

Measuring Regulatory Focus

by

Steve VanKrevelen

B.S., Noth Central University, 1992

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Psychological Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2017

Approved by:

Major Professor
Clive Fullagar, Ph.D.

Copyright

© Steve VanKrevelen 2017.

Abstract

Regulatory focus has emerged as an important construct in the organizational sciences. In the past ten years more than 200 papers have been published applying regulatory focus to a wide variety of contexts ranging from marketing and persuasion to feedback and performance appraisal (Johnson et al., 2015). Despite the ubiquity of RFT's application, only a few studies have targeted the psychometric properties of measures of regulatory focus; and the findings thus far suggest that improvement is needed. Haws (2010) evaluated five measures of regulatory focus and concluded that they differed substantially with respect to their theoretical content, and that most demonstrated unacceptably low internal consistency. Summerville & Roese (2008) drew similar conclusions in their evaluation of the Regulatory Focus Questionnaire (RFQ) and the General Regulatory Focus Measure (GRFM) and added that the two scales might actually be measuring different underlying constructs. Given the inconsistencies and problems associated with existing measures of regulatory focus, the purpose of the current research is to extend the critical evaluation of existing measures of regulatory focus and then to propose the development of a new measure based on rigorous scale development practices like those set forth in Hinkin, (1995) and Crocker & Algina, (1986). A new scale of Regulatory Focus was developed designed to measure all aspects of RFT and to test whether a two-factor or a four-factor SEM model fit the data best. The final scale consisted of 14 items. CFAs were used to test whether a two-factor or a four-factor model of regulatory focus fit the data best. Results suggested that both models fit the data equally well. However, for parsimony reasons and given that one of the latent factors of the four-factor model contained only two items (making any estimates of internal consistency difficult) the two factor model of regulatory focus was retained as the preferred model.

Table of Contents

List of Figures	v
List of Tables	vi
Chapter 1 - Overview	1
Understanding Regulatory Focus Theory	3
Chapter 2 - Method	31
Item Generation	31
Participants.....	34
Procedure	35
Chapter 3 - Results.....	36
Chapter 4 - Discussion	47
References.....	55
Appendix A - Promotion Prevention Scale (PPS) & RFQ	59
Appendix B - Initial PPS Items.....	61

List of Figures

Figure 1. The theoretical underpinnings of regulatory focus theory.	8
Figure 2. Standardized parameter estimates for a first-order confirmatory factor analysis (AMOS) for the Promotion Prevention Scale (PPS).....	9
Figure 3. Standardized parameter estimates for a first-order confirmatory factor analysis (AMOS) for the Promotion Prevention Scale (PPS).....	42

List of Tables

Table 1. Subject matter expert ratings of initial regulatory focus items.....	32
Table 2. Descriptive statistics for the final items of the Promotion Prevention Scale (PPS)	40
Table 3. Chi-square and model fit indices for two and four factor models of regulatory focus...41	
Table 4. Regression predicting promotion regulatory focus from positive and negative affect...44	
Table 5. Regression predicting prevention regulatory focus from positive and negative affect...45	

Chapter 1 - Overview

“When I am old and dying, I plan to look back on my life and say “wow, that was an adventure” not “wow, I sure felt safe”

–Tom Preston-Werner, Github Cofounder (2008)

“I don’t look to jump over 7-foot bars; I look around for 1-foot bars that I can step over”

–Warren Buffet, Chairman and CEO of Berkshire Hathaway (BrainyQuote, n.d.)

Individuals often have similar goals but use very different means to obtain them. The quotes above (Johnson, Smith, Wallace, Hill, & Baron, 2015) are both from highly successful businesspersons but represent different perspectives on the kinds of strategies they use to achieve their goals. The first perspective represents a businessperson who emphasizes risk over safety when pursuing goals; the second represents a more cautious businessperson who values safety over risk-taking. Presumably, both share the ultimate goal of financial success, but the strategies and tactics they use to obtain their goals differ. Understanding the foundation of these differences is of central import to this proposal because measuring a construct is dependent upon an accurate understanding of its theoretical content.

Regulatory focus theory (RFT) (Higgins, 1997, 1998) is a motivational theory that helps explain why the first businessperson in the quote above values risk over safety, and businessperson two values safety over risk. RFT posits that individuals self-regulate their behavior during goal pursuit using two distinct self-regulatory systems. The promotion-focused system is sensitive to gains and opportunities in the environment and helps regulate behavior to fulfill nurturance needs. The prevention-focused system, on the other hand, is sensitive to loss and risk in the environment and helps regulate behavior to fulfill safety and security needs. Despite being orthogonal sub-dimensions of regulatory focus, individuals tend to have chronic

tendencies to pursue goals using either a promotion or a prevention focus. Thus, businessperson one above likely has a stronger promotion orientation and so eagerly pursues accomplishments with little attention paid to risk while guarding against errors of omission; the second businessperson likely has a stronger prevention orientation and so is risk averse and pursues goals cautiously by guarding against errors of commission.

Regulatory focus has emerged as an important construct in the organizational sciences. In the past ten years more than 200 papers have been published applying regulatory focus to a wide variety of contexts ranging from marketing and persuasion to feedback and performance appraisal (Johnson et al., 2015). Researchers have also utilized regulatory focus as a moderator variable to help explain inconsistent findings in the relations between important variables. For example, a central finding from Kluger and DeNisi, (1996) is that work performance feedback given to workers actually impedes work motivation roughly thirty-percent of the time. Subsequent studies however, (e.g., Van-Dijk & Kluger, 2004) showed that regulatory focus moderates the relationship between feedback and motivation such that prevention focused workers are more motivated by negative feedback (what they did wrong) whereas promotion focused workers are more motivated by positive feedback (what they did correctly). Additionally, Shah & Higgins, (1997) presented evidence that challenged the classic interaction effect between expectancy and value on goal commitment (how committed a person is to a goal is a byproduct of how much they value the goal and how likely it is that they can achieve the goal). In a series of four studies they found that the interaction between expectancy and value on goal commitment was positive for promotion focused individuals, but negative for prevention focused individuals. This is likely because as the value of a goal increases for a prevention focused individual (meaning the person feels more obligated to complete the goal) the impact of

expectancy on motivation decreases. In other words, when a prevention oriented person feels a strong obligation to complete a goal, their perception regarding their ability to actually complete the goal becomes less relevant. Thus, regulatory focus can be considered an important and relevant motivational construct with broad and varied potential applications.

Despite the ubiquity of RFT's application, only a few studies have targeted the psychometric properties of measures of regulatory focus; and the findings thus far suggest that improvement is needed. Haws (2010) evaluated five measures of regulatory focus and concluded that they differed substantially with respect to their theoretical content, and that most demonstrated unacceptably low internal consistency. Summerville & Roese (2008) drew similar conclusions in their evaluation of the Regulatory Focus Questionnaire (RFQ) and the General Regulatory Focus Measure (GRFM) and added that the two scales might actually be measuring different underlying constructs.

Given the inconsistencies and problems associated with existing measures of regulatory focus, the purpose of the current research is to extend the critical evaluation of existing measures of regulatory focus and then to propose the development of a new measure based on rigorous scale development practices like those set forth in Hinkin, (1995) and Crocker & Algina, (1986).

Understanding Regulatory Focus Theory

RFT (Higgins, 1997, 1998) developed based on the integration of two lines of research conducted by Higgins beginning in the early 80's and extending into the late 90's. The first was an elaboration of the hedonic principle – the simple motivational idea that that people are ultimately motivated to pursue pleasure and avoid pain. Higgins, (1997) argued that theories using hedonism as a basis for analysis fall short because they do not explain *how* or *why* people seek pleasure and avoid pain, only that they *do* (Johnson et al., 2015). Based on a series of

studies, Higgins, (1997, 1998) demonstrated that there are qualitative differences in the way people approach goals. Persons with a dominant promotion orientation work to achieve desired end states by maximizing their likelihood for achievement and gains. Persons with a dominant prevention focus, on the other hand, work to achieve desired end states by minimizing risks and the likelihood of making mistakes. So, *how* persons go about pursuing a desired goal varies drastically based on their regulatory focus orientation.

Higgins (1997) posits that these unique regulatory systems emerge within individuals based on differential operation of the hedonic principle. The argument he makes is that the hedonic principle operates differently in people when serving qualitatively different survival needs like nurturance vs security. Given that meeting both nurturance and security needs are crucial for human survival, children learn early-on how to regulate their behavior in order to meet these needs. But how children regulate their behavior differs based on the kind of need they are trying to meet – self-regulation in pursuit of nurturance needs differs from self-regulation of security needs. According to RFT, regulation around nurturance needs involves a promotion focus, whereas regulation around security needs involves a prevention focus. But why? What is it about certain caretaker-child interactions that shape self-regulation differently for nurturance vs. security related needs?

According to Higgins, (1997) caretaker-child interactions that evolve into a promotion regulatory focus involve the prominence of pleasurable experiences as rewards for success and the removal of these pleasurable experiences as punishment for failure. For example, the caretaker hugs and kisses the child when the child behaves in a desired manner, or ends a meal when the child throws food at the table. Implied in both interactions is the message that what matters is attaining accomplishments or fulfilling hopes. The presence or absence of positive

experience becomes the shaping mechanism. Thus, over time, the child learns to self-regulate using a promotion focus—an emphasis on and preoccupation with achievement, growth, and accomplishment.

Higgins (1997) also argues that caretaker-child interactions that evolve into a prevention regulatory focus involve an emphasis on the absence of negative experiences as rewards for success and the presence of negative experiences as punishments for failure. For example, the caretaker childproofs the house and teaches the child to mind their manners or yells at the child for making a mistake or for being irresponsible. Implied in both of these interactions is the message that what matters is avoiding pain and maintaining the status quo by being responsible and meeting obligations. The absence or presence of negative experience becomes the shaping mechanism, and, over time, the child learns to self-regulate using a prevention focus—an emphasis on and preoccupation with obligation, duty, and safety.

Thus both promotion and prevention self-regulatory systems develop based on the interaction style between the child and his or her caregiver. And over time the dominance of either the promotion regulatory system or the prevention regulatory system shapes how individuals approach the pursuit of goals.

The second line of research that heavily influenced RFT is Self-discrepancy theory (Higgins et al., 1987). Self-discrepancy theory posits that individuals have various “selves” that help guide behavior. One’s “actual” self is made up of self-perceptions about who that person really is — the characteristics they actually possess. One’s “ideal” self is made up of perceptions about how that person would like to be ideally, i.e. their self-related dreams and aspirations. One’s “ought” self is comprised of perceptions about how that person should be, i.e. their self-related duties and obligations in the world. According to Self-discrepancy theory, individuals

make comparisons between their actual/ideal selves and their actual/ought selves, and these discrepancies motivate discrepancy-reducing behavior in the individual. Further, Self-discrepancy theory posits that the salience of discrepancies between actual/ideal vs. actual/ought differs by individual. Some emphasize and are more sensitive to reducing actual/ideal discrepancies, while others emphasize and are more sensitive to reducing actual/ought discrepancies.

A key component of Self-discrepancy theory (Higgins et al., 1987) is its postulation that the kind of discrepancies that are salient to the individual matter motivationally because they are associated with different kinds of affective states. Succeeding or failing at reducing actual/ideal discrepancies *feels* different than succeeding or failing at reducing actual/ought discrepancies. When people succeed in reducing gaps between their actual and ideal selves they experience emotions associated with happiness; when they fail they experience emotions associated with sadness or dejection. Conversely, when people succeed in reducing gaps between their actual and ought selves they experience emotions associated with calmness or quiescence; when they fail they experience emotions associated with fear or agitation.

The integration of self-discrepancy theory with an elaboration of the hedonic principle resulted in the formation of regulatory focus theory. Individuals are motivated by the hedonic principle to pursue desired goals, but *how* they pursue is dictated by whether they emphasize and seek to reduce gaps between their actual/ideal selves or their actual/ought selves and by whether they are pursuing nurturance related or security related goals. In RFT terms, promotion focused individuals view goals primarily as meeting nurturance needs and are sensitive to reducing gaps between their actual and ideal selves. Prevention focused individuals view goals primarily as meeting security needs and are sensitive to reducing gaps between their actual and ought selves.

Hopefully this brief discussion of RFT has served to provide insight into some of the potential measurement challenges surrounding RFT. As shown in Figure 1 below, RFT is the byproduct of integrating two lines of research—one emphasizing self-regulation governed by survival needs (nurturance or security) and the other emphasizing discrepancy reduction between one’s actual and either their ideal or ought selves. Thus, the conceptual and theoretical underpinning of promotion regulatory focus is self-regulation around nurturance needs and actual-self / ideal-self discrepancy reduction. And the theoretical underpinning of prevention regulatory focus is self-regulation around security needs and actual-self / ought-self discrepancy reduction. Higgins’ theory suggests that nurturance needs and ideal goals are conceptually (and empirically) the same. And that security needs and ought goals are conceptually (and empirically) the same. If Higgins is correct that nurturance and ideal goals tap into a single latent construct and that security and ought goals also tap into a second latent construct, then (from a factor analysis perspective) a two-factor model should yield optimal fit statistics. However, it is also possible that nurturance and ideal goals, as well as security and ought goals, are empirically distinct. If that is true then a four-factor model should yield better fit statistics than the two-factor solution. As we will see when we discuss the measurement of regulatory focus in more detail, the extant empirical evidence warrants an exploration of alternative models of regulatory fit.

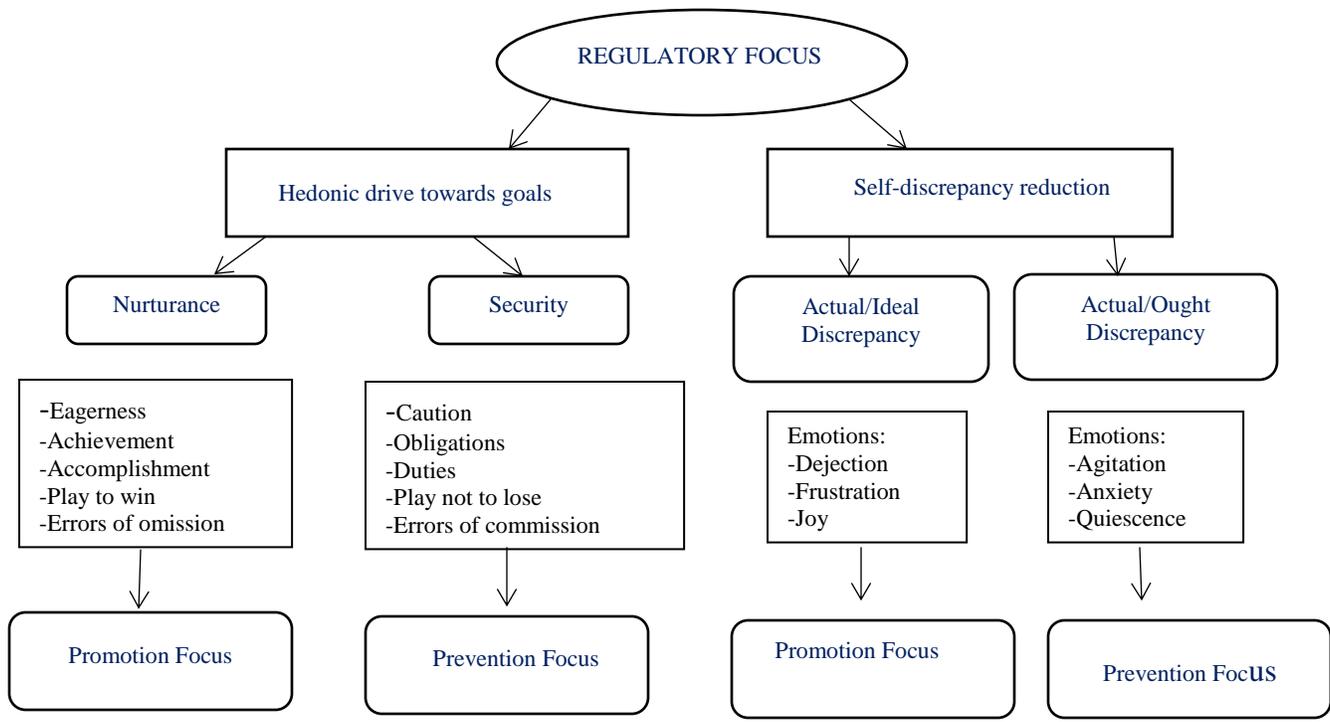


Figure 1. The theoretical underpinnings of regulatory focus theory

In sum, research on regulatory focus theory has evolved over time. Initial perspectives (Higgins et al., 1987) emphasized the reduction of gaps that exist between one's actual-ideal, actual-ought selves. Later perspectives (Scholer & Higgins, 2011) retained some of the language of the earlier work, but reframed regulatory focus as a mechanism used in the hedonic-driven self-regulation of two fundamental survival needs: nurturance and security. These shifts in perspective do not, in and of themselves, undermine the credibility of regulatory focus theory. Construct definitions do and should change over time (Cronbach & Meehl, 1955). However, the fact that differing perspectives and operational definitions exist does contribute to confusion in the literature about how to measure regulatory focus. A new measure of regulatory focus is needed that captures the multidimensional nature of regulatory focus in a way that is both reliable and valid.

Measuring Regulatory Focus

Given the definitional complexity of regulatory focus, it is probably not surprising that approaches to measuring regulatory focus have varied. Most of the early work on regulatory focus used a qualitative measurement instrument called the Selves Questionnaire (SQ; Higgins, Klein, & Strauman, 1985) whereas, as shown by Summerville & Roese (2008), more recent work has primarily used two measures, the Regulatory Focus Questionnaire (Higgins et al., 2001), and the General Regulatory Focus Questionnaire (Lockwood, Jordan, & Kunda, 2002).

The Selves Questionnaire (Higgins et al., 1985) is a qualitative instrument designed to measure discrepancies that may exist between a person's actual self and their ideal or ought selves. The measurement procedure associated with the Selves Questionnaire is complicated. First, respondents complete a free-form questionnaire which asks about various aspects of their self-concept from their own perspective and from the perspective of their mother, father or best friend. Instructions read:

“In the following questionnaire, you will be asked to list [up to ten] attributes of the type of person you (or mother/father/friend) think you *actually*, *ideally*, and *ought* to be:

-*Actual self*: Your beliefs concerning the attributes you (or mother/father/friend) think you actually possess.

-*Ideal self*: Your beliefs concerning the attributes you (or mother/father/friend) would ideally like to possess; your ultimate goals for yourself.

-*Ought self*: Your beliefs concerning the attributes you (or mother/father/friend) believe you should or ought to possess; your normative rules or prescriptions for yourself.”

This procedure results in lists of attributes regarding the actual, ideal, and ought self-concept from four different perspectives: one's own perspective, the mother's perspective, the father's

perspective, and a close friend's perspective. After respondents complete the attribute lists, raters then look for self-concept matches and mismatches from the various lists (rater agreement was .80 for one particular study, but I could not find any other quantitative evaluation of the SQ). Attributes are rated a "match" if the same word (or synonym) is used; attributes are rated a "mismatch" if the opposite word (antonym) is used; attributes are rated "other" if they neither match nor mismatch. Self-discrepancy scores are then created by subtracting the total number of matches from mismatches. This results in an actual/own-ideal/own discrepancy score, an actual/own-ought/own discrepancy score, an actual/own-ideal/other discrepancy score, and an actual/own-ought/other discrepancy score (for scoring simplicity reasons, the researchers used the attributes from the 'other' list deemed most important by the study participant). These scores are then combined to form a total self-concept discrepancy score. For the purposes of this proposal, it is important to note that most of the citations from the seminal regulatory focus papers (Higgins, 1997, 1998) use this operationalization of regulatory focus. Prevention focused individuals are those whose self-discrepancy scores were greater for actual-ought than for actual-ideal. Conversely, promotion focused individuals are those whose self-discrepancy scores were greater for actual-ideal than for actual-ought. As we will see, this operationalization of regulatory focus is quite different from the more recent operationalization found in the RFQ and the GRFQ.

The Regulatory Focus Questionnaire (RFQ) (Higgins et al., 2001) is a traditional Likert-style questionnaire consisting of eleven items designed to measure promotion or prevention regulatory focus. Respondents rate the frequency of certain behaviors on a five-point Likert-type scale (1-Never or seldom to 5-Very often). An example of a prevention item is "Did you get on your parents' nerves often when you were growing up?" An example of a promotion item is "How often have you accomplished things that got you 'psyched' to work even harder" (See

Appendix A for the full list of RFQ items). Thus, the RFQ operationalizes regulatory focus as a two-dimensional construct consisting of something like “obedience to parents” plus an item that seems to measure “caution” (prevention focus), and another that seems to measure “accomplishment” (promotion focus). This is quite different than the operationalization of regulatory focus found in the Selves Questionnaire. The SQ is a *relative* measure, i.e., it taps into the discrepancies between one’s actual self and one’s ideal or ought self. Thus, if one strongly endorses having, e.g., “made progress towards being successful in their life” on the RFQ, that would contribute towards a promotion-oriented regulatory focus. However, using the SQ, one would only get credit for a promotion focus if there was a *gap* between how successfully one perceives themselves and how successful they would like to be ideally. A similar challenge is present in the General Regulatory Focus Measure (GRFM; Lockwood et al., 2002) which is discussed next.

The GRFM (Lockwood et al., 2002) is also a Likert-style scale designed to measure promotion and prevention regulatory focus. But the scale items, especially for the prevention focus subscale, are quite different from those on the RFQ. Respondents answer 18 questions by estimating how “like them” certain behaviors are (1-Not at all true of me to 9-Very true of me). An example of a prevention item is: “I am anxious I will fall short of my responsibilities and obligations.” An example of a promotion item is: “I typically focus on the success I hope to achieve in the future” (see Appendix A for a full list of GRFM items). The GRFM operationalization of regulatory focus is more closely aligned with the operationalization found in the SQ. Both measures use something like “striving for ideals and aspirations” as an operationalization of promotion focus. And both use something like “working to avoid failure and meet obligations” as an operationalization of prevention focus. However, as mentioned in

the discussion of the RFQ, the SQ is still a relative measure that taps into self-discrepancies, so the concerns discussed previously also apply to a comparison of the GRFM and the SQ—directly measuring an individual characteristic using a Likert-style scale is very different than measuring the gap that may exist between one’s actual-ideal or actual-ought self.

Curiously, given the rather obvious differences in the operationalization of regulatory focus in the SQ, the RFQ, and the GRFM, few very studies have empirically examined the psychometric properties or the content validity of these scales. In fact, based on a fairly exhaustive review of the regulatory focus literature, only two studies have examined these questions; and the results are less than promising.

Summerville & Roese (2008) utilized an exploratory factor analysis to compare responses to items from both the RFQ and the GRFM and found that the items from the two scales loaded on different factors that were largely uncorrelated. Further, they found that prevention focus items from the two scales were actually *negatively* correlated to one another. They concluded that “this dissociation of these measures suggests that the two [scales] may represent two unique constructs rather than a single phenomenon” (p. 253). The discussion section of their paper also recommends an exhaustive study pertaining to the construct validity of regulatory focus based on these findings.

Haws (2010) evaluated the SQ, RFQ, and GRFM plus two other, less utilized, measures of regulatory focus. Consistent with the findings from Summerville & Roese (2008), Haws found average to low internal consistency reliability coefficients and poor correlations between promotion subscales as well as low correlations among prevention subscales. The SQ alpha coefficients ranged from .54 to .64 for the promotion and prevention subscales. The RFQ alpha coefficients were better ranging from .64 to .80 for the promotion and prevention subscales. The

Lockwood scale (GRFM) was the best of the three with alphas ranging from .77 to .85 over their two studies.

Correlations between the various subscales were also low and in some cases negative. For example, the correlation between the prevention subscales of the RFQ and GRFM ranged from -.14 to .18. The correlations among the promotion subscales were higher ($r = .25$ to $.27$), but still low with only eight percent shared variance.

The Haws (2010) studies also evaluated the representativeness / content validity of each of the scales i.e., how well the scale items reflect the conceptual and theoretical construct definition. Based on their findings, Haws concludes that “each regulatory scale appears to...emphasize different aspects of regulatory focus theory” (p. 972). This interpretation seems gracious. Given the collection of less-than-impressive reliabilities and subscale correlations, it is more likely that existing scales are comprised of poorly written and content-deficient items.

Given that regulatory focus theory posits relatively stable trait-like behavior of the promotion and prevention sub-dimensions, Haws (2010) evaluated the test-retest reliability of the measures using a five-week gap between administrations. Results for the SQ were poor with coefficients of .47 and .40 for the promotion and prevention subscales respectively. The GRFM test-retest reliability was better (.67 - promotion and .62 - prevention) but still low, especially given the relatively short time between administrations. The RFQ had the best test-retest reliability with .69 and .75 for the promotion and prevention subscales respectively, but when compared to other dispositional measures (e.g., the NEO scale showed a six-year test-retest reliability of .83), the RFQ test-retest performance is modest at best.

In addition to the psychometric problems of all three measures, it seems likely that there are content validity problems with both the RFQ and GRFM. According to RFT described above,

there are qualitatively different concerns that are central to promotion or prevention regulatory focus. Promotion focus is about fulfilling nurturance needs and about achievement and avoiding errors of omission and is engendered within individuals based on caretaker-child interactions that overtly reward desired behavior and withhold those rewards for negative behavior. According to RFT, prevention focus is about fulfilling security needs and about duty and obligation and avoiding errors of commission and is engendered within individuals based on caretaker-child interactions that tacitly reward desired behavior (by not yelling etc.) and overtly punish for negative behavior. It follows then that items written to measure promotion focus should have content related to nurturance needs, achievement, missing out on opportunities, and about how their caretaker rewarded/punished desired or undesired behavior. Similarly, items written to measure prevention focus should contain content related to security needs, and about fulfilling duties, avoiding risk, and about caretaker-child interactions. A review of the promotion and prevention items of the RFQ suggests that they may be grossly deficient from a content perspective. None of the items tap directly into any of the central concepts of promotion or prevention focus. Instead they ask (e.g.) about getting what one wants out of life, being psyched during activities, task performance, getting on parents' nerves, and about obeying rules. Better items can be written that more directly tap into the latent promotion and prevention focus construct. Similar but less egregious problems exist related to the content of the GRFM. Items on the GRFM almost exclusively tap into the promotion and prevention related differences in pursuing goals as ideals versus pursuing goals as obligations, but neglect item content related to nurturance or security needs.

Together, results from the two empirical examinations of frequently used measures of regulatory focus are concerning. Indices of internal consistency and test-retest reliability are only

average at best, and the low correlations between subscales of the scales suggest that improvement is needed in the self-report based scales of regulatory focus. Further, the low correlations between subscales (which should be measuring the same latent variables) of regulatory focus suggests that popular existing measures are tapping into unique latent constructs. Finally, both the RFQ and GRFM lack content validity, and could be improved by rewriting existing items and adding additional items that tap into the full range of RFT content. The proposed research seeks to remedy these problems by creating a new measure of regulatory focus following the rigorous scale development practices described earlier and to test the dimensionality of regulatory focus.

Chapter 2 - Psychological Measurement

The importance of psychological measurement to the social sciences can hardly be overstated. Indeed, sometimes measurement can have life or death consequences. As pointed out by Furr & Bacharach, (2008), in North Carolina persons who are mentally challenged (IQ score below 70) cannot receive the death penalty. Thus precise measurement of (in this case) intelligence becomes so crucial that person's lives depend on it, literally. In this section we briefly define psychological measurement and related challenges and then discuss its central role in the social sciences in general and psychology in particular. I then conclude this section by explaining psychological measurement best practices and how they might be applied to measuring regulatory focus.

So what is meant by *psychological measurement*? Volumes have been written on the topic, including debates about how various ontological perspectives (e.g. empirical operationalism, pragmatism, constructive realism) impact measurement definitions and approaches (e.g. see Maul, Irribarra, & Wilson, 2016). An explication of the details of these arguments is beyond the scope of this paper, but it is important to understand, at least at a high level, what is meant by psychological measurement. Research psychologists have offered various and mostly similar definitions over the years. Stevens (1946) defined measurement as “the assignment of numerals to objects or events according to rules” (p. 677) – a somewhat vague definition that was later made slightly more specific by Lord, Novick, & Birnbaum (1968), and Torgerson (1958) by their assertion that numbers were assigned to the *attributes* of objects and not to the objects themselves. Consider the realm of physical measurement; when a scientist measures the mass of some compound the measures taken are of a specific attribute of the object, not of the actual object itself. Similarly, psychologists do not take measurements of actual individuals (typically), rather they measure attributes of the individual, for example a person's

intelligence, anxiety level, motivation, etc. Unlike physical attributes, these psychological attributes cannot be measured directly. They are constructs, i.e., hypothetical concepts created from the informed imaginations of research psychologists for the purpose of explaining human behavior. Given that constructs are made-up, they can never be absolutely confirmed. Thus, how closely a particular psychological construct depicts an individual can only be determined by observations of his or her behavior (Crocker & Algina, 1986).

Crocker & Algina, (1986) provide an example of the process of construct formation and how that leads to the measurement of attributes. Consider a developmental psychologist who works with children in a preschool on a daily basis. The psychologist notices that certain children frequently try to direct the activities of the other children during recess. After observing similar behavior consistently for the same children, the psychologist labels the behavior “socially dominating.” The psychologist has created (constructed) a theoretical construct represented by a collection of behaviors. However, inventing a construct is not the same as measuring the construct. Before measurement can occur the theoretical construct must be “operationalized,” i.e., some rules must be created linking the theoretical idea to tangible, observable behaviors. In this case, measuring social dominance requires that the psychologist determine which specific behaviors in the preschool setting are considered “dominating.” The psychologist must then come up with a plan for systematically obtaining samples of the children’s behavior. This systematic plan for obtaining samples of the children’s behavior is a test; and when a quantitative value is assigned to the sampled behavior that is measurement. In other words, measurement has occurred when, for example, the psychologist counts and records the number of dominating behaviors displayed by the child in a specified period of time. Based on these measurements of

observations the psychologist might make inferences about how much of the theoretical construct characterizes the child.

The example above represents a somewhat simplified rendering of the overall scientific process. Indeed, as pointed out by Cliff (1993), the central challenge of science in general is the identification of the principle variables and the demonstration of which things are the same and which things are different. This is particularly challenging in the social sciences because social scientists often observe similar behaviors but label them differently. The psychologist in the above example observed children's behavior and called it "social dominance;" another psychologist might see similar behavior and call it "social assertiveness," etc. This all too frequently leads to the problem of construct proliferation in which similar or identical theoretical ideas are given unique labels and are treated as separate constructs. Construct proliferation in the social sciences is detrimental because it has the long-term effect of impeding progress towards identifying the principle variables that make up human behavior.

Construct proliferation is not the only problem associated with measuring psychological constructs. Crocker & Algina, (1986) discuss five measurement problems that are common to all psychological assessments. First, there is seldom universal agreement among researchers about which observable behaviors best depict a construct. Because measurements of psychological variables are always indirect, differences of opinion often exist regarding the behaviors that best represent the construct under study. For example, let us say that a researcher is interested in measuring a student's skill at long division. Given that it is impossible to look into the student's head to determine how much long division he or she knows, the researcher must select observable behaviors that (or that the researcher believes) represent long-division skill. The researcher then makes inferences about the student's skill in long division based on their

performance on the selected behaviors. However, different operational definitions of long division are possible. One approach might be to ask students to solve a series of long division problems; another could be to ask the students to describe, sequentially, the steps involved in long division; yet another is to ask them to detect errors in a series of long division solutions. Different measurement procedures would be required for these differing operational definitions, which would likely lead to different conclusion about the student's level of skill.

A second measurement problem is that psychological measurements, by definition, are never based on an exhaustive set of behaviors. In the example above, it would be impossible to present to students all possible long division problems, so only a limited sample of behavior is used. Determining the variety of content needed to adequately represent the skill of long division is a significant challenge in developing measurement procedures (Crocker & Algina, 1986).

Measurement error is a third problem facing any psychological assessment (Crocker & Algina, 1986). Continuing with the example above, if a student takes a long division test two or three times in succession, it is unlikely that he or she will obtain the identical score. Students experience fatigue, boredom, carelessness, misscoring, etc., or they take a different form of the test with slightly different content and, as a consequence, their scores differ. These score inconsistencies must be regarded as error

A fourth problem in psychological measurement relates to defining and labeling the properties of the measurement scale (Crocker & Algina, 1986). If one student scores a 5 on the assessment designed to measure long division skill and another scores 10, and still another scores 15, what can be concluded from these scores? Does one student possess twice the skill as the other? Are the three students' equally spaced on the ability/skill continuum? What should be concluded if an examinee answers none of the questions? Does that student have zero skill with

regard to long division? Thus, defining the properties of the measurement scale, accurately labeling the scale units, and properly interpreting the values obtained from the test are serious challenges that must be addressed in any measurement procedure.

A final problem highlighted by Crocker & Algina, (1986) is that constructs cannot be defined in isolation, they must also have demonstrated relations with other variables, and those relations should coincide with expectations generated from theory. Sometimes called a nomological network of relations, any construct must be defined based on its logical or mathematical relation to other constructs that are in the same theoretical system. Continuing with our example above, high scores on a long division test should correlate fairly well with other measures of like skills (e.g., multiplication skill). Additionally, the high scores should also *not* correlate with completely dissimilar constructs (e.g., political conservatism).

In sum, measurement is a serious and challenging endeavor for everyone involved in the scientific enterprise. This is especially true for scientists working in psychology or other social sciences because their entire content domain is made up. Conceptual and operational definitions of constructs literally give life to new constructs/variables that hopefully represent fundamental aspects of the human condition. Obtaining quality measurements of constructs requires careful consideration and resolution of several problems that challenge every measurement situation.

Unfortunately, research on regulatory focus has largely ignored these measurement fundamentals and the quality of regulatory focus scales has suffered as a result. Research attention is needed to precisely specify the theoretical underpinnings of regulatory focus. For example, researchers should explain why self-regulation around nurturance needs and actual-ideal discrepancies are considered conceptually similar. Or, if empirical evidence suggests (as it seems to) that the two concepts are independent of one another, then RFT should be updated so

that regulatory focus is defined as a higher-order construct consisting of four sub-dimensions not two. Additionally, researchers should explain whether and how regulatory focus differs from other similar variables like approach/avoidance motivation, fear-of-failure vs. desire-to-succeed, or need for achievement (to name a few). Some work has been done in this area, looking at personality antecedents and work-related consequences of regulatory focus (Gorman et al., 2012), but more specific studies should be conducted, directed at the overall construct validity of regulatory focus. Once the theory and conceptual definition of regulatory focus is improved it is likely that some of the other measurement fundamentals mentioned above could improve as well. Improved consensus about what specific behaviors represent regulatory focus may emerge, which could lead to a wider range of tools to measure regulatory focus and even to reduced measurement error. Having defined measurement and discussed some of the fundamental challenges, I now move on to discuss the procedures that should be followed in order to develop a high quality measurement instrument.

Measurement best practices

Given the importance and centrality of measurement to the scientific process, one might expect that most social scientists would pay close attention to rigorous procedures in developing measures of theoretical constructs. In fact, as we have seen in our previous discussion of RFT, this is not always the case. (Hinkin, 1995) reviewed scale development practices of researchers and found that inadequate measurement development procedures are utilized far too frequently. A common practice in developing new scales is to simply write items that seem to represent the construct of interest, and then call it a measure without any attention paid to reliability and validity concerns. Hinkin, (1995) presents a model for scale development that includes three stages: item generation, scale development, and scale evaluation. I am including a fairly detailed

discussion of Hinkin's recommendations below because I plan to follow these proscriptions in the development of a new scale measuring regulatory focus.

Stage 1 - Item Generation: According to Hinkin's (1995) study, the first stage of good measure development is item generation. Item generation is an iterative process in which items are written to comprehensively represent the theoretical construct of interest. The measurement instrument must also contain as little extraneous content as possible. The measure is said to achieve the primary goal of stage one (content validity) if items are written that successfully and comprehensively cover the domain of interest without adding extraneous content. And even though there is no quantitative index associated with content validity, experts seem to agree that there are two primary approaches to achieving content valid measure: deductive approaches and inductive approaches.

Deductive approaches (the most frequently used) use a typology or classification schema to guide the creation of items. The schema/typology is informed through a careful review of the theoretical literature related to the construct. After developing items based on a conceptual understanding of the theory, the items then should be independently reviewed by subject matter experts who can judge whether the items reflect the construct. An excellent example of this process is described in detail in MacKenzie, Podsakoff, & Fetter (1991). The authors first developed items representing the five dimensions of organizational citizenship based on the organizational citizenship theory (Organ, 1988). Ten faculty members and graduate students then performed a content validity analysis on the items by classifying them into one of the five categories or an "other" category. Those items that were correctly classified into the appropriate dimension up to 80% of the time were retained for use in the measure. This is a great example of a proper item development process.

Inductive approaches to item development do not utilize theory as a guide for developing items; rather they attempt to generate items based on interviews from individuals possessing the attitude/trait/behavior. So, for example, with the regulatory focus construct, inductive item development might involve asking individuals questions about the attachment style they had with their parents, and about whether they have a risk aversion or gains and opportunities orientation to the world. Responses can then be organized/categorized using content analysis in which items are grouped according to key words or themes. After item generation, whether a deductive or inductive approach is used, the items should be screened again for conceptual inconsistency.

According to Hinkin (1995) researchers frequently violate these and other item development best practices. For example, instead of explaining the theoretical basis and item development procedures, Greenhaus, Parasuraman, & Wormley (1990) simply state that measures were “developed expressly for this study” omitting any other details about how items were developed. This is problematic because it eliminates any ability to critically evaluate the author’s procedures and because sometimes item content that appears to represent the domain(s) of interest actually does not. For example, Ettlíe & Reza (1992) developed a scale to measure “customer integration” as part of a larger organizational behavior study, but the scale’s internal consistency was low ($\alpha = .48$) which means that over 50% of score variability is a function of random error. Hinkin (1995) also noted that even well-thought-out item development procedures can produce scales that do not withstand the empirical scrutiny of factor analysis or other techniques.

The wording of items is also an important consideration. Specifically, Hinkin (1995) recommends against negatively worded items that would require reverse scoring. Despite the historical use of reverse-score items to attenuate response bias (see Idaszak & Drasgow, 1987)

the more recent consensus is based on work that suggests that negatively worded items introduce more problems than they solve, including reduced validity (Schriesheim & Hill, 1981) systematic error (Jackson, Wall, Martin, & Davids, 1993) and negatively-worded-item response factors (Schmitt & Stults, 1985).

Stage 2 – Building the Scale: After completing the item development stage, the researcher should have a set of content valid items relating to the theoretical construct of interest. The second stage in the measure development process involves building the scale, which is comprised of three steps: 1) designing the scale development study, 2) scale refinement and reduction, and 3) reliability assessment.

Stage 2, Step 1 - The scale development study step involves organizing the items into a scale which can then be administered and evaluated psychometrically and in terms of how well the scale relates to other similar and dissimilar constructs. Several specific concerns are relevant to this step. First, items should be administered to a sample that is representative of the population of interest and to which the results will be generalized. Arguments should be constructed explaining the rationale for the sample chosen. Additionally, the sample size should be large enough so that the various multivariate statistical techniques commonly used will have sufficient power (Tabachnick & Fidell, 2007).

Second, scale length (i.e., the number of items used) should receive careful consideration. Generally speaking, single-item measures should be avoided, despite arguments to the contrary (Wanous, Reichers, & Hudy, 1997), because internal consistency estimates cannot be obtained for single-item measures. Internal consistency statistics (like Chronbach's Alpha) serve as estimates of the amount of random variance within a set of scale scores. The inability to derive these estimates in single item measures means that one has no error estimate for that measure.

Despite the fact that adding items to a scale does improve scale reliability (Crocker & Algina, 1986), scales that are too long can also be problematic in terms of response fatigue or response bias (Anastasi, 1976). More practically, Carmines & Zeller, (1979) discussed how time (and therefore cost) increases are associated with development and administration of longer measures. Keeping scales short attenuates problems like response bias and fatigue, but measures that are too short may result in problems with overall construct validity as well as reliability.

A third consideration during the scale development step pertains to the scaling of individual items. The main issue is that the scale must generate sufficient variance in scale responses to accommodate the various statistical analyses that will be run on the data in later stages. Consensus among researchers suggests that when using Likert (or Likert-like) scales, five to seven response anchors are optimal (Hinkin, 1995).

As briefly alluded to above, sample size is another consideration during the scale development step. Both confirmatory and exploratory factor analyses are vulnerable to sample size effects (Hinkin, 1995). Large sample sizes allow for stable estimates of standard error which increases confidence that factor loadings represent population values. Hinkin's (1995) review suggests that minimum sample sizes of between 150 to 200 observations should be sufficient for most scale development efforts, although some researchers (Crocker & Algina, 1986) recommend a ratio of 10:1 (observations to items) as a best practice.

Stage 2, Step 2 - The second step associated with scale development is scale refinement and reduction. This primary purpose of this step is to use exploratory and confirmatory factor analyses to examine the structure of the scale. Several specific best practices are noted in Hinkin (1995). First, factor analysis eigenvalues (an eigenvalue is a statistic in factor analysis that indicates the relative proportion of total scale variance that is explained by the factor) of greater

than one should be used as a criterion for factor retention (but the higher the eigenvalue the better). This criterion is used because, in factor analysis, the variance of each individual variable is one. Thus, it is reasonable to only retain factors which explain more variance than a single item would explain.

Next, assuming one has progressed from use of an exploratory factor analysis and is now conducting a confirmatory factor analysis to confirm the factor structure of the new scale on a fresh sample, any items that do not load on the predicted factors should be omitted from the scale. As Hinkin (1995) points out, retaining items that do not load as predicted can lead to low scale internal consistency reliability. Additionally, one should only retain items with factor loading equal to or greater than .40 (the factor loading is the correlation between the item and the factor, so low correlations likely indicate that the item does not “belong” with the factor) and all factor retention and loading information should be succinctly described. Researchers should minimally present the following five pieces of information after conducting the necessary factor analysis: 1) items and related factor loadings; 2) factoring and rotational method; 3) criteria used to determine the number of factors retained; 4) the percentage of variance accounted for both by factor and by the total scale; and 5) the rationale for retention and deletion of items.

Once the internal scale integrity is confirmed, a third concern while building the scale is assessing the model fit to the data. CFA “fit” involves comparing the sample-derived covariance matrix to the estimated population covariance matrix. A “good fitting model” is one in which the sample and estimated covariance matrices approximate one another. So far, the only actual test statistic associated with model fit is the chi-square test. Unfortunately, chi-square is particularly sensitive to sample size, so even slight differences between matrices are often deemed significant (indicating model non-fit). To address this problem researchers have developed a number of

different, descriptive, model fit indices. According to Hinkin (1995) there is no real consensus on what fit indices are best, however most recommend using Chi-square, Chi-square-/DF, GFI, CFI, SRMR,, and the RMSEA. As a general rule of thumb, the closer the Chi-square statistic is to the number of degrees of freedom the better, the closer GFI, CFI, and RFI are to one the better, and the lower the RMSEA is the better (.05 is a commonly used threshold).

As mentioned, the overall objective of this step is to evaluate the structure and integrity of the newly formed scale and to be as transparent as possible with the various decisions made to retain or delete items, model fit, etc. If the researcher chooses to delete items based on low factor loadings or other relevant reasons, the “new” scale should be analyzed again using a new sample. This practice safeguards against constructing and fitting the new scale to the characteristics of one specific sample.

Stage 2, Step 3 - The third step associated with the scale development stage pertains to assessment of the reliability of the scale. Primarily at this step one is concerned with two things: the internal consistency of the new scale (or of its subscales if the construct is multidimensional), and how stable it is over time. Coefficient alpha is the most frequently used statistic for assessing internal consistency, and Nunnally, (1994) recommends .70 as a minimum cutoff. It is important to note that .70 is a minimum cutoff. Keep in mind that an alpha of .70 means that 30% of the variance in scale scores is due to random error. In some measurement circumstances, this much error may be unacceptable and thus higher standards should be retained. A note of caution is in order: Hinkin (1995) asserts that low scale reliability is largely a byproduct of problems at the item generation and scale construction steps. Careful attention to these steps can yield scales with high levels of internal consistency.

Test-retest reliability should be used to assess the stability of the scale over time. As pointed out by Hinkin (1995) however, the test-retest statistic only makes sense when dealing with construct that, theoretically, should not change over time. For example, a measure of a personality trait, like regulatory focus, is theorized to be relatively stable over time and thus should be verified using test-retest reliability statistic. Conversely, a more fluid construct like work engagement would be expected to vary over time, and thus test-retest reliability assessment may not be appropriate. Please note that establishing test-retest reliability is beyond the scope of the current research, and thus is left for future research.

Stage 3 – Scale Evaluation After completing the scale development stage, attention shifts to the evaluation of the new measure. The main concern here is beginning to establish construct validity. Establishing construct validity is an ongoing process and not something that can be obtained in a single study. Indeed, a primary technique used to establish construct validity is to show how the new measure relates to other similar and dissimilar constructs. The nomological network of the construct is a tapestry of construct relations formed over time by a collection of research. In fact, the validity of constructs only emerges as this network of relations is established. An important component of the nomological network is demonstrating the discriminant and convergent validity of the new measure. Similar constructs should have moderate to high correlation coefficients and dissimilar constructs should have low, negative, or zero correlation coefficients. According to Hinkin (1995) less than 25% of the over 270 scales he evaluated addressed construct validity at all. In fact, many researchers he evaluated erroneously concluded that demonstrating a stable factor structure of a new measure is evidence of construct validity. In the current research a measure of positive and negative affectivity (PANAS; Watson, Clark, & Tellegen, 1988) is used to test the discriminant validity of the new regulatory focus

measure. Prior research (Haws, 2010; Summerville & Roese, 2008) has suggested that regulatory focus is unrelated to both positive and negative affect. An important feature of the current work is to test whether the relation (or lack of) is maintained with the new measure.

In sum, building a new measurement instrument is an involved and time-consuming process. Items must be written carefully and precisely to adequately reflect the underlying theoretical concept; these items must be organized into a scale that is not too long, nor too short, with response anchors that will generate variability; and then they must be administered to the appropriate sample based on the anticipated future use of the new measure. Factor analyses should be used to assess the structure of the new scale and fit indices like Chi-Square should be used to confirm the model fits the data. The new scale must have at least adequate reliability so that scores can be interpreted confidently. And concerns related to construct validity must be attended to continuously over time until the new measure is situated within its nomological network of similar and dissimilar constructs.

Applying these measurement best practices to the RFT literature illuminates many problems that the proposed research seeks to remedy. First, researchers associated with the scales mentioned above all failed to describe any of their scale development procedures. There is no discussion of how items were developed, no explanation regarding exactly what theoretical content the scales were trying to measure, and no description of the scale development study (e.g., sample used, scale reduction techniques etc.). Second, all of the measures discussed in this proposal lack evidence of content validity. This problem can be remedied by following Hinkin's best practice to write new items that comprehensively tap into the theoretical dimensions of regulatory focus. Third, negatively worded and double-barrel items found in both the RFQ and GRFM should be eliminated. This should help to improve the internal consistency of the

measure. Fourth, none of the scales mentioned above were scrutinized after development with regard to their construct validity. Beginning to establish the construct validity of regulatory focus is important. For example, how does regulatory focus differ from other similar constructs like positive or negative affect, approach/avoidance motivation, fear-of-failure vs desire-to-succeed, or need for achievement to name a few. In the current research, I include evidence suggesting that the new regulatory focus scale is not related to positive or negative affect (why this is important is explained in more detail in the results section); a prediction established in existing regulatory focus theory and extant empirical research (Haws, 2010; Summerville & Roese, 2008).

Chapter 2 - Method

Item Generation

A pool of 59 items was generated by two researchers with expertise in regulatory focus theory and scale development. Items were constructed deductively i.e., based on a comprehensive review of the regulatory focus theoretical literature (Higgins, 1997, 1998; Higgins et al., 1985) and were designed to measure all aspects of regulatory focus theory. The nurturance aspect of RFT was reflected in items like: “*Whenever possible, I try to improve my current situation*”. The security aspect of RFT was reflected in items like: “*I prefer to maintain the status-quo*”. The actual-ideal aspect of RFT was reflected in items like: “*I strive to be my ideal self*”. The actual-ought aspect of RFT was reflected in items like: “*I view goals as duties I must complete*”. The initial item pool (see Appendix B for full list) included 16 items designed to measure nurturance, 14 items designed to measure security, 14 items written to measure actual-ideal orientation, and 15 items designed to measure actual-ought orientation. Prior research (Haws, 2010; Summerville & Roese, 2008) has questioned the psychometric properties and item quality of existing measures of regulatory focus. Thus, for the present study, no items from existing measures of RFT were used; all new items were written. The item pool contained only forward-scored items as prior research suggests that reverse-scored items may lower scale validity (Schriesheim & Hill, 1981), introduce systematic error to a scale (Jackson et al., 1993) and may even result in an artificial response factor consisting of negatively worded items (Schmitt & Stults, 1985). All items were written to be comprehensible to individuals with no prior experience or exposure to regulatory focus theory.

All 59 items were sent to five subject matter experts for independent review. The subject matter experts (SME) included three academic psychologists with expertise in regulatory focus

theory, one academic psychologist familiar with regulatory focus theory and with expertise in testing and assessment, and one doctoral candidate in I/O Psychology who is familiar with regulatory focus theory and with good item writing techniques. The SME's read a brief description of regulatory focus theory and then rated each item based on whether they thought it represented a promotion item, a prevention item, or other. SME's also provided feedback on the quality and clarity of the wording of each item. Items that exceeded 80 percent agreement among raters and that contained no wording ambiguity were retained for the final item pool. The final item pool consisted of 39 items and is shown in Table 1 (shaded items were retained, non-shaded items were eliminated). Each reviewer's rating of the items is presented (pro = promotion item; pre = prevention item; oth = other) along with an aggregate agreement rating (a rating of 5 indicates 100% consensus on the item), notes, and whether the item was included in the final scale (indicated by a Y or N in the Final Scale column).

Table 1. *Subject matter expert ratings of initial regulatory focus items.*

Item #	Item	Final Scale?	Rev 1	Rev 2	Rev 3	Rev 4	Rev 5	Agree	Notes
38	When I succeed I feel a sense of joy	Y	pro	pro	pro	pro	pro	5	
1	I think about how I will achieve my hopes and aspirations	Y	pro	pro	pro	pro	pro	5	
3	I strive towards my ideals	Y	pro	pro	pro	pro	pro	5	
4	I feel a strong sense of duty to those around me	Y	pre	pre	pre	pre	pre	5	
5	I am more focused on achieving success than avoiding mistakes	Y	pro	pro	pro	pro	pro	5	
6	I hate making mistakes	Y	pre	pre	pre	pre	pre	5	
8	I strive to meet my obligations	Y	pre	pre	pre	pre	pre	5	
10	I play not to lose	Y	pre	pre	pre	pre	pre	5	
11	Success to me is achieving my aspirations	Y	pro	pro	pro	pro	pro	5	
13	I am comfortable with change	Y	pro	pro	pro	pro	pro	5	
14	I prefer to maintain the status-quo	Y	pre	pre	pre	pre	pre	5	
15	I think about how to achieve an optimal life	Y	pro	pro	pro	pro	pro	5	
17	I like winning more than I dislike losing	Y	pro	pro	pro	pro	pro	5	
18	I do not like change	Y	pre	pre	pre	pre	pre	5	
20	Success to me is meeting my responsibilities	Y	pre	pre	pre	pre	pre	5	

23	I pursue goals eagerly	Y	pro	pro	pro	pro	pro	5
24	I pursue goals cautiously	Y	pre	pre	pre	pre	pre	5
30	I pursue ideals in life	Y	pro	pro	pro	pro	pro	5
33	I dislike losing more than I like winning	Y	pre	pre	pre	pre	pre	5
34	I view goals as opportunities not to be missed	Y	pro	pro	pro	pro	pro	5
35	I view goals as duties I must complete	Y	pre	pre	pre	pre	pre	5
36	I like new opportunities	Y	pro	pro	pro	pro	pro	5
39	When I succeed I feel a sense of calm	Y	pre	pre	pre	pre	pre	5
40	I play to win	Y	pro	pro	pro	pro	pro	5
42	When I fail I feel dejected	Y	pro	pro	pre	pro	pro	5
43	When I fail I feel anxiety	Y	pre	pre	pre	pre	pre	5
46	I strive to be my ideal self	Y	pro	pro	pro	pro	pro	5
47	I am concerned that I will fall short of my duties	Y	pre	pre	pre	pre	pre	5
49	I think about how I can prevent failures in my life	Y	pre	pre	pre	pre	pre	5
50	Advancing in life is important to me	Y	pro	pro	pro	pro	pro	5
52	Whenever possible, I try to improve my current situation	Y	pro	pro	pro	pro	pro	5
53	I am focused on preventing negative events in my life	Y	pre	pre	pre	pre	pre	5
54	I savor winning	Y	pro	pro	pro	pro	pro	5
57	I am a cautious person	Y	pre	pre	pre	pre	pre	5
59	I am really hard on myself when I make mistakes	Y	pre	pre	pre	pre	pre	5
60	I have lofty aspirations for life	Y	pro	pro	pro	pro	pro	5
9	When I make a mistake I feel sad	Y	pro	pro	pro	pro	pro	5
7	I tend to be a spontaneous person	N	oth	pro	pro	pro	oth	3
16	I tend to be a responsible person	N	oth	pre	pre	pre	oth	3
19	I live a fast-paced life	N	oth	pro	oth	pro	pro	3
27	I pay attention to the big picture	N	oth	pro	oth	pro	oth	3
37	I pay attention to details	N	oth	pre	oth	pre	pre	3
56	I am a carefree person	N	oth	pro	pro	pro	pro	3
58	I like to strike out on my own by setting my own goals	N	oth	pro	oth	pro	pro	3
21	I like changes in life	N	pro	pro	pro	pro	oth	4
22	I prefer not to take risks in life	N	oth	pre	pre	pre	pre	4
25	I see mistakes as opportunities to grow	N	oth	pro	pro	pro	pro	4
26	I live a slower-paced life	N	oth	pre	pre	pre	pre	4
28	I take risks in life	N	oth	pro	pro	pro	pro	4
32	I do not like missing opportunities	N	pro	pro	pro	pro	pre	4
41	I usually follow the rules	N	oth	pre	pre	pre	pre	4
44	I am achievement oriented	N	oth	pro	pro	pro	pro	4
48	I hate missing out	N	pro	pro	pro	pro	pre	4
55	My responsibilities are important to me	N	pre	pre	pre	pro	pre	4

2	Whenever possible, I try to maintain my current situation	N	pre	pre	pre	pre	pre	5	vague
12	I see mistakes as betraying my duty	N	pre	pre	pre	pre	pre	5	awkward
29	I like the routines of life	N	pre	pre	pre	pre	pre	5	vague
31	I pursue obligations in life	N	pre	pre	pre	pre	pre	5	vague
45	I am more focused on avoiding mistakes than achieving success	N	pre	pre	pre	pre	pre	5	
51	I strive to be the person I am obligated to be	N	pre	pre	pre	pre	pre	5	awkward

Participants

An online (Qualtrics) sample of 619 individuals responded to the initial set of scale items during the week of April 24th 2017. Researchers (Hinkin, 1995; Tabachnick & Fidell, 2007) recommend a subject to item ratio of at least 10:1 for scale development studies. This is especially important when using multivariate techniques like factor analysis because results have been shown to be sample specific with smaller sample sizes (Schwab, 1980). Utilizing a large sample size may ameliorate some of these issues. Further, a large sample is necessary so that final CFA results can be validated against an adequately sized hold out sample. Scholars have also raised concerns about online samples (e.g., Couper, 2000), arguing mainly that they suffer from self-selection bias and thus limit generalizability. However, Highhouse & Gillespie (2009), cite findings from five meta-analyses and make an empirical argument that the specific nature of the sample does not impact the (theoretical) generalizability of research findings. These findings, coupled with the ease of data collection associated with online samples, and the fact that online samples are likely more diverse (demographically) than a typical introductory psychology course sample, justifies the use of an online sample for the current research.

On average, participants were female (70.3%), employed either part or full time or retired (76.5%), 48 years old, making between \$50,000 and \$60,000 annually. A hold-out sample of 151 participants was created by randomly sampling participants from the main participant pool.

There were no significant demographic differences between the main sample and the holdout sample. The hold-out sample was used to validate the final CFA models described below.

Procedure

Survey respondents were asked to complete all questions and were given the following prompt “The following section contains questions pertaining to your general approach to life. When answering each question think about what is typically true for you. For example if a statement is very typical of you, answer ‘strongly agree’; if a statement is not very typical of you, answer ‘strongly disagree’. Participants were asked to rate their responses on a Likert scale with the following options: 1 (Strongly Disagree), 2 (Disagree), 3 (Neutral), 4 (Agree) or 5 (Strongly Agree). Administration of the online survey constituted the data collection procedures approved by the Kansas State University Institutional Review Board and participants indicated informed consent before participating. The order of the items presented to the participants was randomized to prevent any potential problems with order effects.

Chapter 3 - Results

Item reduction and preliminary model testing

A reliability analysis of the original pool of 39 items was examined to guide the first stage of item reduction. Ten items that were lowering the internal consistency of the promotion and prevention subscales ('alpha if removed' > .80) were eliminated. Two survey items were redundant (i.e., the items were identical) to other items and were also eliminated. Thus, a preliminary item pool of 27 items was retained. Data screening procedures following Tabachnick & Fidell (2007), were conducted using SPSS programs. There were no missing data. Linearity, normality, and homoscedasticity assumptions were checked and confirmed using a normal probability plot of standardized residuals (to test linearity), a regression of standardized residuals onto a dummy dependent variable (to test homoscedasticity), and a visual check of the normal curve superimposed over histograms for each item to "test" normality. Multicollinearity was evaluated using a conservative variance inflation factor (VIF) threshold of three (larger VIF values indicate higher standard error). Using an iterative process, 38 of the 39 independent variables (IV) were regressed onto one IV which was used as a dependent variable (DV). All VIF values were below 3 indicating no issues with multicollinearity in the data set. To check for univariate outliers, item scores were converted to zscores. Outliers were those cases that exceeded the $z \pm 3.29$ criterion. There were 43 univariate outlier cases which were deleted from the data set. Multivariate outliers were checked using Mahalanobis Distance (MD), which were computed by regressing all 27 scale items against a dummy dependent variable (ID#). MD values were then compared against a chi square distribution table. Any values that fell below a $p < .001$ were identified as outliers. There were 25 multivariate outliers which were deleted from

the data set. Thus, 551 cases were retained for model development and confirmatory factor analysis (CFA).

As mentioned, 551 cases were retained for model development and CFAs. Factor models were tested with the structural equation modeling software program AMOS 18 (Arbuckle, 2009) using maximum likelihood estimation (MLE). MLE is the most commonly used estimation technique and made sense in the current research because data was normally distributed. A primary concern in CFA is assessing the fit of the proposed model(s) to the data (more specifically, the fit between the sample and estimated population covariance matrices). “Fit”, as noted by Tabachnick & Fidell (2007), is commonly tested against a non-significant χ^2 as the criterion. However, with large sample sizes (such as with the current research) trivial differences between the sample and estimated covariance matrices are often significant because the calculation of χ^2 involves multiplying the minimum of the function by $N - 1$ (in the current research $N - 1 = 550$). Thus, researchers have proposed a variety of other descriptive model fit indices; and which indices are used is a matter of personal preference for the researcher (Tabachnick & Fidell, 2007). Fortunately, good-fitting models tend to produce consistent result on the various indices, so if several indices suggest good model fit to the data, chances are the other indices will be good as well. For the current research, the following test statistic and indices were used to assess goodness of fit: χ^2 , χ^2 / df comparative fit index (CFI), goodness-of-fit index (GFI), root mean squared error of approximation (RMSEA), and the standardized root mean square residual (SRMR). The CFI is important because it is a relative index i.e., it compares the estimated population covariance matrix against both an independent model (comprised of unrelated variables) and a saturated (perfect fitting) model. The resulting fit index lies in the 0 to 1 range with values approaching 1 considered optimal. Hu & Bentler (1999),

recommend a CFI value of .95 (or higher) as indicative of a good fitting model. RMSEA is also important because it provides a slightly different fit assessment. RMSEA estimates the lack of model fit compared to a saturated (perfect) model. When model fit is perfect the RMSEA yields a value of zero; as model misspecification increases the RMSEA value increases indicating a poor fitting model. Hu & Bentler (1999), recommend a RMSEA value of .06 (or less) as indicative of a well-fitting model. SRMR is also commonly used in assessing model fit because it uses the residual values from the comparison of the sample and estimated population covariance matrices. Good fitting models will have small residual values because the sample covariance matrix and the estimated population covariance matrix approximate one another. The SRMR values range from 0 to 1, with values of .08 or less indicating good model fit (Hu & Bentler, 1999). The final fit index used in the current research (GFI) is analogous to R^2 in regression in that it calculates the proportion of variance in the sample covariance matrix that is explained by the estimated population covariance matrix (Tabachnick & Fidell, 2007). Values range from 0 to 1 with values of .95 or higher indicating greater variance explained, and thus a better fitting model. In sum, the current research uses a variety of model fit indices as criteria for testing model fit: χ^2 , χ^2/df $F \leq 2$ (Tabachnick & Fidell, 2007); for CFI, GFI, RMSEA, and SRMR the current research follows the cut-off recommendations of Hu & Bentler (1999) mentioned above.

The twenty seven items that were retained from the initial item reduction were divided into their theoretically-derived categories to test both a four factor and a two factor model of regulatory focus. The first CFA models were tested on the main sample (N=400) with two first-order latent factors (promotion regulatory focus and prevention regulatory focus) each indicated by the measured items that reflect promotion and prevention regulatory focus. Both factors were tested simultaneously. The 27-item model demonstrated moderate fit.

After exploring the patterns of item loadings in the two-factor model, items were eliminated in a series of iterative factor models. Items were eliminated if they demonstrated poor loadings on their hypothesized factor. Per Comrey & Lee (1992), factor loadings exceeding .71 are considered excellent, .63 very good, and .55 good. Consequently, a factor loading cut-off of .50 was used so that the shared item-factor variance was greater than 25 percent. All tested models were congeneric, which means that items were allowed to load on only one of the two first-order factors and measurement error of the items was not permitted to be correlated. Through this iterative process, a final model emerged that was a good fit to the data: [$\chi^2(76)=162.00, p=.000; \chi^2/DF = 1.97; RMSEA=.053; SRMR=.04; CFI=.96; GFI=.95$]. This model appears in Figure 2. The RMSEA, SRMR, CFI, and GFI were within the cut-off range recommended by Hu and Bentler (1999). The chi-square test was significant (an indication of model non-fit); however, it should be noted that the chi-square statistic has been criticized as an overly-sensitive test that frequently suggests rejecting potentially useful models, particularly as N becomes large (Bollen, 1989).

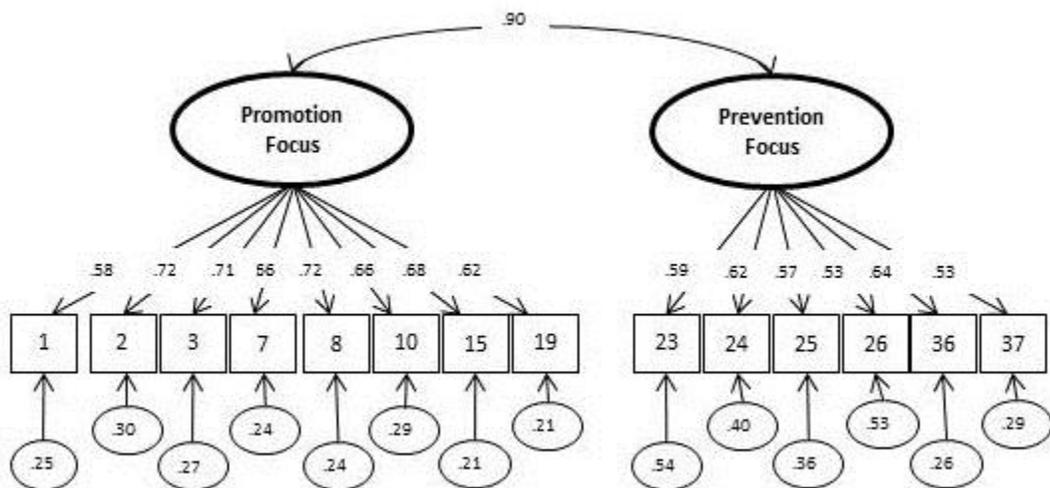


Figure 2. Standardized parameter estimates for a first-order confirmatory factor analysis (AMOS) for the Promotion Prevention Scale (PPS). Numbers in boxes correspond to PPS questionnaire items presented in Table 2.

Consistent with the initially proposed two-factor model, this model consisted of two first-order latent factor (promotion regulatory focus and prevention regulatory focus). The final 14 retained items appear in Table 2 along with their descriptive statistics across samples. In the final model, the promotion subscale had eight items, and the prevention subscale had six thus satisfying the recommended minimum number of observed variables per latent variable (Kline, 2004; Wegener & Fabrigar, 2000). All retained items had standardized loadings on their first-order factors above .50 (range = .53 to .72).

Table 2. *Descriptive statistics for the final items of the Promotion Prevention Scale (PPS)*

	Sample 1 (N = 400)				Sample 2 (N=151)			
	M	SD	Skew	K	M	SD	Skew	K
1. When I succeed I feel a sense of joy	4.34	0.56	-0.12	-0.72	4.34	0.65	-0.47	-0.69
2. I strive towards my ideals	3.98	0.66	-0.3	0.28	3.99	0.69	-0.37	0.24
3. I think about how to achieve an optimal life	3.83	0.77	-0.38	0.07	3.79	0.76	-0.56	0.79
7. I strive to be my ideal self	4.06	0.66	-0.27	0.07	3.99	0.67	-0.26	0.09
8. I think about how I will achieve my hopes and aspirations	3.94	0.71	-0.49	0.31	3.9	0.75	-0.52	0.33
10. Success to me is achieving my aspirations	4.01	0.7	-0.31	-0.02	3.89	0.7	-0.31	0.11
15. I view goals as opportunities not to be missed	3.9	0.75	-0.3	-0.16	3.76	0.72	-0.16	-0.17
19. Whenever possible, I try to improve my current situation	4.18	0.64	-0.46	0.59	4.11	0.63	-0.25	0.19
23. I strive to meet my obligations	4.24	0.67	-0.48	-0.13	4.19	0.64	-0.5	0.75
24. Success to me is meeting my responsibilities	4.13	0.65	-0.3	-0.04	4.05	0.66	-0.48	0.76
25. I view goals as duties I must complete	3.72	0.89	-0.44	-0.28	3.64	0.91	-0.76	0.33
26. When I succeed I feel a sense of calm	4.03	0.71	-0.46	0.21	3.98	0.67	-0.52	0.88
36. I think about how I can prevent failures in my life	3.74	0.83	-0.51	0.12	3.79	0.85	-0.62	0.26
37. I am focused on preventing negative events in my life	3.78	0.87	-0.45	-0.26	3.87	0.89	-0.96	1.46

A primary purpose of the current research was to compare the fit between a two-factor model of regulatory focus (with promotion and prevention regulatory focus as latent factors) and a four-factor model of regulatory focus. As discussed in the introduction, a four-factor solution may offer superior fit because regulatory focus might be better conceptualized as consisting of four latent factors: nurturance, actual-ideal, security, and actual-ought orientations. Thus, using the 14 remaining items from the initial CFA, a second CFA model was tested on the main sample with four first-order latent factors (nurturance, actual-ideal, security, actual-ought) each indicated by

the measured items that reflect the four potential dimensions of regulatory focus. All factors were tested simultaneously. The 14-item four-factor model demonstrated nearly identical fit to the two-factor model with all fit indices within the cutoff ranges suggested by Hu & Bentler (1999), [$\chi^2(71)=140.32$, $p<.000$; $\chi^2/DF = 1.97$; RMSEA=.05; SRMR=.04; CFI=.97; GFI=.95]. This model is shown in Figure 2. The chi-square test was also significant with this model, but, reiterating from above, the chi-square statistic has been criticized as an overly-sensitive test that frequently suggests rejecting potentially useful models, particularly as N becomes large (Bollen, 1989). A comparison of model fit statistics and indices for the two and four factor models is presented in Table 3, below.

Table 3. *Chi-square statistics and model fit indices for two and four factor models of regulatory fit.*

		Sample 1 (N = 400)					
		χ^2	χ^2/DF	RMSEA	SRMR	CFI	GFI
Two							
Factor							
Model	162	1.97	0.05	0.04	0.96	0.95	
Four							
Factor							
Model	140.3	1.97	0.05	0.04	0.97	0.95	

In sum, both model fit statistics and indices were very good. However, given that the four-factor model had only two items reflecting one of the sub-dimensions (security) making any internal consistency estimates for this factor suspect, and for reasons of parsimony (all other things being

equal, simple models are preferred over more complex), the two-factor model is preferred.

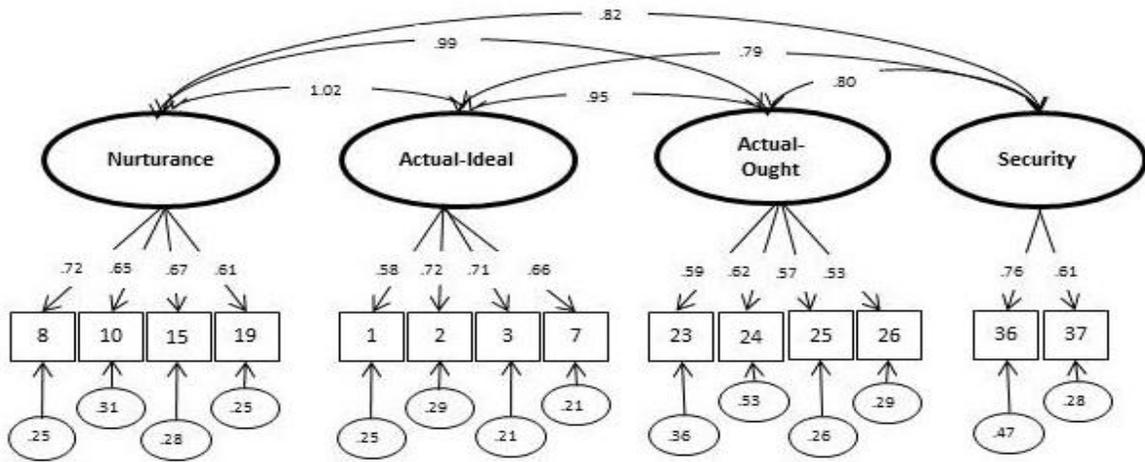


Figure 3. Standardized parameter estimates for a first-order confirmatory factor analysis (AMOS) for the Promotion Prevention Scale (PPS). Numbers in boxes correspond to PPS questionnaire items presented in Table 2.

Cross-validation

Because the model pruning approach can result in a model that achieves acceptable fit by capitalizing on chance error, both the two-factor and four-factor CFAs were performed on data drawn from an independent hold-out sample (N=151) to test the replicability of the models derived from the preliminary analyses. MLE was again used. The fit indices suggest that the correspondence between the proposed models and the data was acceptable. The test statistic and fit indices for the two-factor model were similar to those from the first sample: [$\chi^2(76)=120.93$, $p<.001$, $\chi^2/DF = 1.59$; RMSEA = .063; SRMR = .05; CFI = .95; GFI = .90]. Also, the test statistic and fit indices for the four-factor model were similar to those from the first sample: [$\chi^2(71)=112.54$, $p<.001$, $\chi^2/DF = 1.59$; RMSEA = .062; SRMR = .05; CFI = .96; GFI = .90]. A summary of all model fit statistics and indices are presented in Table 3. Given the similarity of model fit between the first and second samples, it is reasonable to be confident that the proposed model solution approximates the population.

Table 3. *Chi-square and model fit indices for the two and four factor models of regulatory focus.*

	Sample 1 (N = 400)						Sample 2 (N=151)					
	X ²	X ² /DF	RMSEA	SRMR	CFI	GFI	X ²	X ² /DF	RMSEA	SRMR	CFI	GFI
Two Factor Model	162	1.97	0.05	0.04	0.96	0.95	120.93	1.59	0.063	0.05	0.95	0.9
Four Factor Model	140.3	1.97	0.05	0.04	0.97	0.95	112.54	1.59	0.062	0.05	0.96	0.9

Internal consistency and factor inter-correlation

The 14 retained items associated with the two-factor model of regulatory focus demonstrated good levels of internal consistency (promotion subscale, $\alpha = .91$, prevention subscale, $\alpha = .75$).; Four factor model: nurturance subscale, $\alpha = .86$; ideal subscale, $\alpha = .81$; ought subscale, $\alpha = .60$; security subscale, $\alpha = .64$). Because the alpha coefficient is influenced by the number of items included in the measure (John & Benet-Martinez, 2000), it is not surprising that the actual-ought and security subscales of the four-factor model had low internal consistency. This is consistent with findings from previous research – the prevention subscales of the RFQ and GRFM also had lower internal consistency estimates. The magnitude of the intercorrelations between the latent variables, calculated in the two-factor CFA in which the latent variables were allowed to covary were large with $r = .90$ (Cohen, 1992).

Discriminant validity

A key finding from previous research on regulatory focus is that regulatory focus is unrelated to both positive and negative affect. This is an important distinction because if most of

the variance in regulatory focus can be accounted for by affect, then perhaps the regulatory focus construct is redundant and unnecessary. Previous research (Haws, 2010; Summerville & Roese, 2008) suggest that older measures of regulatory focus (like the RFQ and GRFM) are unrelated to positive or negative affect as measured by the PANAS scale (Crawford & Henry, 2004). Thus, an important step in the current research is to determine whether the PPS behaves similarly. To test this question, two, three-stage hierarchical regression analyses were conducted using the holdout sample of 151 participants to test the discriminant validity of the new measure.

Composite scores/variables were created for both promotion and prevention regulatory focus by adding the subscale scores and dividing by the total number of items in the subscale. The same procedure was used to create a composite score for positive and negative affect (PA, NA). In the first hierarchical regression, promotion was the dependent variable. Prevention was entered at stage one of the regression to control for any potential overlap of prevention with promotion. PA was entered at stage two and NA at stage three. In the second hierarchical regression, Prevention was the dependent variable. Promotion was entered at stage one of the regression to control for any potential overlap of promotion with prevention, and PA and NA were entered at stage two and three respectively. A summary of the regression statistics are presented in Table 4 and 5 below.

Table 4. Regression analysis predicting promotion regulatory focus from positive and negative affect.

Predictor	R^2	ΔR^2	B	Std Error	β
Step 1					
Prevention	.57	.57	.79	.06	.76
Step 2					
Prevention	.63	.06	.70	.06	.67
PA			.17	.03	.26
Step 3					
Prevention	.63	.001	.71	.06	.67
PA			.16	.04	.26
NA			-.02	.03	-.02

Note: **Bold** items indicate $p < .05$.

Table 5. Regression analysis predicting prevention regulatory focus from positive and negative affect

Predictor	R^2	ΔR^2	B	Std Error	β
Step 1					
Promotion	.57	.57	.72	.05	.76
Step 2					
Promotion	.57	.00	.74	.06	.78
PA			-.03	.04	-.05
Step 3					
Promotion	.58	.00	.74	.06	.77
PA			-.02	.04	-.03
NA			.05	.03	.08

Note: **Bold** items indicate $p < .05$.

The first hierarchical multiple regression revealed that at stage one, prevention contributed significantly to the regression model, ($F(1, 149) = 197.07, p < .05$) and accounted for 57% of the variation in promotion. Introducing PA in stage two explained an additional 6% of variation and this change in R^2 was significant ($F \text{ change}(2, 148) = 126.32, p < .05$). Finally, adding NA to the regression model explained less than 1% additional variance in promotion and this change in R^2 was not significant. Together the three independent variables accounted for 63% of the variance in promotion, with the majority of the variance explained by prevention.

The second hierarchical multiple regression revealed that at stage one, promotion contributed significantly to the regression model, ($F(1, 149) = 197.07, p < .05$) and accounted for 57% of the variation in promotion. Introducing PA in stage two explained less than 1% of variation in prevention and this change in R^2 was not significant. Finally, adding NA to the regression model also explained less than 1% additional variance in prevention and this change in R^2 was not significant. Together the three independent variables accounted for 58% of the variance in prevention, with virtually all of the variance explained by promotion.

Taken together these findings largely support previous findings (Haws, 2010; Summerville & Roese, 2008) that PA and NA are unrelated to promotion and prevention regulatory focus. Despite the fact that PA predicted promotion regulatory focus in the first regression analysis, the effect size ($R^2 = .06$) was so small that Cohen (1992), would consider it meaningless.

Chapter 4 - Discussion

Regulatory focus has emerged as an important individual difference variable in organizational research (Lanaj, Chang, & Johnson, 2012). However, despite its ubiquity of use, the psychometric properties of existing measures of regulatory focus have received relatively sparse research attention. Haws (2010) and Summerville (2008) evaluated several existing measures of regulatory focus and concluded that the various measurement approaches to regulatory focus orientation are very different theoretically and empirically, and have very little overlap. In fact a key finding from Summerville (2008) is that the prevention subscales of the RFQ and GRFM (the two most widely used measures of regulatory focus) are actually negatively related to one another. Clearly, these two scales are not measuring the same latent factors. Thus, the overall purpose of the current research was to develop a new measure of regulatory focus (Higgins, 1997, 1998) that was more grounded in extant regulatory focus theory and that followed current scale development best practices (Hinkin, 1995).

A key assumption of regulatory focus theory is that regulatory focus consists of two sub-dimensions: a promotion regulatory focus and a prevention regulatory focus. However, to the best of this author's knowledge, no direct empirical test of the dimensionality of regulatory focus has been conducted. The seminal regulatory focus papers (Higgins, 1997, 1998) discuss promotion focus as consisting of self-regulation around nurturance needs and ideal goals, and prevention focus as self-regulation around security needs and ought goals, but offer no empirical evidence that nurturance needs and ideal goals or security needs and ought goals represent single factors. It is conceivable that promotion focus (nurturance orientation and ideal goals) and prevention focus (security orientation and ought goals) are more accurately represented as composite constructs i.e., that promotion is comprised of two facets (concern with nurturance

and ideal goals) and that prevention is also comprised of two facets (concern with security and ought goals).

The current research sought to remedy this gap in the literature by developing a new measure of regulatory focus consisting of items written to measure all aspects of the theory and to test the dimensionality of regulatory focus i.e., whether a two-factor or a four-factor model fit the data best. Scale development best practices (see Crocker & Algina, 1986; Hinkin, 1995) were followed and 59 new items were written to measure all aspects of regulatory focus (Higgins, 1997, 1998). After pruning poor performing items, a final scale consisting of 14 items emerged. CFAs were used to test whether a two-factor or a four-factor model of regulatory focus fit the data best. Results suggested that both models fit the data equally well. However, given that the two-factor model is more parsimonious and given that one of the latent factors of the four-factor model contained only two items (making any estimates of internal consistency difficult) the two factor model of regulatory focus was retained as the preferred model.

The current research makes several important theoretical contributions to the existing regulatory focus literature. First, researchers heretofore have merely assumed and accepted that promotion and prevention regulatory focus are one-dimensional factors comprised of nurturance concerns and the pursuit of goals as ideals and security concerns and the pursuit of goals as obligations, respectively. Results from the current research provide the first empirical support that items written to measure nurturance concerns and ideal goal orientation, and items written to measure security concerns and ought goal orientation do, in fact, group together to form two regulatory focus sub-dimensions.

Second, the newly developed measure of regulatory focus (the PPS) represents a substantial improvement over existing measures both in terms of reliability, and in terms of

content validity. The RFQ, for example, has reliability coefficients in the .70 range and is missing items to measure affective components of regulatory focus. The reliability of the PPS is better, with promotion subscale alpha coefficient at .91, and the prevention subscale alpha coefficient at .75. The PPS also includes items to measure affective aspects of regulatory focus.

A third theoretical contribution involves the relation of regulatory focus to positive and negative affect (PA and NA). Regulatory focus theory posits and empirical studies (Haws, 2010; Summerville & Roese, 2008) have substantiated that regulatory focus is unrelated to both PA and NA. A shortcoming of those studies, however, is that they used older measures of regulatory focus (RFQ, GRFM) that lack items written to tap into the affective components of regulatory focus. The current research is the first to demonstrate empirically that a regulatory focus measure containing affective items can maintain its distinct relation to PA and NA.

A final theoretical consideration pertains to the orthogonal nature of promotion and prevention focus. Prior research (Haws, 2010; Higgins, 1997, 1998; Johnson et al., 2015) suggests that promotion and prevention are orthogonal sub-dimensions of regulatory focus. In contrast, findings from the current research suggest that promotion and prevention are actually strongly correlated ($r = .90$) to one another. This may suggest that promotion and prevention as operationalized in the PPS are really measuring the same underlying latent factor. Or it may indicate that dimensions of regulatory focus somehow function simultaneously, and that conceptualizing individuals as functioning with *either* a promotion focus *or* a prevention focus does not accurately represent how human beings function in real life. Future empirical and theory development work is needed to reconcile these incongruent findings.

A new more reliable and (hopefully) content valid measure of regulatory focus also has practical implications. First, given the improved reliability associated with the PPS, researchers

can be more confident in the precision with which they are measuring regulatory focus. Given the ubiquity of regulatory focus research of late, that is a substantial benefit. Potentially harmful consequences could come from a measure of regulatory focus (like the RFQ) where close to 30 percent of score variability is due to random error.

Second, given a more psychometrically sound measure, human resource practitioners may feel more comfortable including regulatory focus as a performance predictor in certain jobs (e.g., jobs that require attention to safety may benefit from a prevention oriented employee). If regulatory focus measures predict job performance then it would make sense to consider incorporating a measure of regulatory focus into corporate selection models.

Finally, if the current empirical results are replicated in future research, then it may suggest that regulatory focus is better conceptualized as consisting of a hybrid between promotion and prevention rather than an “either / or” perspective. Existing regulatory focus research (e.g., Halverson & Higgins, 2013) acknowledges that individuals are governed by both foci, but suggest that they operate independently of one another. A good test of this assumption would be to compare the model fit of a one-factor model of regulatory focus to the two and four factor models. If a one-factor model fits well, then that may indicate that more theoretical development of regulatory focus is needed to conceptualize regulatory focus as a unidimensional construct. If the two factor model continues to fit the data best, then consensus grows regarding the dimensionality of regulatory focus.

The current research must also be considered in light of several limitations. First, given the high correlation between the PPS promotion and prevention subscales, it is reasonable to question whether the new measure is really discriminating between individuals with different regulatory focus orientations. Perhaps a new measurement approach could be used in which

respondents are given scenarios and then asked to respond to items. For example, a scenario might involve purchasing a new car. Regulatory focus theory suggests that prevention oriented individuals would be more likely to purchase a car based on safety and reliability features, whereas a promotion person would be more impressed with features like 0-60 speed, or fine leather upholstery. Perhaps a measure that included scenarios like these followed by behaviorally anchored questions, could more precisely discriminate between promotion and prevention individuals.

A second limitation concerns the cross-sectional nature of the current research. Given that study participants responded to items at one discrete point in time, estimates of test-retest reliability were not possible. The PPS is designed to measure trait regulatory focus, which suggests that individuals have a relatively stable dispositional tendency towards either promotion or prevention regulatory focus. Administering the PPS to the same group of participants over varying time intervals, and obtaining consistent results, would be evidence that would support the stability of the regulatory focus over time. Conversely, if test-retest reliability coefficients were consistently low, then it might suggest that regulatory focus functions more like a state than a trait. If evidence supported the latter conclusion then regulatory focus theory would need to be reconsidered almost completely.

Third, following suggestions of other researchers (e.g., Haws, 2010) potential interaction effects associated with regulatory focus should be explored. For example, do (and if so how) promotion and prevention regulatory focus interact with each other in certain circumstances. It is possible that someone with a strong promotion orientation may languish when performing tasks that would require a strong prevention focus to do well (e.g., a tax auditor, or editor of a newspaper) and vice versa. Further, trait-like regulatory focus should be investigated in the

context of state-like regulatory focus. The current research focused exclusively on the measurement of trait regulatory focus, however, a small amount of prior research (e.g., Shah, Higgins, & Friedman, 1998) suggests that context can actually activate a promotion or prevention focus within an individual. If this is true, then future research should explore the interaction effects of trait and state regulatory focus.

Given this newly established and more reliable measure, researchers should also continue working to establish the overall construct validity of regulatory focus. Attention should be paid to various criterion related validity as well as to the theoretical content of regulatory focus as empirical evidence continues to either support or challenge existing theory. An important part of establishing construct validity is developing the construct's nomological network, which consists of established empirical relationships with other like and unlike constructs. For example, given that regulatory focus is a motivational construct, there should be at least some overlap between regulatory focus and other motivational constructs like work engagement, work commitment, or even flow. One might also expect that regulatory focus would be related to job satisfaction. In contrast, regulatory focus should be negatively related (or not related at all) to constructs like general intelligence or turnover intentions. A step towards developing the nomological network of regulatory focus was made in the current research by showing that regulatory focus is not related to PA or NA in any meaningful way; however the network of construct relations should continue to be built if the construct validity of regulatory focus is to be well established.

Other more general challenges to RFT also remain unexplored. For example, given RFT's reliance on caretaker-child interactions as the formative process driving the development of promotion and prevention focus, an obvious question is how does RFT relate to Bowlby's Attachment Theory (Bowlby, 1969, 1973)? This question has not been addressed in the extant

literature. Longitudinal studies should be designed to track infant/parent interactions to determine if they do in fact impact the development of regulatory focus later in life. Currently it is unclear what happens to one's regulatory focus development if, for example, one parent uses a more authoritarian parenting style (which theoretically should relate to a prevention focus) while the other parent uses a more authoritative style (which should relate to a promotion focus). Or what happens if parents inconsistently mix authoritative and authoritarian parenting styles? Does that lead to a more balanced promotion/prevention orientation, or is one style more influential than the other?

Additionally, the seminal regulatory focus papers (Higgins, 1997, 1998) cite studies conducted out of Higgins's lab at Columbia University to substantiate many of the predictions of regulatory focus; unfortunately, most (perhaps all) of those older studies use an operationalization of regulatory focus based on the Selves Questionnaire (mentioned above under measurement approaches). As may be recalled from discussion above, the Selves Questionnaire is a highly unreliable and subjective measure, and any conclusions drawn from its use should be made cautiously. A program of research is needed to retest some of the main predictions of regulatory focus theory using reliable and valid measures of regulatory focus.

Finally, RFT seems to suggest that the entirety of human goal-striving can be subsumed under promotion or prevention regulatory focus. But are there other fundamental human needs that operate outside of the parameters of regulatory focus? Are nurturance and security needs the only human survival mechanisms driving behavior? What about goal setting? An interesting program of research would be to integrate regulatory focus theory with goal setting theory (Locke & Latham, 2006). Both models of motivation have goals at their center, but curiously, the work of Latham and Locke and the work of Higgins make no mention of one another. One

possible reason for this may be seen in a scathing critique of control theories (Carver & Scheier, 1982) (regulatory focus theory is considered a control theory by most accounts) by Locke back in the early 90's (Locke, 1991). Essentially his critique of control theories is that they misrepresent human behavior in response to goals. It is true, as control theories suggest, that humans have goals and seek to reduce discrepancies between their goal and their current state. However, as Locke points out, humans are also deficit *creating* by nature too in that they actively and routinely create goals for themselves that they then try to achieve. Locke & Latham (2004) suggest a model of (work) motivation that includes a variety of characteristics all thought to be salient to motivating behavior. Perhaps additional work should be conducted to include RFT into this model.

All of these kinds of questions should be explored and having a quality measurement tool available is a great and necessary first step.

References

- Anastasi, A. (1976). *Psychological Testing* (4th ed.). Oxford, England: Macmillian.
- Arbuckle, J. L. (2009). AMOS (Version 18). Chicago: IBM SPSS.
- Bowlby, J. (1969). *Attachment and loss: Vol 1. Attachment*. New York, NY: Basic Books.
- Bowlby, J. (1973). *Attachment and loss: Vol. 2. Separation anxiety and anger*. New York, NY: Basic Books.
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and Validity Assessment*. New York, NY: Sage Publications.
- Carver, C., & Scheier, M. (1982). Control theory: A useful conceptual framework for personality–social, clinical, and health psychology. *Psychological Bulletin*, 92(1), 111–135.
- Cliff, N. (1993). What is and isn't measurement. *A Handbook for Data Analysis in the Behavioral Sciences, 1*, 59–93.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Comrey, A. L., & Lee, H. B. (1992). Interpretation and application of factor analytic results. In *A First Course In Factor Analysis*.
- Couper, M. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly*, 64, 464–494.
- Crawford, J. R., & Henry, J. D. (2004). The positive and negative affect schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *The British Journal of Clinical Psychology*, 43, 245–65.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory. Introduction to classical and modern test theory*. Fort Worth: Holt, Rinehart, and Winston, Inc.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302.
- Ettlie, J. E., & Reza, E. M. (1992). Organizational integration and process innovation. *Academy of Management Journal*, 35(4), 795–827.
- Furr, R. M., & Bacharach, V. R. (2008). Psychometrics and the importance of psychological measurement. In *Psychometrics: An introduction* (pp. 1–14). Thousand Oaks, CA: Sage Publications.
- Gorman, C. A., Meriac, J. P., Overstreet, B. L., Apodaca, S., McIntyre, A. L., Park, P., &

- Godbey, J. N. (2012). A meta-analysis of the regulatory focus nomological network: Work-related antecedents and consequences. *Journal of Vocational Behavior, 80*(1), 160–172.
- Greenhaus, J. H., Parasuraman, S., & Wormley, W. M. (1990). Effects of race on organizational experiences, job performance, evaluations, and career outcomes. *Academy of Management Journal, 33*, 64–86.
- Halverson, H. G., & Higgins, E. T. (2013). *Focus*. New York, NY: Penguin Group.
- Haws, K. L. (2010). An assessment of chronic regulatory focus measures. *Journal of Marketing Research, (October)*, 967–982.
- Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*(12), 1280–1300.
- Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. *Advances in Experimental Social Psychology, 30*, 1–46.
- Higgins, E. T., Friedman, R. S., Harlow, R. E., Idson, L. C., Ayduk, O. N., & Taylor, A. (2001). Achievement orientations from subjective histories of success: Promotion pride versus prevention pride. *European Journal of Social Psychology, 31*(1), 3–23.
- Higgins, E. T., Horwitz, E. K., Horwitz, M. B., Cope, J., Horwitz, E. K., Horwitz, M. B., & Cope, J. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*(3), 319–340.
- Higgins, E. T., Klein, R., & Strauman, T. (1985). Self-concept discrepancy theory: A psychological model for distinguishing among different aspects of depression and anxiety. *Social Cognition, 3*(1), 51.
- Highhouse, S., & Gillespie, J. (2009). Do samples really matter that much? In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and Methodological Myths and Urban Legends: Doctrine, Verity and Fable in the Organizational and Social Sciences* (pp. 249–268). New York: Routledge.
- Hinkin, T. (1995). A review of scale development practices in the study of organizations. *Journal of Management, 21*(5), 967–988.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.
- Idaszak, J. R., & Drasgow, F. (1987). A revision of the Job Diagnostic Survey: Elimination of a measurement artifact. *Journal of Applied Psychology, 72*, 69–74.

- Jackson, P. R., Wall, T. D., Martin, R., & Davids, K. (1993). New measures of job control, cognitive demand, and production responsibility. *Journal of Applied Psychology, 78*, 753–762.
- Johnson, P. D., Smith, M. B., Wallace, J. C., Hill, A. D., & Baron, R. D. (2015). A review of multilevel regulatory focus in organizations. *Journal of Management, 41*(5), 1501–1529.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 8–23.
- Lanaj, K., Chang, C. D., & Johnson, R. E. (2012). Regulatory focus and work-related outcomes: a review and meta-analysis. *Psychological Bulletin, 138*(5), 998–1034.
- Locke, E. A. (1991). Goal theory vs. control theory: Contrasting approaches to understanding work motivation. *Motivation and Emotion, 15*(1), 9–28.
- Locke, E. A., & Latham, G. P. (2004). What should we do about motivation theory? Six recommendations for the twenty-first century. *Academy of Management Review, 29*(3), 388–403.
- Locke, E. A., & Latham, G. P. (2006). New Directions in Goal-Setting Theory. *Current Directions in Psychological Science, 15*(5), 265–268.
- Lockwood, P., Jordan, C. H., & Kunda, Z. (2002). Motivation by positive or negative role models: Regulatory focus determines who will best inspire us. *Journal of Personality and Social Psychology, 83*(4), 854–864.
- Lord, F. M., Novick, M. R., & Birnbaum, A. (1968). *Statistical theories of mental test scores*. Oxford, England: Addison-Wesley.
- MacKenzie, S., Podsakoff, P., & Fetter, R. (1991). Organizational citizenship behavior and objective productivity as determinants of managerial evaluations of salespersons' performance. *Organizational Behavior and Human Decision Processes, 50*, 123–1991.
- Maul, A., Irribarra, D. T., & Wilson, M. (2016). On the philosophical foundations of psychological measurement. *Measurement, 79*, 311–320.
- Nunnally, J. (1994). Reliability. In *Psychometric Theory*. New York, NY: McGraw-Hill.
- Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington, MA: Lexington Books.
- Schmitt, N. W., & Stults, D. M. (1985). Factors defined by negatively keyed items: The results

- of careless respondents? *Applied Psychological Measurement*, 9, 367–373.
- Scholer, A. A., & Higgins, E. T. (2011). Promotion and prevention systems: Regulatory focus dynamics within self-regulatory hierarchies. In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of Self-Regulation* (2nd ed., pp. 143–161). New York, NY: The Guilford Press.
- Schriesheim, C. A., & Hill, K. (1981). Controlling acquiescence response bias by item reversal: The effect on questionnaire validity. *Educational and Psychological Measurement*, 41, 1101–1114.
- Schwab, D. P. (1980). Construct validity in organization behavior. In *Research in Organizational Behavior*, Vol. 2 (pp. 3–43). Greenwich, CT: JAI Press.
- Shah, J., & Higgins, E. T. (1997). Expectancy \times value effects: Regulatory focus as determinant of magnitude and direction. *Journal of Personality and Social Psychology*, 73(3), 447–458.
- Shah, J., Higgins, T., & Friedman, R. S. (1998). Performance incentives and means: How regulatory focus influences goal attainment. *Journal of Personality and Social Psychology*, 74(2), 285–293.
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103(2684), 677–680.
- Summerville, A., & Roese, N. J. (2008). Self-report measures of individual differences in regulatory focus: A cautionary note. *Journal of Research in Personality*, 42(1), 247–254.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics* (5th ed.). Boston, MA: Pearson.
- Torgerson, W. S. (1958). *Theory and methods of scaling*. Oxford, England: Wiley.
- Van-Dijk, D., & Kluger, A. N. (2004). Feedback Sign Effect on Motivation: Is it Moderated by Regulatory Focus? *Applied Psychology*, 53(1), 113–135.
- Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: how good are single-item measures? *Journal of Applied Psychology*, 82(2), 247–252.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070.

Appendix A - Promotion Prevention Scale (PPS) & RFQ

Promotion Prevention Scale (PPS)

The following questions pertain to your general approach to life. Please answer based on what is typically true for you. 1-Strongly Disagree to 5-Strongly Agree

1. When I succeed I feel a sense of joy...
2. I strive towards my ideals...
3. I think about how to achieve an optimal life...
4. I strive to be my ideal self...
5. I think about how I will achieve my hopes and aspirations...
6. Success to me is achieving my aspirations...
7. I view goals as opportunities not to be missed...
8. Whenever possible, I try to improve my current situation...
9. I strive to meet my obligations...
10. Success to me is meeting my obligations...
11. I view goals as duties I must complete...
12. When I succeed I feel a sense of calm...
13. I think about how I can prevent failures in life...
14. I am focused on preventing negative events in my life...

Regulatory Focus Questionnaire

Higgins, E. T., Friedman, R. S., Harlow, R. E., Idson, L. C., Ayduk, O. N., & Taylor, A. (2001). Achievement orientations from subjective histories of success: Promotion pride versus prevention pride. *European Journal of Social Psychology, 31*(1), 3–23.

Items 1, 3, 7, 9, 10, 11 – Promotion scale
 Items 2, 4, 5, 6, 8 – Prevention scale items

Table 1. Event reaction questionnaire

This set of questions asks you about specific events in your life. Please indicate your answer to each question by circling the appropriate number below it.

1. Compared to most people, are you typically unable to get what you want out of life? [−0.65]	1	2	3	4	5
never or seldom			sometimes		very often
2. Growing up, would you ever “cross the line” by doing things that your parents would not tolerate? [−0.80]	1	2	3	4	5
never or seldom			sometimes		very often
3. How often have you accomplished things that got you “psyched” to work even harder? [0.37]	1	2	3	4	5
never or seldom			a few times		many times
4. Did you get on your parents’ nerves often when you were growing up? [−0.65]	1	2	3	4	5
never or seldom			sometimes		very often
5. How often did you obey rules and regulations that were established by your parents? [0.56]	1	2	3	4	5
never or seldom			sometimes		always
6. Growing up, did you ever act in ways that your parents thought were objectionable? [−0.84]	1	2	3	4	5
never or seldom			sometimes		very often
7. Do you often do well at different things that you try? [0.54]	1	2	3	4	5
never or seldom			sometimes		very often
8. Not being careful enough has gotten me into trouble at times. [−0.55]	1	2	3	4	5
never or seldom			sometimes		very often
9. When it comes to achieving things that are important to me, I find that I don’t perform as well as I ideally would like to do. [−0.51]	1	2	3	4	5
never true			sometimes true		very often true
10. I feel like I have made progress toward being successful in my life. [0.81]	1	2	3	4	5
certainly false					certainly true
11. I have found very few hobbies or activities in my life that capture my interest or motivate me to put effort into them. [−0.53]	1	2	3	4	5
certainly false					certainly true

Appendix B - Initial PPS Items

Nurturance Items	Security Items	Ideal Items	Ought Items
I take risks in life	I like the routines of life	I pursue ideals in life	I pursue obligations in life
I do not like missing opportunities	I dislike losing more than I like winning	I view goals as opportunities not to be missed	I view goals as duties I must complete
I like new opportunities	I pay attention to details	When I succeed I feel a sense of joy	When I succeed I feel a sense of calm
I play to win	I usually follow the rules	When I fail I feel dejected	When I fail I feel anxiety
I am achievement oriented	I am more focused on avoiding mistakes than achieving success	I strive to be my ideal self	I am concerned that I will fall short of my duties
I hate missing out	I think about how I can prevent failures in my life	Advancing in life is important to me	I strive to be the person I am obligated to be
Whenever possible, I try to improve my current situation	I am focused on preventing negative events in my life	I savor winning	My responsibilities are important to me

Nurturance Items	Security Items	Ideal Items	Ought Items
I am a carefree person	I am a cautious person	I like to strike out on my own by setting my own goals	I am really hard on myself when I make mistakes
I think about how I will achieve my hopes and aspirations	Whenever possible, I try to maintain my current situation	I strive towards my ideals	I feel a strong sense of duty to those around me
I am more focused on achieving success than avoiding mistakes	I hate making mistakes	I tend to be a spontaneous person	I strive to meet my obligations
I usually don't follow the rules	I play not to lose	Success to me is achieving my aspirations	I see mistakes as betraying my duty
I am comfortable with change	I prefer to maintain the status-quo	I think about how to achieve an optimal life	I tend to be a responsible person
I like winning more than I dislike losing	I do not like change	I live a fast-paced life	Success to me is meeting my responsibilities
I always like life to be changing	I prefer not to take risks in life	I eagerly pursue goals	I cautiously pursue goals
I see mistakes as opportunities to grow			I live a slower-paced life
I pay			

Nurturance Items	Security Items	Ideal Items	Ought Items
attention to the big picture			