td>
<td>A Drive System response that protects a person’s sense of status, esteem, self, or identity by avoiding or deflecting the blame or harm, by taking a protective stance against the inner-critic, or by indicating a sense of surgency in an effort to continue goal pursuit.</td>
<td>*Disagree with criticism and state positive attributes (I am not ___ I am ___) competent, strong, worthy, capable, smart, etc *It wasn’t a big deal, I shouldn’t get upset *You shouldn’t talk this way to yourself *You are obviously not (critical statement) dumb, incompetent, etc</td>
</tr>
<tr>
<td><strong>Submission</strong></td>
<td>A Threat System response whereby attention is drawn to one’s inadequacies and failures, emotions involve shame, powerlessness, and futility, and action tendencies involve agreeing with one’s inner-critic, rumination, resigned persistence, or avoidance and withdrawal. Submission is never adaptive in itself but it is an important step in taking responsibility, self-soothing, and learning from mistakes. Submission is maladaptive when it is not coupled with self-soothing and learning and when it becomes a primary way of self-responding.</td>
<td>*I agree with myself / my critic *I give into it *I accept the criticism as true</td>
</tr>
</tbody>
</table>
Table 2. Intercorrelations Among Scales and Descriptive Statistics.

<table>
<thead>
<tr>
<th></th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
<th>7.</th>
<th>8.</th>
<th>9.</th>
<th>10.</th>
<th>11.</th>
<th>12.</th>
<th>M(SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-Report Measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Self-Criticism</td>
<td>.50 (.20)</td>
<td>.48 (.13)</td>
<td>.54</td>
<td>-.71</td>
<td>.29</td>
<td>.48</td>
<td>.35</td>
<td>-.37</td>
<td>-.27</td>
<td>.68</td>
<td>.55</td>
<td>23.1 (10.9)</td>
</tr>
<tr>
<td>2. Interpersonal Submission¹</td>
<td>.28</td>
<td>-.31</td>
<td>-.48</td>
<td>.15</td>
<td>.24</td>
<td>.08</td>
<td>-.16</td>
<td>-.09</td>
<td>.55 (.27)</td>
<td>.47 (.11)</td>
<td>27.6 (8.6)</td>
<td></td>
</tr>
<tr>
<td>3. Interpersonal Submission²</td>
<td>-.47</td>
<td>-.52</td>
<td>.11</td>
<td>.24</td>
<td>.17</td>
<td>-.48</td>
<td>-.03</td>
<td>.51 (.15)</td>
<td>.45 (.04)</td>
<td>1.7 (.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Self-Compassion¹</td>
<td>.69</td>
<td>-.17</td>
<td>-.35</td>
<td>-.34</td>
<td>.53</td>
<td>.17</td>
<td>-.57</td>
<td>-.27</td>
<td>19.8 (4.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Self-Compassion²</td>
<td>-.23</td>
<td>-.36</td>
<td>-.29</td>
<td>.29</td>
<td>.20</td>
<td>-.58</td>
<td>-.42</td>
<td>21.8 (4.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Self-Submission</td>
<td>.30</td>
<td>.15</td>
<td>.09</td>
<td>-.10</td>
<td>.12</td>
<td>.26</td>
<td>21.8 (8.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Interview Measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Self-Criticism</td>
<td>.36</td>
<td>-.19</td>
<td>-.16</td>
<td>.40</td>
<td>.33</td>
<td>5.5 (2.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Self-Submission</td>
<td>-.47</td>
<td>-.68</td>
<td>.42</td>
<td>.21</td>
<td>3.0 (3.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Self-Compassion</td>
<td>.32</td>
<td>-.45</td>
<td>-.15</td>
<td>3.2 (3.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Attack-Resistance</td>
<td>-.18</td>
<td>-.07</td>
<td>5.1 (3.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Outcome Measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Negative Affect - Trait</td>
<td>.55</td>
<td>24.4 (7.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Negative Affect - State</td>
<td>15.7 (5.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes.** Interpersonal Submission¹ (values in parentheses partial out interpersonal distress; Allan & Gilbert, 1997); Interpersonal Submission² (values in parentheses are ipsatized values; Hennig & Foti, 2017); Self-Compassion¹ (Raes, Pommier, & Neff, 2011); Self-Compassion² (Gilbert, 2004); correlations reported rs ≥ .27 significant at p < .05, those ≥ .33 significant at p < .01, and those ≥ .42, p < .001.
<table>
<thead>
<tr>
<th>Trigger</th>
<th>Percent (Frequency)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not meeting academic expectations / bad marks</td>
<td>70.18 (40)</td>
</tr>
<tr>
<td>Procrastinating, handing things in late, scheduling issues</td>
<td>35.09 (20)</td>
</tr>
<tr>
<td>Causing relationship conflict or dissolution</td>
<td>28.07 (16)</td>
</tr>
<tr>
<td>Being socially awkward</td>
<td>24.56 (14)</td>
</tr>
<tr>
<td>Idiosyncratic</td>
<td>24.56 (14)</td>
</tr>
<tr>
<td>Work (employment)</td>
<td>19.29 (11)</td>
</tr>
<tr>
<td>Applying for something / working towards a future program</td>
<td>14.04 (8)</td>
</tr>
<tr>
<td>Extracurricular Activities</td>
<td>12.28 (7)</td>
</tr>
<tr>
<td>Exercise and nutrition</td>
<td>10.52 (6)</td>
</tr>
<tr>
<td>Body image</td>
<td>7.02 (4)</td>
</tr>
</tbody>
</table>
Table 4

Qualitative Descriptives for the Interview-Based Coding

<table>
<thead>
<tr>
<th>Category by Dimensional Cut-off</th>
<th>Low (≤ 4)</th>
<th>Moderate</th>
<th>High (≥ 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Criticism</td>
<td>22 (38.6%)</td>
<td>6 (10.5%)</td>
<td>29 (50.9%)</td>
</tr>
<tr>
<td>Submission</td>
<td>22 (39.6%)</td>
<td>5 (8.8%)</td>
<td>30 (52.6%)</td>
</tr>
<tr>
<td>Compassion</td>
<td>35 (61.4%)</td>
<td>3 (5.3%)</td>
<td>19 (33.3%)</td>
</tr>
<tr>
<td>Resistance</td>
<td>38 (66.7%)</td>
<td>4 (7.0%)</td>
<td>15 (26.3%)</td>
</tr>
</tbody>
</table>

High on 0 Adaptive Dimensions: 29 (50.9%)
High on Compassion OR Resistance: 22 (38.6%)
High on Compassion AND Resistance: 6 (10.5%)

Note. Dimensions were coded on a 9-point Likert scale (0 – 9); Total N = 57.
Table 5

Hierarchical Regression Analyses Predicting Self-Reported Trait Negative Affect

<table>
<thead>
<tr>
<th>Variables</th>
<th>ΔR²</th>
<th>F change</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Reported Predictors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1. Self-Criticism</td>
<td>.46</td>
<td>47.35***</td>
<td>.52</td>
</tr>
<tr>
<td>Step 2. Self-Submission/Problems Resisting Self-Compassion</td>
<td>.06</td>
<td>3.35***</td>
<td>.42***</td>
</tr>
<tr>
<td>Interview-Coded Predictors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1. Self-Criticism</td>
<td>.16</td>
<td>10.60**</td>
<td>.27**</td>
</tr>
<tr>
<td>Step 2. Attack Resistance Self-Compassion Self-Submission</td>
<td>.17</td>
<td>4.51**</td>
<td>.16</td>
</tr>
</tbody>
</table>

Notes. Standardized βs presented in the third column are those of the final model.
*p < .05  **p < .01  ***p < .001
APPENDIX A: Compassionate Imagery Script

Hello. Over the next few minutes we will be doing some guided imagery together. You will be asked to pay attention to your breathing, to make different facial expressions and notice how they feel, and then we will begin imagining some things together.

To begin get comfortable in your chair. You will be sitting like this for a few minutes so get comfy. Now close your eyes and begin breathing in and through your nose, focusing on the feeling of your breath. As you breathe in and out, notice the journey the air travels. Does it stop at your lungs or does it go deeper into your belly? If it stops at your lungs try gently allowing your breath to get longer and bring the air right down to your stomach. Allow your breath to slow down. Slow, slow, slowing down. As your breath slows down notice the feeling of gravity pulling you towards the earth. You may notice your feet feel glued to the floor, or you may feel the force pulling you down and that the chair is underneath supporting you.

As you feel this sense of slowing I want you to completely relax your face. Relax your forehead, relax your eyes, relax your cheeks, relax your tongue, relax your jaw. Relax. Now focus on how it feels to have a completely neutral face with all the muscles relaxed. Not I want you to wear a half smile. Not a smile of great joy, but a soft smile like you are meeting an old friend. Some people call this the Buddha smile. Notice how it feels to smile. Now drop the smile and return to the completely neutral face. Now wear the smile again and notice how that feels.

Wearing this half smile and breathing slowly we are going to begin to imagine something together. Remember, it does not matter if it is true, we are just going to pretend. I want you to imagine that you are the most compassionate person in the world. Now whatever that looks like I am going to bring a few qualities to mind. Imagine that as the perfectly compassionate being you have a deep wisdom. You understand the causes of suffering and you understand how to stop suffering. You understand that we all just find ourselves here with a body we never chose, and a brain we never chose, dealing with challenges that are difficult. You have a deep wisdom. Now imagine you have a perfect sensitivity to the suffering of others. You know when they are in pain. You can sense and feel their suffering. Lastly I want you to imagine that you have strength. A great strength. With this strength you can help bring people out of their suffering. You can actually help them. So remember you are the most compassionate person in the world. You have a deep wisdom and you understand the causes of suffering. You have perfect awareness of the suffering of others, you can tolerate their pain and sit in it with them. You have a great strength and you can help them alleviate their suffering.

Now I want you to imagine someone you care about. Bring their face to mind. I want you to imagine that you can help this person with their difficulties. Remember in this image exercise you are the most compassionate person in the world. Imagine that you have a deep wisdom and you understand this person’s suffering, you are aware of their suffering and can sit with them in their painful feelings, and you are strong. You can actually help this person reduce some of their pain.

As you imagine these things I want you to say something in your head. Say the person’s name and then say “may you find peace, may you be happy”. Slow your breathing and with each
outbreath, in the warmest and most loving voice you can imagine say “may you find peace, may you be happy”. Keep repeating this with each breath until the bell rings. ‘Ring’
APPENDIX B: Attack-Resistance Imagery Script

To begin get comfortable in your chair. You will be sitting like this for a few minutes so get comfy. Now close your eyes and begin breathing in and your through your nose, focusing on the feeling of your breath. As you breathe in and out, notice the journey the air travels. Does it stop at your lungs or does it go deeper into your belly? If it stops at your lungs try gently allowing your breath to get longer and bring the air right down to your stomach. Allow your breath to slow down. Slow, slow, slowing down.

Together we are going to try and create the sense of inner peace and strength. A type of relaxation that is strong. Like a peaceful warrior or martial artist. So first, now that you are slowing down, I want you to go completely floppy. Relax your shoulders and arms, slouch in your chair, and let you head dangle. Notice how that feels. Now I want you to sit up straight in your chair and hold your head high. Breath slowly and feel how the breath gives you strength. Now go floppy again, let you head go limp, your belly relax, your shoulders relax. Notice how that feels. Now again enter a strong posture. Straighten your back and hold your head high. Your breath is slow and you are relaxed but you are also strong and confident.

Now I want you to imagine that you are a strong and loving person. It does not matter if you think it is true, remember we are just imagining together. Imagine that you are this strong, confident, loving person. Now I want you to focus on three qualities in particular. Perseverance. You have a great perseverance. No matter what challenges you face you are able to get up and overcome them. You know that in the battle between the rock and the river, the river always wins. Not through strength but through perseverance. You will persist until you succeed. Now imagine that you have a great intellect and an amazing grasp of logic. Using logic and reason you can see a solution to every problem. The hardness of rational thought allows you reach your goal. Lastly, imagine you have a great strength. With this strength you can help bring people with their problems.

So remember you are a strong and loving person. You have a great ability to persevere until you succeed. You can use logic and reason to solve problems. And you have an amazing strength and ability. With these you can help people solve their problems.

Now I want you to imagine someone you care about. Bring their face to mind. I want you to imagine that this person is in trouble. Someone at their work, their school, or their personal life is bullying them. Now imagine that you can protect this person and solve their social problems. You will persevere until their world is better, you will use logic and reason to find a solution, and you have strength and ability to carry out your plan with success. Your breathing is slow and deep. Your posture is strong.

As you imagine these things I want you to say something in your head. Say the person’s name and then say “together I will help you solve your problem. I will help you. We will succeed”. Slow your breathing and with each outbreath, in the strongest and most loving voice you can imagine say “together I will help you solve your problem. I will help you. We will succeed”. Keep repeating this with each breath until the bell rings. ‘Ring’
APPENDIX C: Raven's Matrices

Item 1

Item 2

Item 3

Item 4
<table>
<thead>
<tr>
<th>APPENDIX D</th>
</tr>
</thead>
</table>

Date of transcription:

Transcriber name:

Participant ID:

<table>
<thead>
<tr>
<th>Celebrate (1-5)</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Are there times when you are self-critical (Y/N/spontaneous answer)</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>When is this most likely to happen</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>What words go through your head when you are self-critical</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Tone of self-criticism (1-5)</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Response after that (D5a-5c &amp; E)</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>i.e. what do they say in response to their self-criticism</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Specific questions on responding</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Do you ever defend, comfort, submit etc (F)</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Tone positive responses (1-5)</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>If other people said that to you (H-I)</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Tone father</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Tone mother</th>
<th></th>
</tr>
</thead>
</table>

| Awareness of self-critical process |   |
Rate:

How comfortable are they in responding in the interview? (1-9 and qualitative impressions)

What extent are they censoring themselves? (1-9 and qualitative impressions)

What extent are they aware of the back and forth in their mind? (1-9 and qualitative impressions)

Rate the severity of the inner critic (1-9)
APPENDIX E

Awareness of Self-Critical Process Interview (ASCPI) Coding Guide

Step 1: Read the transcript
Read the whole transcript. Look to see if the question of responding to self-criticism is revisited. If the response to self-criticism is revisited later, include this portion in the coddle area.

Step 2: Highlight code-able units
Look for specific responses that fit into the different categories (see chart).

To code the responses, highlight sentences or parts of sentences that fit into each of the categories found in the table. ***If there is another response capturing self-to-self interactions, make sure to highlight and make a note of the participant ID

Step 3: Rate the participant’s response style
These ratings will be used to carry out most of the statistics. In this part of the coding you are required to think about the person’s response holistically, as a gestalt. Do NOT just count the number of different highlights. Instead, think about the overall impression you get from the participant’s response in addition to the number of discrete codable statements. You may need to re-read this portion of the transcript before answering.

To what degree did the participant respond to their self-criticism with:

1) Self-criticism
2) Self-Submission
3) Self-Compassion
4) Attack-Resistance

Use the attached table template (at end of document) for each rating.
Participant ID:
Coder:
Date of coding(year/mm/dd):

<table>
<thead>
<tr>
<th>Rate*</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>none</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>very much</td>
</tr>
</tbody>
</table>

Rate the level of the participant's **Self-Criticism**

Consider:

- Frequency of self-critical statements
- Whether it reflects a stable characteristic (I am lacy VS I slacked off)
- Whether it puts down the person or is derogatory

Overall, to what degree did the participant respond to their self-critical statement...

**With self-submission**

Consider:

- Whether the person agrees with the self-criticism with a sense of passivity or withdrawal (e.g. I totally agree with it [code sub] VS I know I messed up but everyone makes mistakes so I try not to be hard on myself [not submission])
- Seems beaten down or put down by the self-criticism
<table>
<thead>
<tr>
<th>With self-compassion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consider:</td>
</tr>
<tr>
<td>- Kindness, common humanity, mindfulness</td>
</tr>
<tr>
<td>- Attempt to soothe or take care of the criticized person with warmth</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>With Attack-Resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consider:</td>
</tr>
<tr>
<td>- Does the person rebut the self-criticism in some way?</td>
</tr>
<tr>
<td>- Is there a cold logical approach used?</td>
</tr>
<tr>
<td>- Does the person bolster themselves as a way of ignoring the self-criticism (e.g. I am not dumb, I’m the best)</td>
</tr>
<tr>
<td>- Does the person’s response seem to focus on ignoring the criticism in favor of goal pursuit</td>
</tr>
</tbody>
</table>

Indicate which overall style they respond with the most. i.e. what is their predominant response style?

- self-compassionate
- self-defending
- submissive
- mixed (indicate which mix)

* 0 = the attribute was not displayed at all, totally absent
* 3 = somewhat but not very much
* 5 = a moderate amount
* 7 = a fair bit
* 9 = very much (almost prototypical response and frequent)
APPENDIX F

After you are self-critical what goes through your head next?

A1. Do they provide an answer that BOTH captures what happens after the self-criticism is over AND that is a mental event (i.e., not physical) that is NOT an action tendency?
   If Yes go to A2.
   If No go to A3.

A2. Does their response contain details and spontaneous elaboration of the mental process?
   If Yes score 5
   If NO score 4

A3. Does their response contain EITHER a statement indicating that their self-criticism continues after the initial statement OR a response about a behaviour or action tendency?
   If YES go to A4.
   If NO score 0 and discontinue the Awareness Scoring

A4. After being prompted/redirected do they provide an answer that BOTH captures what happens after the self-criticism is over AND that is a mental event where that event is NOT an action tendency?
   If YES score 3
   If NO go to A5

A5. After being prompted/redirected a second time do they provide an answer that BOTH captures what happens after the self-criticism is over AND that is a mental event?
   If Yes Score 2.
   NOTE: an action tendency related to preventing the same mistake is scored as 2. Otherwise go to A6.
   If NO go to A6.

A6. Do they ever provide a response that BOTH captures what happens after the self-criticism is over AND that is a mental event?
   If Yes score 1
   If no Score 0.