CAN HAPPINESS BE SUCCESSFULLY PURSUED? A RANDOMIZED
CONTROLLED TRIAL OF THE PURSUIT AND ASSESSMENT OF HAPPINESS

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by

Cody Daniel Christopherson

_____________________________
George S. Howard, Director

Graduate Program in Psychology
Notre Dame, Indiana
July 2011
CAN HAPPINESS BE SUCCESSFULLY PURSUED? A RANDOMIZED
CONTROLLED TRIAL OF THE PURSUIT AND ASSESSMENT OF HAPPINESS

Abstract

by

Cody Daniel Christopherson

The worst way to achieve some goals is by directly pursuing them. In the present study, the pursuit of happiness, or subjective well-being, is examined as a possible ironic process. This is done using the logic of a randomized controlled trial in a factorial design, N = 413. As a subset of this possibility, in separate experiment, the assessment of happiness is examined as a potential source of decreased happiness, N = 175. Results show that the pursuit of happiness is not self-defeating when effort is isolated as a variable. It is further found that the assessment of happiness is not self-defeating compared to control assessments. Both experiments were conducted in in-vivo settings and used email and online surveys. Implications and further ways of advancing the field with regard to effort and subjective well-being are discussed.
CONTENTS

TABLES .................................................................................................................. iii

FIGURES ................................................................................................................. iv

ACKNOWLEDGMENTS .............................................................................................. v

CHAPTER 1: INTRODUCTION ................................................................................... 1

CHAPTER 2: METHOD ............................................................................................... 40

CHAPTER 3: RESULTS .............................................................................................. 46

CHAPTER 4: DISCUSSION ......................................................................................... 73

APPENDIX A: NO EFFORT/CONTROL ACTIVITY
GROUP INSTRUCTIONS .............................................................................................. 85

APPENDIX B: WILLFUL EFFORT/CONTROL ACTIVITY
GROUP INSTRUCTIONS .............................................................................................. 86

APPENDIX C: NO WILLFUL EFFORT/THREE GOOD THINGS
GROUP INSTRUCTIONS .............................................................................................. 88

APPENDIX D: WILLFUL EFFORT/THREE GOOD THINGS
GROUP INSTRUCTIONS .............................................................................................. 90

APPENDIX E: THE WILLFUL EFFORT/NO ACTIVITY
GROUP INSTRUCTIONS .............................................................................................. 92

REFERENCES ........................................................................................................... 93
TABLES

Table 3.1: Descriptive Statistics of All Five Experimental Groups on All Three Measures of SWB at Post-Test. .................................................................48

Table 3.2: Satisfaction With Life Means and Marginal Means at Post-Test..................49

Table 3.3: Positive Affect Means and Marginal Means at Post-Test ..................50

Table 3.4: Negative Affect Means and Marginal Means at Post-Test..................50

Table 3.5: Mean Differences, Standard Errors, P Values, and Confidence Interval for Contrasts.................................................................55

Table 3.6: Mean Differences, Standard Errors, P Values, and Confidence Interval for Contrasts Testing Main Effect of Effort Across Levels of Activity ..........56

Table 3.7: Between and Within Subject Effects for Factors Time, Activity, Try and Interactions Time * Activity, Time * Try, and Time * Activity * Try..........58

Table 3.8: Between Subject Effects at One Week Follow-up ..................................59

Table 3.9: Between Subject Effects at One Month Follow-up .................................60

Table 3.10: Descriptive Statistics at Post-Test on Each Dependent Measure ..........62

Table 3.11: Descriptive Statistics of Between-Subject Effects at Follow up 1 ..........64

Table 3.12: Descriptive Statistics of Between-Subject Effects at Follow up 2 ..........66

Table 3.13: Descriptive Statistics of Each Group on Each Measure at Time Point ........68

Table 3.14: Pairwise Comparisons Between Experimental and Control Groups for Each Measure and at Each Time Point with Bonferroni Adjusted CI ....69
FIGURES

Figure 3.1: Positive affect at post-test .................................................................52
Figure 3.2: Satisfaction with life at post-test ....................................................53
Figure 3.3: Negative affect at post-test ..............................................................54
Figure 3.4: Post-test and follow-up scores on negative affect ............................63
Figure 3.5: Positive affect in project 2, happiness measures compared to control measures .................................................................70
Figure 3.6: Negative affect in project 2, happiness measures compared to control measures .................................................................71
Figure 3.7: Satisfaction With Life in project 2, happiness measures compared to control measures .................................................................72
ACKNOWLEDGMENTS

This dissertation represents the culmination of my formal education. There are many people who have led, nudged, directed, and supported me on this journey. Though I cannot name them all here, I owe each a debt of gratitude. First, thank you to my advisor, George Howard. I sought mostly autonomy when I first began my search for a graduate advisor. I found a model of autonomy in George and to my surprise I also found a kind and decent friend.

Many faculty at Notre Dame have left their impression on me. I wish to thank my dissertation committee in particular. Anita Kelly’s provocative questions provided the spark that led to this dissertation in particular. She exemplifies the excitement and vibrancy that can be a part of the life of a scholar. Scott Maxwell possesses the rare dual gift of knowledge and the kindness that comes with remembering what it is like not to know. His nuanced and sensitive approach to statistical analysis provides a high benchmark for me and for all social scientists. Megan Brown helped me to appreciate the perspective of those researchers who had gone before.

Many fellow students have been cheerleaders and commiserators. I am equally grateful for both roles. Michael Lau took me under his wing from my first moments on the Notre Dame campus and has been a good example, worthy interlocutor, and dependable confidante. Steven Tueller was always willing to give brief but frequent statistical consultations over the last four years. Families Degruccio, Mason, and Andrew,
as well as Clint Elison collectively made South Bend a welcoming and lovely place for our family.

The anonymous staff at the library who scan and send articles and our own department staff of Rhonda Singleton, Judy Spiro, and Judy Stewart were each crucial. I appreciate the financial support from Notre Dame which made full time scholarship possible.

Three professors in my undergraduate years made particularly significant nudges that led to my current path. Brent Slife remains a shining example of courageous scholarship. I continue to wrestle with ideas he taught me over several courses, meetings, and personal conversations. Michael Lambert generously provided an early entry in to the world of empirical research. Though quite different in their approaches, each of these men showed devotion to their clients and in balancing research and practice. Claudia Clayton spoke words of encouragement and provided teaching opportunities at crucial developmental moments. Thank you too to five early cultivators of young minds: Mary Ellen Piscopo, Betty Miller, Frank White, Karen Del Vecchio, and Joe Jackson.

My parents and extended family have been nothing but encouraging in all my endeavors. My mom and her husband, Christie and Rob Berg, my dad, Harold, and my brother, Jesse, have each made a point of expressing their love and pride. My father-in-law, Gordon Rands, has been a helpful sounding board and unofficial advisor in all things academic.

My single greatest support has been my wife, Jennie. We moved to South Bend with just one another and we leave with each other, but also with Juniper, Olive, and Henry. I am thankful that God’s hand has blessed us both to thrive and to struggle. I am
thankful to her for growing with me and for pulling it all off, many times from behind the scenes. This dissertation is dedicated to her.
CHAPTER 1:
INTRODUCTION

The worst way to achieve some goals is by directly pursuing them. Most famously, we have learned that not thinking of white bears is nearly impossible under certain conditions—namely, when we are prompted to so do (Wegner, Schneider, Carter, & White, 1987). The phenomenon of a task becoming more difficult as it is pursued more rigorously is known as the ironic process of mental control (Wegner, 1994). Other examples include relaxation (Heide & Borkovec, 1983), trying to fall asleep (Ascher & Efran, 1978), concentration (Wegner & Erber, 1992), and pain control (Cioffi & Holloway, 1993).

Is the pursuit of happiness one such ironic process? There is theoretical and philosophical evidence which suggests that the pursuit of happiness is self defeating (see for example Aristotle, 1994; Frankl, 1963; Kierkegaard, 1997; Mill, 1873). More recently, there has empirical scientific evidence from laboratory, field, and in vivo experiments indicating that attempting to increase happiness can result in decreased happiness (Schooler, Ariely, & Loewenstein, 2003; Wegner, Erber, & Zanakos, 1993). Researchers have long been aware of barriers to increasing happiness. These barriers include the high degree of heritability in genetic studies of happiness (Lykken & Tellegen, 1996), the association of high levels of positive affect with stable personality traits such as high extroversion and low neuroticism (Diener & Lucas, 1999), and the
relationship between happiness and slowly changing life circumstances such as income (Diener, Lucas, & Scollon, 2006), marriage, and religiosity (Diener, 1984).

Despite these ironic effects and barriers, there appear to be specific interventions that have been shown to increase happiness in real life settings. For example, some claim that happiness can be increased by expressing gratitude (Seligman, Steen, Park, & Peterson, 2005), counting blessings (Emmons & McCullough, 2003), setting goals (MacLeod, Coates, & Hetherton, 2008, Sheldon et al., 2010), positive self-visualization (Sheldon & Lyubomirsky, 2006), kind acts (Lyubomirsky, Sheldon, & Schkade, 2005), reflecting on positive experiences (Lyubomirsky, Sousa, & Dickerhoof, 2006), identifying your personal strengths (Seligman et al., 2005), and mindfulness (Zautra et al., 2008). There is some indirect evidence that effort itself might be important to successfully increasing happiness (Lyubomirsky, Dickerhoof, Boehm, & Sheldon, 2011).

Thus, there are two opposing viewpoints with regard to effort and happiness. In achieving happiness, effort is either self-defeating on one hand, or necessary on the other. The main purpose of this introduction is to analyze these viewpoints as thoroughly as possible given current empirical research. I will do so in four parts. First, I will begin with a very brief introduction to happiness and subjective well-being studies. Second, I will review theoretical and scientific findings supporting the idea that the pursuit of happiness is difficult, impossible, or even self-defeating. I will expand on a line of reasoning which attempts to demonstrate that hedonic appraisal is different than hedonic experience, appraisal is detrimental to experience, and that pursuing happiness can be self-defeating. This general reasoning was originally presented by Schooler et al.
I will expand on it by updating it and including additional supporting studies for the main arguments. Third, I will review the literature that has attempted to show the possibility of achievable increases in happiness. I will review both heavily flawed and well-designed studies in order to show the state of the field and for an attempt at completeness. Overall, there is weak evidence for the possibility of improving happiness but there are some promising individual interventions. Finally, I will describe the way in which the present study fills the gap in knowledge with regard to the pursuit of happiness.

Happiness and Subjective Well-Being

Humanity has been wrestling with questions about happiness for at least 2500 years. Aristotle wrote and debated prodigiously on the nature of eudaimonia—the good life. He is not alone. There is perhaps no topic for which more ink has been spilled than human happiness. However, a scientific approach to understanding human flourishing, strengths, and virtues has emerged only in the last century. The current incarnation of this movement, positive psychology, can be said to have its roots in William James, the humanist movement of the 1950s and Maslow in particular (Froh, 2004), and possibly even the mind cures of the 19th century and the mental hygiene movement of the early 20th century (Becker & Marecek, 2008). “Positive psychology” was launched in earnest by Martin Seligman during his tenure as APA president at the end of the 1990s. Positive psychology is meant to be “a science of positive subjective experience, positive individual traits, and positive institutions [which] promises to improve quality of life and prevent the pathologies that arise when life is barren and meaningless” (Seligman & Csikszentmihalyi, 2000, p. 5). Although leading subjective well-being researcher Ed
Diener started work on it in 1981 before the term “positive psychology” was in use, the study of subjective well-being is now considered to be an important part of the positive psychology movement.

Subjective well-being (SWB) is a technical term for the colloquial term “happiness.” In this paper, the two terms will be used interchangeably. Subjective well-being relies on the subjective experience of the individual and consists of three separable components: high positive affect (PA), low negative affect (NA), and satisfaction with life (SWL) (Diener, 2000). Subjective well-being studies typically have focused on basic research like defining and measuring SWB, studying the demographics and behaviors of the happy, and exploring how and why we adapt to changes in SWB. In the last ten years, studies have begun to emerge which test methods to systematically increase SWB. This type of research has been slow to develop, possibly because of the well-known barriers to increasing levels of SWB.

Pursuit of Happiness: Bad Idea?

Evidence that we cannot consciously improve levels of subjective well-being is somewhat indirect, given the difficult nature of “proving a negative,” or finding enough test cases to adequately demonstrate the lack of a relationship (in this case, between conscious awareness/effort and increased SWB). Therefore, because it is impossible to gather an infinite number of test cases, evidence demonstrating this effect is reliant on inference. In this section I will review four different types of evidence: first, explanations for variations in happiness that clearly do not involve conscious effort, namely, genetics; second, difficulties in measuring happiness; third, evidence that
assessment of hedonic experience negatively influences the hedonic experience; and finally, some limited direct experimental evidence that the pursuit of happiness can be self-defeating.

*Genetic studies of happiness.* If differences in happiness can be accounted for by variables over which one has no control, there is reason to believe that we cannot consciously improve subjective well-being. Genetic studies of happiness are one major source of evidence that differences in happiness among individuals are not the result of conscious effort. In their classic twin study, using the Minnesota Twin sample of identical twins who had been raised apart, Lykken and Tellegen (1996) reported that between 44 and 52% of the variance in happiness in their sample was based on genetic variation (compared to less than 3% of the variation in happiness explained by socioeconomic status, marriage status, education, income, and religious commitment). Furthermore, they estimated that the stable component of well-being approached 80%. This led to their famous quip that trying to be happier may be as futile and counterproductive as trying to be taller. Though this study has not definitively shown that happiness is non-effortful, it raises serious doubt as to how much of our happiness we are able to control.

*Do we know how happy we are? Difficulties in measuring happiness.* Problems with measuring happiness are important for two reasons: first, studies of interventions heavily rely on such measures. If these measures are not valid, the entire body of happiness interventions is compromised; second, understanding the difference between hedonic experience and assessment of that experience highlights one possible mechanism for understanding why the pursuit of happiness is self-defeating: pursuing happiness may
lead to the too-frequent assessment of happiness, which causes decreases in happiness.

To flesh out these reasons, in this section I will examine the relationship between happiness and assessment of happiness.

Subjective well-being researchers typically begin their studies with the explicit assumption that face valid self-report measures of happiness are valid measures of happiness. The justification for this assumption is given in two parts: First, in order for the study of “happiness” to be meaningful, we must assume that happy and sad individuals are conscious of their emotional state (Diener, 2000; Lyubomirsky, 2008). If conscious awareness of emotion is completely separate from the emotional state itself, study of the emotional state is impossible. Second, self-reports of happiness have been shown to be relatively valid and reliable (Kahneman & Krueger, 2006). One meta-analysis of SWB measures showed that SWB was somewhat less reliable than many other microeconomic variables but had adequate reliability to support research based on comparing group means (Krueger & Schkade, 2007). Self-reports of overall, global happiness tend to correlate well with assessments made by friends and spouses ($r = .41 - .66$) (Lyubomirsky & Lepper, 1999).

Why, then, would one doubt that humans can assess their own happiness? There is evidence that many factors bias the assessment of global subjective well-being. For example, the way researchers ask participants to evaluate their happiness can have a profound impact on global SWB reports. For example, one way to introduce bias is through order effects (Strack, Martin, & Schwarz, 1988). One way the impact of order effects has been demonstrated is by comparing two groups of college students asked an identical set of two questions. The first group of college students was asked first how
satisfied they were with their lives overall and then how many dates they had been on recently. The correlation was -0.16. Startlingly, when the next group was first asked how many dates they had been on recently and then asked about overall SWL, the correlation increased to 0.66. Other contextual variables that have been shown to have an impact on global happiness self-reports include current weather (Schwarz & Clore, 1983), finding a dime (Schwarz, 1987 as cited in Kahneman & Krueger, 2006), and forcing people to smile (Strack, Martin, & Stepper, 1988). The results of these simple manipulations suggest that instead of actually assessing their own emotional/hedonic state, participants look for external evidence or “objective cues” about whether they are feeling happy or sad.

In response to the “objective cue” criticism, SWB researchers claim that people are subject to many biases when summing up their experiences over time, but can provide accurate feedback about their immediate emotional state. The researcher can then average the reports to gain an accurate picture of their overall level of well-being (Kahneman, 1999). In this way, SWB theorists distinguish between visceral experience (experiential consciousness) and reflective appraisal (meta-consciousness). They suggest that we have continuous hedonic experience but only intermittent awareness of said experience (intermittent reflective appraisal/meta-awareness/meta-consciousness). That is to say, we always are feeling some degree of pleasure or displeasure—we can call this our hedonic experience. Our hedonic experience is registered as subjective visceral feelings which are further revealed by cardiovascular flow, eye blink response, and brain activation (Blascovich & Tamaka, 1996). Even though these things are continuous through every waking moment, we do not have the resources to continuously monitor
ourselves. It is possible that someone can feel happy or sad and not realize it until a spouse, coworker, or therapist points it out (Schooler, et al., 2003). Research on flow shows that we typically lack meta-consciousness of our subjective emotional experience (Csikszentmihalyi, 1999).

The idea that we cannot directly assess our emotions but must infer them is as old as psychology itself. William James first suggested that we infer our emotions by observing our actions. In his classic example, a man runs from a bear and only later infers that he is feeling fear. He does not feel fear until he is aware of it (James, 1884). An alternative explanation is that he feels fear as he is running from a bear (continuous hedonic experience), but does not explicitly appraise the feeling as fear until he observes his own behavior. He feels fear before he is aware of it.

This alternative theory is Bem's self-perception theory, summed up by this quote: “Individuals come to 'know' their own attitudes, emotions, and other internal states partially by inferring them from observation of their own overt behavior and circumstances in which this behavior occurs” (Bem, 1972, p. 2). Self-perception theory is important to this discussion because experimental tests of self-perception theory shows that people can and often do lack awareness of their own internal states (See Laird, 2007 for a review).

Another area of research that demonstrates lack of continuous hedonic awareness is misattribution theory. Similar to self-attribution theory, this theory shares the assumption that people have an underlying subjective experience that is subject to interpretation and, therefore, error. For example, studies have shown that when given misleading external cues, subjects misattribute the source of their arousal. Two
demonstrated effects are the mistaking of exercise arousal for anger and fear for sexual arousal (Schwarz, Survay, & Kumpf, 1985). This supports the notion that explicit appraisals of hedonic state are made using inferences, and not by directly accessing one's ongoing underlying experience.

One important assumption in the social sciences is that humans tend to pursue pleasure and avoid pain (Slife & Williams, 1995). Amazingly, we might have difficulty even knowing whether an experience is positive or negative without relying on external cues such as anchoring effects. In one informal study, a professor used two different classes as the participants in two experimental conditions. He asked one class whether they would pay $1 to listen to him read a prose piece entitled *Leaves of Grass* for three minutes and one class whether they would listen to three minutes of *Leaves of Grass* for $1. Next, he asked both classes how much they would pay or be paid to listen to ten minutes of *Leaves of Grass*. On average, the group to which he offered money demanded more to listen to ten minutes of the piece, and the group from whom he requested money were willing to pay more to listen to ten minutes (Ariely, Loewenstein, & Prelec, 2006). This effect has been replicated using a slightly different methodology. Participants listened to a single tone for 30 seconds and then were randomly asked one of two questions: Would you listen to this tone for 30 seconds for ten cents? Or, Would you listen to this tone for 30 seconds for ninety cents? All participants then entered a real, online negotiation with a computer. Those who were asked if they would listen for ten cents ended up getting less money than those who were asked whether they would listen for ninety cents (Ariely et al., 2000). Thus it is apparent that we rely on outside cues even to determine whether a particular experience is enjoyable or burdensome.
The main contribution of the above research is that it has shown that the ongoing hedonic experience and the explicit evaluation of that experience are two separate phenomena. This assertion raises two additional questions: does assessing happiness have an effect on happiness? And is this effect reliably negative? I will show evidence that the answer to both questions is yes.

*Effects of hedonic introspection.* Even if people are capable of knowing how happy they are, the act of finding out may have a polluting and negative effect on their happiness. John Stuart Mill (1873) wrote “Ask yourself whether you are happy, and you will cease to be so.” Over one hundred years later, this pithy quote has found empirical support.

There is evidence that reflection and introspection dulls sensual acuity. Moment-to-moment evaluation can dull sensitivity to changes in underlying hedonic experience (Wilson & Schooler, 1991). For example, when subjects were asked to reflect on the amount of pain they felt when shocked, they became less sensitive to difference in magnitude amongst the shocks (Ariely, 1998). This effect has been demonstrated in positive hedonic experience as well. In just one example (Wilson et al., 1993), participants were shown a variety of posters. Each participant in the control group first ranked the posters and then selected a poster to take home. Each participant in the introspective (experimental) group first analyzed how he or she felt about the posters and why, and then ranked them and took home his or her top choice. Two weeks later, each participant was asked two questions: “How satisfied are you with your selection?” and “Do you have your poster hung up?” Control participants were both more satisfied and more likely to have the poster displayed. These studies support the theory that
introspection can dull sensitivity to underlying hedonic experience and that when we introspect, we make worse hedonic choices.

Engaging in appraisal could also impact the hedonic experience itself. There is correlational evidence that chronically happy people are less introspective (Lyubomirsky & Lepper, 1999). Unhappy people are more self-focused and ruminative (Ingram, 1990). Self-awareness can cause bad experiences because people come to recognize the gap between their aspirations and their attainments (Duval & Wicklund, 1972).

Noting the effects of introspection, some researchers have developed techniques to overcome the dulling impact of explicit appraisal. For example, forcing gut-level, fast assessments helps people tap into their hedonic experience (Wilson et al., 2000). Researchers can also help increase the accuracy of hedonic appraisals by providing a mirror in the room where participants are completing hedonic evaluations (Schier & Carver, 1977).

Assessment of happiness is necessary for increasing happiness, and it is problematic in many ways. Beyond mere assessment, what happens when people actually try to increase their happiness? I will now present evidence that pursuing happiness is a difficult, impossible, or even self-defeating process.

Impact of the Pursuit of Happiness on the Attainment of Happiness

Although most people are happy, most people would like to be happier (Diener & Diener, 1996). However, becoming happier is no easy task. There is scientific evidence that we fail in happiness pursuits in predictable ways. The main problem is that we are subject to innate cognitive errors, or biases, which impact our ability to make a good
prediction about the hedonic effects of a decision. These biases lead to faulty personal theories of happiness. The biases in these theories include impact bias, projection bias, distinction bias, memory bias, and belief bias (see Gilbert, 2006 and Hsee & Hastie, 2006 for reviews and further elaboration on these biases).

One idea, the *hedonic treadmill* theory, has been particularly fruitful in helping researchers understand why our faulty personal theories fail. When making plans to be happier, we either underestimate or are unaware of the hedonic treadmill (Brickman & Campbell, 1971). That is to say, we habituate to both good and bad changes rather rapidly and return to a hedonic set-point soon after a good or bad special event in our lives. Because we do not intuitively understand the hedonic treadmill, when trying to increase our long-term happiness, we consistently set goals that increase our short-term happiness temporarily before we tumble back to our baseline (Gilbert et al., 1998). One example of a goal that is often thought to increase long-term happiness but that only has short term effects is an increase in material gains. (Evidence for the failure of the pursuit of money as strategy for happiness is further supported by research showing that people who wish to accumulate wealth are less happy (Kasser & Ryan, 1993)). It is possible that the widespread awareness of the hedonic treadmill theory among happiness researchers has prevented rigorous studies to increase happiness (Diener, Lucas, & Scollon, 2006).

The hedonic treadmill/adaptation theory of happiness helps explain many of the mysteries of subjective well-being research. In fact, all demographic variables combined including age, race, sex, religion, and geographic location account for only 10-20% of the variance in happiness (Argyle, 1999; Campbell, Converse, & Rodgers, 1976). One
might then expect that the well-educated, rich, or beautiful people among us are the happiest. On the contrary, income in the U.S. (Diener, Sandvik, Seidlitz, & Diener, 1993), objective physical attractiveness (Diener, Wolsic, & Fujita, 1995), and objective measures of health (Okun & George, 1984) all have been found to have very low correlations with happiness. These weak relationships make sense if we all are only experiencing happiness in brief spurts before returning to baseline. The hedonic treadmill is, therefore, one more theory to explain why successfully pursuing happiness is a non-intuitive and perhaps difficult or impossible task.

*Theoretical evidence of the harm in pursuing SWB.* Aside from being merely fruitless, there is theoretical evidence that the pursuit of happiness actually can be self-defeating. This is particularly true when people see goals as means to an end (happiness) instead of valuable activities in and of themselves. Performing a task or completing a plan for an external reward takes away from the intrinsic benefit of the activity (Deci et al., 1999). An example from Deci et al. is that rewarding school children with money or candy for reading can make reading for its own sake difficult to enjoy. In a similar way, participating in activities such as service to other people in order to make oneself feel happier may take away from the intrinsic happiness that comes from acting to the end of helping another person. The participant’s goal changes from helping another person to making the participant (the self) happier. By so doing, one robs oneself of the rewards for achieving goal A in order to gain the rewards of goal B. Further supporting this point, people who linked their lower order goals (“get my grades up”) with higher order goals (“...so I will be happier”) are more prone to rumination and depression than those who make no such link between higher and lower order goals (McIntosh, Harlow, & Martin,
Selfishness is in fact negatively correlated with happiness. There is survey research which supports this point. People rated others they knew on various character dimensions such as happiness and selfishness. Those who were rated as more selfish also were rated as less happy (Myers, 2000).

Thus, faulty theories regarding adaptation to short-term happiness can lead a person down the wrong path, but pursuing goals for the sake of happiness itself can rob the goal of its intrinsic value. Though the evidence cited thus far is theoretical and correlational, there is also experimental data to support this assertion.

Experimental evidence. Though the topic is under-developed, there is some limited experimental evidence that the pursuit and assessment of happiness can be self-defeating. This evidence falls in to three categories: experimental studies attempting to demonstrate the efficacy of happiness interventions but which backfire and demonstrate a decrease in happiness, laboratory experiments focused on examining the negative impact of pursuing happiness, and field experiments looking at real life instances of the ironic effects pursuing happiness.

The first category, failed attempts to increase happiness, is the weakest evidence of the three categories. Interpreting a null result is notoriously problematic. There are many possible causes of a null result, including insufficient power and invalid or unreliable measures. (See Cohen (1994) for a thorough discussion of the problems in interpreting significance tests in general and null results in particular.) However, given the relatively small body of studies of interventions to increase happiness, it may be a fruitful exercise to look at null results obtained (or results in the unexpected direction)
and attempt to understand which interventions appear to have failed. I explore this evidence more thoroughly in a review of the intervention literature below.

The second way of looking at evidence that the pursuit of happiness is self-defeating is to examine experiments which demonstrate such an effect. I will detail three studies which attempted to show whether the pursuit of happiness is self-defeating.

Attempting to self-induce happiness while being cognitively distracted creates an ironic effect. In a series of two experiments, Wegner, et al., (1993) demonstrated that subjects under a cognitive load (memorizing a seven-digit number) experienced an ironic effect when they tried to re-live a happy or sad moment. That is to say, participants who memorized the seven digit number experienced a change in mood in the opposite direction of what they intended. Those attempting to be happier became more sad, and those attempting to be sadder became more happy. According to the ironic process theory of mental control, an attempt to control our own experience involves two separate processes: first, an operating process which brings about the intended change by identifying elements of experience consistent with the desired change; second, a monitoring process which identifies elements of experience inconsistent with the desired change to gauge how much of the operating process is needed to work (Wegner, 1994). When burdened with a cognitive task, the operating process is burdened but the monitoring process is not. This is because the operating process works at a conscious level but the monitoring process does not. The monitoring process does a good job of finding evidence that the operating process is needed, but the operating process is busy remembering the seven digit number. As a result, the monitoring process over-influences our experience and produces ironic effects. Though researchers were successful in
ironically inducing happiness for those trying to be sad, this discovery has not been used as a method to increase happiness to date!

Sometimes the pursuit of happiness can backfire even without a cognitive load. There is experimental lab evidence and field study evidence to support this point. The hypothesis for the lab experiment was that the pursuit and assessment of happiness can sometimes be self-defeating (Schooler, et al., 2003). Subjects listened to a nine-minute piece of classical music selected because it was thought to be “hedonically ambiguous.” Before they listened, they were divided into four groups: no instructions (control), instructions to try to be as happy as possible, instructions to continually monitor their happiness throughout the piece via a movable scale, and instructions to both try to be happy and to continually monitor their happiness. The researchers used three measures of happiness. Embedded within many distracter questions about music to protect the control group from contamination, answered on a Likert-type scale were these critical questions: “How happy do you feel right now?” and “What is your mood right now?” They also used a schematic face that could be adjusted from frowning to smiling to indicate their mood. Each participant completed these measures before and after listening to the music.

Across all three measures, those who were in the combined monitor/try to be happy condition felt worse after the intervention and those who were in the no instructions condition felt better after the intervention. However, although it is generally true that those trying to be happy, monitor their happiness, or both did worse than controls, the results were somewhat mixed depending on the outcome measure. This may be due in part to the use of ad hoc measures of happiness which have no previously
established evidence of validity or reliability. Furthermore, details about the number of participants, means and standard deviations of each group, effect sizes and significance levels were not included in the published results. For these reasons, this should be considered pilot-like data that has shown that directly pursuing and monitoring happiness can have adverse effects.

If conscious attempts to increase happiness undermine our ability to become more happy, then events intentionally designed to increase happiness also may have an ironic effect. If true, this is valuable to the pursuit of happiness researchers because expectancy could serve as a mechanism for the failure of attempts to increase happiness. In fact, this effect has been demonstrated both in a field experiment (Schooler, et al., 2003) and in a lab experiment (Della Porta, Sin, & Lyubomirsky, 2009). I will both describe the studies and their importance in the broader field of happiness studies.

The field experiment demonstrated an expectancy effect for enjoyment of a major event. Participants were polled the last week of December 1999 regarding their plans for New Year's Eve. They were asked four questions—What size celebration are you planning? How much do you expect to enjoy New Year's Eve? How much time, and how much money are you planning on spending on your celebration? There was just one follow up question—How much did you enjoy New Year's Eve? Over 450 people participated in the study, and 85% of them did not enjoy New Year's as much as they predicted. Those who planned the largest celebrations were the most disappointed, those who planned no celebrations were the least disappointed. Using the difference score between anticipation and enjoyment as the dependent variable, the researchers regressed the difference score on anticipation, time spent, and money spent in preparation. The
model was statistically significant as were the coefficients for anticipation and time spent in preparation, but not money spent. The researchers conclude that expectation leads to disappointment, and suggest that monitoring decreases enjoyment (based on the assumption that those with higher expectations for the night were also more frequently monitoring to decide whether they were enjoying themselves), and that therefore efforts to achieve happiness can backfire. Further, they claim that the results of this study supplement the results of the first study, showing that even when participants select their own goals and pursue happiness by their own chosen method, such efforts can have the opposite result of what is intended.

There are problems with integrating this study into the body of SWB literature. First, the researchers did not assess positive or negative affect, merely anticipated and actual enjoyment of a specific event. They also did not assess cognitive appraisal of SWL. Subjective well-being typically is defined by these three characteristics, and it is unknown how anticipation of enjoyment might be related to subjective well-being. Second, it seems unsurprising that those planning bigger events would have a great gap between expectation and experience simply because there is more possibility for disappointment. This may be the case for any planned activity, regardless of whether the activity is meant to be particularly pleasurable. Nevertheless, this study does hint at the need for further experimental exploration of the topic.

There is also experimental evidence that possible gains in subjective well-being may be undermined by expectancy effects. Using a randomized, placebo-controlled trial, researchers have demonstrated that increasing the expectancy of the effectiveness of a happiness intervention creates an ironic effect for the intervention (Della Porta, et al.,
In one such demonstration, four groups were divided using a 2x2 factorial design with type of intervention (letter writing vs. placebo control) and expectancy (induced vs. not induced) as the factors. Within the placebo-control group, those who were induced to believe that the placebo would be effective in boosting happiness experience no gain or loss on average, while non-expectant placebo group experienced an increase in happiness. The intervention group experienced a decrease in happiness regardless of expectancy, and this result will be discussed in the review of happiness interventions below. Compared to the field study of expectancy effects described above, this experiment is both more tightly controlled and therefore easier to integrate into the body of happiness intervention studies. Expectancy effects should be further investigated as a mechanism for the ironic effects of attempting to increase happiness.

In summary, the two lab experiments, the field study, and the in vivo experiment all found similar results generally. That said, these results are not uniform within or across studies. For example, although a cognitive load induced ironic effects in the pursuit of happiness, it is not clear in the study whether those without a cognitive load were able to significantly increase their happiness (Wegner et al., 1993). It has further been demonstrated that those without a cognitive load can experience ironic effects, but this finding was not consistent across all outcome measures (Schooler, et al., 2003). It is clear that more well-designed experiment with valid and reliable outcome measures and sufficient power are needed to put this hypothesis to a strong test.
Experiments Demonstrating Increases in Happiness

Despite evidence that the pursuit of happiness can be self-defeating, happiness interventions are effective when analyzed in aggregate (Sin & Lyubomirsky, 2009). In order to look carefully at the scientific evidence that happiness can be increased, I will now review studies of happiness intervention validations. I will do so by first, briefly reviewing known correlates of happiness which may be changeable and therefore effortful; next, discussing one methodological issue, use of placebo, for tests of happiness interventions in general; and finally, discussing the body of published happiness intervention studies starting with one early program of research and continuing by categorizing the intervention studies in to six groups, grouped by type of intervention.

In perhaps the most heavily cited documentation of the correlates of happiness, Argyle (1999) summarizes the correlations of happiness with many variables, including age, education, social class, income, marriage, relationships, work, ethnicity, employment, leisure, religion, life events, intelligence, and attractiveness. All together, these variables account for 10-20% of the variability in happiness. However, these correlations are of limited usefulness when it comes to an individual deciding how to maximize her own happiness for three reasons. One, many things for which there have been significant correlations could reasonably be explained by a third variable. For example, age, employment, and income are all closely related variables. Two, happiness can be a cause of some of these variables and not strictly a consequence (Lyubomirsky, King, & Diener, 2005). Third, there is a great amount of individual variability in happiness, making correlations low in general and not necessarily applicable or relevant for individuals (Diener, 2008).
As explained in the measurement of happiness section above, happiness reports may be particularly sensitive to behavioral cues. Because of this, for studies of happiness interventions to be valid, it is particularly important that they use an appropriate placebo control. One illustration of this point is found in Seligman, et al. (2005). In this randomized trial of five happiness interventions, the placebo produced a large jump immediately after the intervention before returning to baseline and staying there one week later. This illustrates the need for a placebo control in a study because even the placebo group experienced an increase in happiness scores immediately following treatment. If a placebo treatment (a neutral intervention that is hypothesized to be completely unrelated to increases or decreases in SWB) can have a positive effect on SWB, a treatment can only be declared effective if it is significantly more effective than a placebo control. Subjective well-being researchers have not always followed this guideline, however. In one example, the experimental group attended two 7-hour trainings followed by 3 weeks of 45-minute follow-up trainings, five days a week. The control group received only pre and post measures (Papaousek & Schulter, 2008). The intervention was successful in improving SWB but it is impossible to know the cause of this effect given the extreme nature of the intervention and lack of an appropriate placebo control.

Fordyce (1977, 1983) was a pioneer in the field of happiness interventions. He constructed a tri-pronged program of increasing happiness based on what he called “14 fundamental techniques” including increasing positive thinking, eliminating negative rumination, and spending more time socializing. He used a quasi-experimental design with intact college classes instead of random assignment to the experimental and control
groups. He used a placebo control for the control group class and found significant increases in SWB, some of which lasted more than 18 months. Since that time, researchers have generally focused on isolating one intervention at a time.

Interventions that have received experimental attention generally fall into six categories: gratitude, goals, positive self-visualization, prosocial behaviors, reminiscing, and signature strengths. I will review each of these areas below. Because I am interested in possible ironic results, as I review these studies I will include studies yielding both statistically significant and non-significant findings. Positive Psychology differentiates itself from prior similar movements largely on the grounds that Positive Psychology researchers use the traditional natural science paradigm and methods in their investigations. For this reason I will be reviewing those studies which used an appropriately rigorous method. I define rigorous loosely, meaning only that the experimenters used random assignment and a placebo control. If the literature for a particular intervention is especially sparse, I will mention some less rigorous studies with qualification and only in the name of completeness.

**Expressing gratitude, counting blessings, and happiness.** Though it has been the subject of many experimental investigations, the effectiveness of gratitude as an intervention to increase happiness is still very much in question. Experiments using gratitude as an intervention to increase happiness have yielded mixed results. There have likely been many attempts to validate gratitude as a happiness intervention because there is well developed theoretical background for why gratitude interventions should work. An intervention that increases gratitude could increase happiness in several ways. First, gratitude can promote savoring the pleasurable parts of life so that each positive life event
is milked for maximum hedonic benefit. This maximization may, in turn, reduce the effects of the hedonic treadmill by making the hedonic effects of positive life events stretch and last longer, reducing the need for constant variety. Also, gratitude might be incompatible with negative emotion states like anger or envy (Lyubomirsky, Sheldon, & Schkade, 2005).

To empirically test these theories, Emmons and McCullough (2003) conducted a series of three experimental studies using random assignment. First, participants were asked to make weekly reports and were divided into three groups: gratitude reflections, hassles (the experimental groups), and life events (the control group). The study lasted ten weeks and the participants made ten weekly reports. Outcome was measured with a list of 30 affect terms rated on a scale from 1-5, physical symptoms, global appraisals (terrible to delighted for both the past week and next week), and amount of exercise. There was a small and statistically significant boost for the grateful group for global evaluation of life as a whole, optimism for the upcoming week, and physical health symptoms. The largest effect for the grateful group was an increase in the amount of exercise. Study 2 used a daily journal for 2 weeks. The life events control condition was replaced with a downward social comparison condition, in order to create a condition similar to gratitude but which was conceptually distinct. Based on the increase in exercise time for the grateful group in Study 1, a healthy behaviors measure was added. Finally, a two-question measure of pro-social behavior question was added with these two questions: “Did you help someone with a problem today? Did you offer someone emotional support today?” The gratitude intervention did not differ from the control group on any measure. Increases in health and exercise time did not replicate. Study 3
used a sample of adults with neuromuscular disorders. They were divided into two groups: daily journaling of gratitude reflections and a control who only completed the measures (no placebo). They kept a daily journal for 3 weeks. A spousal rating of affect and SWL was added. Tiny statistically significant differences in positive and negative affect in the expected direction were found, but to get this effect researchers more than doubled the amount of participants in the experimental condition by adding in data from another study and then used maximum likelihood factor analysis instead of simply comparing means of positive and negative affect terms. The gratitude group was higher in global assessment of life and optimism for the upcoming week. Spousal ratings indicate higher SWL as well.

The effects of gratitude interventions on happiness were generally small and it is unknown whether they will endure. Although there were differences between the hassle and gratitude groups, there were never affective differences between gratitude and the placebo control condition. In Study 3 there was an affective difference, but there was no placebo control—control group participants simply filled out the measures. In study 2 there also was not a difference in global judgments between the gratitude and the control condition—this difference occurred only in study 1. Overall, comparing the gratitude with the control group, effects were in the predicted direction but failed either to replicate or to reach statistically significant levels. Although the results are mixed and some questionable methods and analyses were used (particularly in Study 3), that the results were in the predicted direction is notable in that future gratitude interventions can be developed using these studies as a foundation. As the researchers acknowledge, their manipulations were fairly minimal and not equivalent to the development of long term
gratefulness or a grateful personality. Although there were some consistent results when comparing the gratitude and the hassles group, it is unclear from these studies alone whether this was due more to the hassles intervention or the gratitude intervention.

Parts of these studies were replicated in surprising ways, including evidence about whether complaining or gratitude was more responsible for the differences between the groups, when a quasi-experimental study was conducted with 221 young adolescents (Froh, Sefrick, & Emmons, 2008). Measures similar to those used in the studies above were used, including a measure of positive and negative affect and an assessment of SWL. Classes were assigned to one of three conditions: daily gratitude exercise, daily hassles exercise, or control (only completing the daily SWB measures). There were no differences among the three groups for the PA measure. For NA, there was a statistically significant difference, with the hassles group being lower than the other two groups—which were statistically the same. However, although they did not differ statistically, the mean NA score was higher for the gratitude condition than the control condition. Furthermore, the gratitude group was not statistically different from either the hassle or control group with regard to SWL, but the control and the hassle group did differ, with the control group being higher. This is evidence that writing about hassles decreases SWL more than writing about gratitude increases it. This replication does not instill more hope for experimental manipulations of gratitude of this type.

Surprisingly, the negative effect of gratitude as a happiness intervention has been partially replicated a third time (Lyubormirsky, Sheldon, & Schkade, 2005). Using a control group that only completed the outcome measures, researchers compared two experimental groups. The first completed a “count your blessings” exercise in which
participants were instructed to “contemplate the things for which you are grateful” once a
week for six weeks. The second group did the same, but contemplated three times a
week. Though the outcome measure is not identified, the control group decreased and the
once a week blessing group increased in “Subjective well-being”. However, the three-
times-a-week blessing group also decreased in SWB. Because the scale, the metric,
significance tests, and effect sizes all are missing from the report, it is difficult to know
how these results compared to the Emmons and McCullough studies. One notable aspect
is that Emmons and McCullough concluded that because their daily journal participants
appeared to reap greater benefits from the intervention than those who made a weekly
report, more frequent interventions were preferable to less frequent interventions. Based
on their data, Lyubomirsky, Sheldon, and Schkade (2005) came to the opposite
conclusion, hypothesizing that the too frequent practice of gratitude exercises triggers the
hedonic treadmill (as described above), and therefore results in diminished SWB returns.

Another attempt to use a gratitude-based intervention to increase SWB failed to
show results. Sheldon and Lyubomirsky (2006) tested gratitude along with a life events
control condition. The intervention failed to increase PA or decrease NA more than a
placebo control condition. Even worse news comes from placebo-controlled experiment
using random assignment with a sample of depressed people (Della Porta, Sin, &
Lyubomirsky, 2008). In this study, writing a gratitude letter caused a statistically
significant decrease in well-being compared to the placebo control group.

Despite these discouraging findings, there may be a purely methodological reason
why gratitude studies have failed to validate gratitude interventions. It is possible that the
studies were simply not long enough. Using a newly constructed outcome measure (the
Steen Happiness Index) and a random assignment, placebo-controlled design, Seligman et al. (2005) examined the effects of five interventions on both SWB and depression symptoms in a non-clinically depressed sample. Each intervention lasted one week, and participants completed one-week, one-month, three-month, and six-month follow ups. The interventions related to gratitude were the gratitude visit and the “three good things” intervention (the results for the other three interventions will be discussed, each in its respective section). The gratitude visit intervention consisted of participants writing and then hand-delivering a letter to someone to whom they felt a debt of gratitude. The three good things intervention consisted of asking participants to write down three things that went well and a causal explanation for why they went well. At immediate post-test, gratitude visit participants experienced the biggest increase in happiness of the six interventions but by the six-month follow up, their happiness and depression scores were identical to the control group. The three things group began with much lower happiness scores at the initial post-test and at the one week follow up, but had statistically significant higher scores at one-month, three-month, and six-month follow ups. These results are promising in the sense that they offer an alternative way to understand the repeated failures of gratitude interventions to show effects. Perhaps the failed gratitude studies would have shown effects if they had assessed their participants for long-term effects.

Overall, evaluating gratitude as an intervention, there is not yet strong scientific support that increases in happiness are possible through gratitude interventions. The evidence is mixed and gratitude may cause decreased SWB for some people. However, based on existing evidence there is some cause for hope and continued investigation of
the long term benefits of daily recognizing three positive things as well as short term benefits of thoughtfully expressing gratitude in person.

Goals, planning, and happiness. Goal pursuit is a particularly interesting intervention because it can be used to illustrate the primary research question of this paper—Can happiness be obtained by pursuing it directly? Whereas setting and achieving goals can be linked to SWB (Houser-Marko & Sheldon, 2008), those who explicitly link successfully accomplishing a goal with happiness tend to be more ruminative and depressed (McIntosh, Harlow, & Martin, 1995). Goal setting also is a promising potential intervention for increasing subjective well-being because it is a process-oriented intervention. It is dynamic and changing, and therefore less likely to be adapted to or flattened by the hedonic treadmill (King, 2008). Unfortunately, a strong and well designed validation of these theories has not yet been performed. The one study with which a control group was used to test the efficacy of goal planning in increasing happiness has serious methodological problems (MacLeod, Coates, & Hetherton, 2007).

Researchers performed two studies in which they asked participants to volunteer either for the experimental or control condition. This is a major methodological flaw which makes the fruitful interpretation of data nearly impossible. Additionally, researchers found differences between those who volunteered for the control and experimental groups at pre-test. In both studies, they found a significant effect for goal-training on SWL, but failed to find a significant difference for positive or NA. Again, based on the flawed design, this data is at best very weak support for goal setting in SWB.

Positive self visualization and happiness. “Best possible self” (BPS) or “imagining yourself at your best” means imagining and sometimes writing about your
own future after you have achieved all of your most important goals. Imagining success in goals is thought to improve actual success rates (Pham, 1999), improve psychological adjustment (Taylor, Pham, Rivkin, & Armor, 1998), give a feeling of control (Lyubomirsky, Sousa, & Dickerhoof, 2006), bring greater awareness to values and priorities (McGregor & Little, 1998), and reduce goal conflict (Pennebaker, 1998). For these reasons, imagining your best possible self seems like a good candidate for a happiness intervention. However, BPS has not been a well validated intervention. One study demonstrated that participants in BPS condition increased in happiness immediately after the intervention, but at the follow up measure lost twice as much as they had gained for a net decrease in PA (Sheldon & Lyubomirsky, 2006). The opposite is true for NA. Participants experienced a decrease following the intervention but at the follow up measurement increased NA by nearly twice as much as they had decreased for a net gain of NA. At the follow up, the control condition had slightly less PA but also had less NA than BPS. These results are similar to one earlier study (King, 2001). Participants in another study were assigned to write about best possible selves or a control topic for 20 minutes a day, four days in a row. Those who participated in a BPS condition experienced a gain in positive mood, but less so than participants in the control condition. Though this difference was not statistically significant, the direction of the results indicate that the intervention made the people less happy than if they would not have had it.

When BPS interventions do yield results, they can be short-lived and can leave the participant worse off up to six months later. A third study follows this pattern, using a one-week intervention. Once again, participants failed to show an increase in happiness
for a “you at your best” activity compared to a placebo control immediately following the intervention. BPS participants also failed to show increases at the one week, one-month, three-month, and six-month follow ups (Seligman et al., 2005). BPS participants did show a significant decrease in depression scale scores immediately following the intervention, but these differences disappeared by the one-week follow up. Six months after the study, BPS participants were less happy than placebo controls (though the significance of this relationship is not reported). Thus, despite good theoretical reasons to believe it might be effective, imagining oneself at one's best has yet to be shown effective as a method to increase happiness and it may in fact cause harm.

_Pro-social behavior and happiness._ Does kindness increase happiness? Published studies tend to suffer from poor designs or inconsistent results. As an example of a poor design, one quasi-experimental study used entire classes for each instead of random assignment to avoid contamination effects. The intervention consisted of asking participants to count their own kind acts for one week and report them at the end of the week (Otake, Shimai, Tanaka-Matsumi, Otsui, & Fredrickson, 2006). The outcome measure was a Japanese version of the Subjective Happiness Scale, which has been shown to have acceptable validity and reliability (Shimai et al., 2004). Those who counted kindness experienced a greater increase than a no-treatment control. However, the hypothesis tested cannot provide evidence for kindness, but simply being aware of kindness. This finding would be more interpretable if they had used random assignment to groups in the experiment design.

Inconsistent findings characterize kindness studies. Another example of this is the Lyubomirsky, Sheldon, and Schkade (2005) six-week experiment. Participants were
assigned to one of three groups: those asked to perform five acts of kindness spread across the week for six weeks; those asked to perform five acts of kindness per week for six weeks, but to perform them all on the same day; and a no treatment control. On average, the control and the “spread out” kindness group experienced decreases in well-being and the single day group experienced an increase. Standard deviations, effect sizes, or even the outcome measure or metric being used on the graph are not reported, making it impossible to contextualize these results. For those reasons and the reasons outlined above, both of these studies can be seen as only preliminary evidence that kindness interventions might increase SWB.

Finally, long term census data has been used to marshal support for the connection between pro-social behavior and happiness. Meier and Stutzer (2004) used data from a large-scale (22,000 participants) German Socioeconomic Panel survey, which ran from 1985-1999. They found that volunteering is correlated with higher SWL and that more frequent volunteering is correlated with higher SWL. This correlation holds even when controlling for many different variables such as family income, unemployment, sex, age, etc. Further, when Germany reunified in 1990, many national social organizations were disbanded and volunteer opportunities disappeared. After reunification, average life dissatisfaction in East Germany decreased. This decrease was disproportional for those who had volunteered and been forced to stop after reunification, even after controlling for other variables like job loss. Though not an experiment, this lends supporting evidence to the hypothesis that volunteering can enhance SWL. The researchers conclude by noting that those volunteers who are primarily extrinsically motivated benefit less from volunteering than those who are primarily intrinsically motivated. On the same note,
they leave the question open as to whether offering institutional incentives (such as college admission) in exchange for volunteering is a self-defeating endeavor.

Volunteering seems promising as a cause of happiness in everyday life, but perhaps not as a happiness intervention due to the decreased benefits when done for extrinsic reasons.

*Reflection on positive experiences and happiness.* The ways in which one reminisces about past life events may impact SWB (Lyubomirsky, Sousa, & Dickerhoof, 2006). However, based on experimental evidence, these differences likely are modest. In a test of the effects of writing, talking, or thinking about a positive past life event, using many outcome measures such as satisfaction with life, PA, NA, and many health dimensions, few differences were found. First, those who talked about their positive life event had a somewhat lower SWL score than the other groups including the control. Second, those who thought about life events had significantly higher SWL scores compared to the other experimental conditions, but not compared to the control group. Notably, the placebo comparison group had the highest PA score of the four conditions. In another study, participants were assigned to think or write about a positive life event, and in so doing to either mentally replay the event or analyze it (Lyubomirsky et al., 2006). Among the many comparisons made, the only difference found was that the think-replay group was somewhat higher in a long-term PA measure. The researchers tested many groups to see whether they were different in a statistically significant way. When researchers make many comparisons in a given sample, alpha levels should be adjusted to account for multiple tests. Failing to adjust alpha levels via such methods as Bonferroni's adjustment increase the risk for a Type I error. No such adjustment was made in this study. Therefore, given the inflated alpha level and the increased possibility of a type I
error, only modest claims regarding the significance of the difference should be made. It is notable that even without controlling alpha level, most of these interventions failed to show an impact on PA, NA, and subjective well-being. Though these results are not promising, others have demonstrated that using imagery to reminisce about positive experiences is more powerful in increasing happiness than using memorabilia, and both are significantly more beneficial than a placebo control (Bryant, Smart, & King, 2005).

Signature strengths and happiness. The positive psychology movement consistently has sought to supplement a focus on disease, negative emotionality and dysfunction in traditional psychology with an opposing focus on optimal functioning, positive emotions, and well-being. To this end, Peterson and Seligman (2004) published a complement to the Diagnostic and Statistical Manual, Classification of Strengths. This book was intended to be used to classify 24 human characteristics thought to be strengths of character, such as curiosity, valor, and zest. As part of the large validation study of 5 happiness interventions Seligman, et al. (2005) used the signature strengths in two different ways as two separate interventions. In the first, participants simply took an assessment of their signature strengths, received the results with regard to their five highest strengths, and were asked to use them more often during the week of the study. In the second, after receiving feedback about their five highest signature strengths, participants were asked to use their signature strengths in new and different ways and given several examples of how they could do so. Although these are presented as two separate interventions, they are only subtly different. Surprisingly, outcomes for the two interventions were dramatically different. For the first intervention, mean happiness score only differed from the control at the post-test. The effect disappeared at all
subsequent follow ups. For the second intervention, there was no detectable positive effect on happiness until the one-week follow up, and effects continued at the three and six-month follow-ups as well. Thus, it appears that identifying signature strengths might be an effective happiness intervention if it is coupled with instructions to use them in new ways and specific examples of how one might do so.

*Mindfulness, happiness, and brain activation.* Mindfulness is the final intervention which has been tested using a placebo control. Mindfulness training has been found to increase subjective well-being amongst rheumatoid arthritis patients when compared to a placebo control group (Zautra et al., 2008). Other mindfulness studies have shown benefit to participants but lacked a controlled comparison group (see Bedard et al., 2003, for one example). Mindfulness is a particularly interesting intervention for the research question at hand, because to be successfully mindful of oneself in a lived moment, it is likely that one would have to let go of thinking about and working toward future happiness. This is an interesting paradox which could be clarified with future research.

Mindfulness has also been studies using electrical brain activity as a dependent variable (Davidson et al., 2003). Twenty-five participants were trained for eight weeks in mindfulness meditation. Compared to a wait-list control group, these participants experienced greater increases in left-side anterior brain activation. This area of the brain has been thought to be associated with PA based on case studies of patients with brain lesions or stroke damage. The use of brain activation to validate measures of happiness as well as to validate happiness interventions is a likely future direction for the field of happiness studies. Though this may be a promising direction for happiness measurement,
it cannot be taken as strong evidence that happiness can be increased for two reasons: first, the study lacks a placebo control. The problems with not using a placebo controlled study in studies of happiness change have been discussed above. The second reason is that meditation participants did not experience an increase in PA compared to wait list participants, but a decrease in NA. Negative affect is not hypothesized to manifest merely as the flip side of PA in brain activation—they appear to be two separate phenomenon involving different parts of the brain (Damasio et al, 2000, George et al., 1995, Lane et al., 1997). This complicates interpreting results from this study because it is unclear how to understand a reported decrease in NA according to a self-report scale and an apparent manifestation of increased PA in an electrical brain activity. Additional research clarifying the relationship between brain activation and self-report measures is needed.

There are a small number of other possible interventions that might show positive benefits, such as forgiveness (Freedman & Enright, 1996) or life coaching (Green, Oades, & Grant, 2006). However, these interventions never have been tested with a placebo control and therefore their relative value is as yet unknown.

One final note regarding all studies of happiness interventions: much of this research is based on a problematic assumption about the proportion of explainable happiness. It is generally believed that 50% of the population variation in happiness is accounted for by a genetic set point (Lykken & Tellegen, 1996), and that mostly stable life circumstances account for about 10% (Argyle, 1999). Thus, some SWB researchers conclude that 40% of our happiness is left up to intentional activity (Lyubomirsky, Sheldon, & Schkade, 2005; Sheldon & Lyubomirsky, 2007). The problem with this logic
is two-fold. One, this is arguing from a variation on the “god of the gaps” -type logic—because something is unaccounted for by mechanistic science, any explanation can be used to fill in the gap in knowledge. While any given explanation is possible, there are many other possibilities as well. Perhaps further mechanistic causes have not yet been discovered, for example. The second problem, as pointed out by Diener (2008), is that although 50% of happiness may be heritable for identical twin samples reared apart, that does not imply that this figure applies to any particular individual, inside or outside the sample. Heritability figures are derived from the variation between subjects in a particular sample and have no necessary relationship with what might make any particular individual more or less happy (Diener, 2008).

The Next Step for Studies of Effort and Happiness

It should be clear that as whole, the body of happiness intervention studies has shown inconsistent support for the notion that happiness can be willfully increased. Part of this inconsistency could be due to persistent methodological problems, measurement issues, and gaps in the reporting of results. Those issues aside, some interventions, such as the “best possible self” intervention, appear to be a complete failure and even possibly cause long-term harm to participants. Other interventions, such as expressing gratitude in person, appear to have some scientific evidence of their efficacy in increasing happiness, if only in the short term. Besides increasing in-person gratitude, using signature strengths, practicing mindfulness, and recording three good things that happen during the day are two other promising happiness interventions. Even if there is just one effective intervention, this is evidence that the pursuit of happiness need not be self-defeating.
Tests of happiness interventions must be combined with tests of intentionality in order to resolve the gap between the two bodies of research I have reviewed. It remains unknown whether these few interventions with the beginnings of empirical support would be equally effective at increasing SWB if they were explicitly performed in the pursuit of happiness. It is possible that an emphasis on increasing happiness would rob them of their intrinsic value and transform the goals from other-focused to self-focused. On the other hand, participants might have been successful in these interventions simply because they wanted to be—the successful interventions simply provided them enough of an external cue that they were able to decide that they were happy. Do people need to put forth effort to increase happiness or will the result in disappointment and ironic effects? At this point, looking at all the data would yield a conflicted answer at best. We know that sometimes some people are helped by certain happiness interventions, and we know that sometimes some people are hurt by trying to be happy. Much correlational and cross-sectional research has been conducted on SWB. Experimental methods including experiments with a longitudinal design are vital to the continuing revision, refining, and sophistication of happiness research (Diener, 2008) as well as answering basic questions such as whether happiness ought to be pursued.

Preliminary experiments have already been done in a laboratory setting and have shown that assessment and direct pursuit of happiness can be self-defeating (Schooler et al., 2003). Placebo-controlled studies with random assignment are needed both for replication as well as for evaluating the power of the effect in a real-life setting. If further evidence of the detrimental effects of pursuing happiness emerges, a scientific test of the theory that the assessment of happiness in a real life setting would help clarify one
possible mechanism. Participants could be randomly assigned to one of two groups: frequent testing and non-frequent testing. The non-frequent testing group is simply given a “placebo” test each day of the study, followed with a SWB measure on the final day. The frequent testing group is given a SWB measure each day of the study. Scores from the final day of the experiment would be compared between groups to determine whether frequent assessment of happiness has a damaging effect. This is one way of answering the question of whether the assessment of subjective well-being causes a decrease in subjective well-being. This is important to the study of increasing happiness because assessment of happiness is necessary to gauge whether an increase has occurred. If the hypothesis that frequent assessment causes decreases in happiness is supported, this provides one possible mechanism to understand why the pursuit of happiness can be self-defeating: Those who seek happiness are continually assessing their own happiness and are thereby dulling their ability to feel happy.

One way to further advance the scientific study of the intentional control of happiness is to isolate intentionality as an independent variable. Intentionality could be experimentally isolated in at least two profitable ways. One experiment to validate intentionality as a method to increase happiness would be to use a design similar to those that have been used to validate happiness interventions. That is, using a placebo control group and random assignment, give instructions to participants to try to be happier. A second experiment would couple explicit intentionality with already known effective happiness interventions. Using random assignment, each participant would be placed in a placebo control group or one of two experimental groups, both with an identical happiness intervention. One group would be encouraged to use the intervention to try to
become as happy as possible. Members of this group would know that the purpose of the intervention is to increase happiness. The second group would be encouraged to complete the intervention but the purpose of the intervention would be disguised. These experiments would provide direct evidence that either the pursuit of happiness is necessary for obtaining happiness or, on the other hand, that the pursuit of happiness poisons itself and creates an impoverished hedonic experience. This is the aim of the current study.
CHAPTER 2:

METHOD

It has not been previously investigated whether the assessment and pursuit of happiness impedes the achievement of increased happiness in a real-life setting. Two experiments were conducted to investigate this question. The first experiment examines the effect of conscious effort on the pursuit of happiness. It explores the question of how conscious effort to be happier and traditional happiness interventions work together and whether conscious effort interferes with happiness. The second experiment deals with the assessment of happiness. It examines the question of whether the assessment of happiness impacts happiness. Both studies were conducted online via qualtrics.com, an online research host.

Participants

Participants were recruited from the Notre Dame Psychology subject pool as well as social media websites (for example, facebook.com), websites devoted to promoting psychology experiments (for example, socialpsychology.org), and related psychology listservs (for example, the APA positive psychology list-serv). Advertisements on these sites and list-servs also encouraged participants to forward the information and the experiment link to their friends. Participants could choose whether to receive entry in a drawing for
one of six $15 gift certificates or course credit for participation. All participants were treated in accord with APA ethical guidelines.

Measures

The dependent variables for each study are the three commonly identified aspects of subjective well-being: PA, NA, and SWL. All participants in all three studies will complete two measures in order for a complete assessment of subjective well being. These measures are the SWL scale and the Positive Affect Negative Affect Schedule.

The SWL scale is a five-item measure of overall SWL. Items are statements such as “The conditions of my life are excellent” and “I am satisfied with my life.” Responses are measured on a 7-point scale ranging from 1 ("strongly disagree") to 7 ("strongly agree"). Many studies have examined the internal consistency and the reports of alpha range from .79 to .89 (Pavot & Diener, 1993). Measures of test-retest reliability vary based on the study and time interval, and \( r \) tends to be above .8 for intervals less than one month (Pavot & Diener, 1993). At the four year interval, \( r = .54 \) (Magnus, Diener, Fujita, & Pavot, 1992). SWL thus appears to be a relatively stable construct in the short term with no intervention. The SWL has been shown to have acceptable sensitivity for use in studies of clinical interventions (Pavot & Diener, 1993).

The PANAS (Watson, Clark, & Tellegen, 1988) is a 20-item scale with two 10-item affect subscales—positive and negative. Test takers are asked to indicate the extent to which he or she has experienced each item over a given period of time. Items consist of one word descriptions of affect such as “interested,” “attentive,” and “strong” (Positive) and “disinterested”, “guilty”, and “afraid” (negative). Responses are given in a
five point scale ranging from 1 (“very slightly or not at all”) to 5 (“extremely”). Many
time frames have been used including “in this moment”, “this past year”, and “in
general.” The positive and negative affect subscales have been shown to be largely
independent, sharing between 1% and 5% of their variance (Watson et al., 1988).
Regardless of time frame, alpha reliabilities have been shown to be consistently
acceptable, ranging from .84 to .90 (Crawford & Henry, 2004). Test-retest reliability at an
eight week interval ranges from .39 to .71 and is strongly related to the time frame in
question (as would be expected, the highest reliability was found for “in general”
followed by “in the last year” and so on, down to “in this moment.”) (Watson et al, 1988).

Experiment 1

Design. Participants were randomly assigned to one of five groups. These groups
were divided along dimensions of effort (either try to be happy or no instructions
regarding happiness) and activity (either the control activity or the three good things
activity also known as the “bona fide” activity). The two levels of these two factors, when
crossed, account for four groups. The fifth group can be termed the “pure try” group—
this group was given the instruction to try to be happy without being given an
accompanying activity. This group was created to isolate the effects of effort as much as
possible. We will employ two factorial designs to account for all five groups. This is
necessary because of non-overlapping conditions amongst the groups. In the first
analysis, the first four groups will be compared in a 2x2 MANCOVA design, with pre-test
SWL score as the covariate, and the three components of subjective well-being as the
dependent variables (Positive Affect and Negative Affect scale scores from the PANAS,
and the SWL scale score). In the second analysis, each of the three groups which were given instructions to “try to be happy” will be compared in a 3x1 MANCOVA analysis with the levels control activity, three good things activity, and no activity (pure try). As with the first analysis, the pre-test SWL scale score will be included as a covariate and the three measures of subjective well-being will be the dependent variables.

Procedure. The central question for experiment 1 is whether people can consciously try to be happier and succeed simply by trying to so be. After consenting to participate and taking the pre-test, participants were randomly assigned to one of five groups: control, try to be happy, intervention, try to be happy/intervention combined, and pure try group. Participants received a reminder e-mail with repeated instructions each day of the study.

The no effort/control activity group were given an assignment based on the recollection of early memories taken from Seligman et al, 2005 (see Appendix A). The willful effort/control activity group was given the instructions found in Appendix B. The no willful effort/three good things group received the instructions found in Appendix C. The willful effort/three good things group received the instructions found in Appendix D. The willful effort/no activity group received the instructions found in Appendix E: After the post-test, there was a manipulation check of two questions to ask all participants whether they participated in the assignment and if so, how often they did it. Following the initial study, participants were asked to report at one week and one month intervals in order to test effects of the intervention beyond immediate impact. Following the final follow-up and debriefing, three winners were randomly selected from among the participants who entered themselves in to the drawing and will be contacted via e-mail.
Experiment 2

Design. The purpose of experiment 2 is to evaluate the impact of measuring subjective well-being on levels of subjective well-being. This was tested using a two group pre-post-follow up design. Each participant took a pre-test of the three components of subjective well-being and was randomly assigned (using block randomization) to either the control or experimental condition. Both groups took an online daily assessment for five days, a different measure each day. In the experimental group, all of the assessments were directly related to subjective well-being. The control group took measures unrelated to subjective well-being. After five days, all participants took a post-test which consisted of the same three measures of subjective well-being. After one week, all participants took an identical one-week follow up assessment. Only participants who complete pre, post, and follow up measures were included in the study.

Procedure. After consenting to participate, participants were randomly assigned to groups. All participants took the SWL and the PANAS as both pre- and post-test. The control group took the pre-test, completed the five control measures five times, one every day, and the post-test. The experimental group took the pre-test, the five happiness measures, one each day, and the happiness post-test. All tests were administered online. Participants received a reminder e-mail each day of the study with a link to the test they were to take that day.

Between the pre and post tests, the control group took five different placebo measures. These measures were selected as control measures because they are theoretically unrelated to well-being, mood, or affect and are thus thought to be inert with regard to impact on SWB. The measures are the following: Curiosity and Exploration
Inventory (Kashdan, Rose, & Fincham, 2004), Activism Orientation Scale (Corning & Myers, 2002), Juror Bias Scale (Kassin & Wrightsman, 1983), Social Desirability Scale (Crowne & Marlowe, 1960), and Extreme Response Style measure (Greenleaf, 1992). In addition to their theoretically inert natures, these measures were selected because they have approximately the same number of items as the measures of subjective well-being in the experimental group.

Between the pre and post tests, the experimental group took five different measures related to subjective well-being, one per day. These measures are the Subjective Happiness Scale (Lyubomirsky & Lepper, 1999), Meaning in Life Questionnaire (Steger, Frazier, Oishi, & Kaler, 2006), Authentic Happiness Inventory (Peterson & Park, 2006), Center for Epidemiological Studies-Depression Scale (CES-D; Radloff, 1977), and Happiness Measures (Fordyce, 1988). These measures were selected because they are directly related to the self-evaluation of happiness. Measures different from those being used at the pre- and post-tests were selected in order to prevent practice effects.

The dependent variable is change in subjective well-being. Participants were asked to report again one week following the post-test. Following the final follow-up, participants were debriefed and those who opted for entry in the drawing were selected.
CHAPTER 3: RESULTS

Experiment 1: Does willful effort undermine happiness?

The purpose of experiment 1 was to determine the impact of willful effort on happiness. In the sample recruited for this study, 413 of the initial 580 people, or 71%, completed all follow up surveys. Only participants who completed the experiment, the week follow-up, and the month follow-up were included in the analysis. Those who dropped out scored, on average, about 1.2 points lower on the SWL at baseline than those who completed the study, drop out $M = 24.2$, $SD = 6.6$; completers $M = 25.4$, $SD = 6.0$; $t(575) = -2.09$, $p = .037$. Those who dropped out did not systematically vary in group assignment, as tested by a chi-square goodness-of-fit $\chi^2(4, N = 167) = 5.99$, $p = .20$.

The sample is mostly single never-married (54.6%) women (67%), mostly white (85%), and US residents (94%). Participants had a mean age of 28.4 (SD = 11.8) and 48% possessed at least 4 year college degree.

Participants were randomly assigned to one of five groups as described in the methods section. Table 3.1 shows the means for each of the five groups on each of the three measures at post-test. Two factorial designs were employed to account for all five groups. This was necessary because of non-overlapping conditions amongst the groups. In the first analysis, the first four groups were compared in a 2x2 MANCOVA design, with pre-test SWL score as the covariate, and the three components of subjective well-
being as the dependent variables. Next, each of the three groups were given instructions
to “try to be happy” in a 3x1 MANCOVA analysis with the levels control activity, bona
fide activity, and no activity (pure try). All analyses were run with and without outliers,
and direction and size of the effects remained the similar so without theoretical
justification to remove them, all participants were included in the analyses reported here.
For clarity, consistency, and convenience, in this report those assigned to the control
activity group will be referred to “no activity”, those assigned to the bona fide activity
group will be referred to as “yes activity,” those assigned to not make an effort to try to
be happy will be referred to as “no try,” and those assigned to make an effort to try to be
happy will be referred to as “yes try.”

As a manipulation check, participants were asked to report how many days out of
six possible days they performed their assigned task. The mean number of days was 4.64,
SD = 1.62, median = 5 days, mode = 6 days. Eight out of 412 participants (1.9 percent)
reported not doing their activity at all. Because the overwhelming majority, 98.1%,
reported performing the activity at least once, and the majority performed that activity
more days than not (76.2%), the manipulation was considered successful. To see whether
the groups differed systematically in how many times they performed their activity
during the week, pairwise comparisons were made between the five groups. After
performing a Bonferroni adjustment to the critical \( p \) value of .05 and arriving at the new
critical \( p \) value of .0125, only one pairwise comparison was significant at the \( p = .05 \)
level: The yes activity yes try (\( M = 6.05, SD = 1.3 \)) vs. the no activity yes try group (\( M =
5.15, SD = 1.8, t(143.68) = 3.58, p < .001 \)). (Note: not assuming equal variances based on
a significant Levene’s test of equal variances, \( p < .001 \).) The yes activity yes try group
TABLE 3.1

DESCRIPTIVE STATISTICS OF ALL FIVE EXPERIMENTAL GROUPS ON ALL THREE MEASURES OF SWB AT POST-TEST.

<table>
<thead>
<tr>
<th>Group</th>
<th>Satisfaction with Life</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Deviation</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>no act no try</td>
<td>25.83</td>
<td>6.129</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>yes act no try</td>
<td>25.99</td>
<td>5.836</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>yes act yes try</td>
<td>26.22</td>
<td>6.434</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>no act yes try</td>
<td>26.26</td>
<td>6.152</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>pure try</td>
<td>27.07</td>
<td>5.147</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>All groups</td>
<td>26.27</td>
<td>5.940</td>
<td>412</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>Positive Affect</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Deviation</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>no act no try</td>
<td>30.17</td>
<td>7.022</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>yes act no try</td>
<td>32.76</td>
<td>6.545</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>yes act yes try</td>
<td>34.43</td>
<td>6.294</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>no act yes try</td>
<td>33.64</td>
<td>7.018</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>pure try</td>
<td>33.94</td>
<td>7.486</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>All groups</td>
<td>33.01</td>
<td>7.000</td>
<td>412</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>Negative Affect</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Deviation</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>no act no try</td>
<td>19.97</td>
<td>6.855</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>yes act no try</td>
<td>19.62</td>
<td>6.423</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>yes act yes try</td>
<td>20.23</td>
<td>7.739</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>no act yes try</td>
<td>20.90</td>
<td>7.630</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>pure try</td>
<td>17.94</td>
<td>5.494</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>All groups</td>
<td>19.73</td>
<td>6.907</td>
<td>412</td>
<td></td>
</tr>
</tbody>
</table>
did their assigned activity nearly one time more, on average, than the no activity, yes try group. The omnibus test, or the test of the overall model, was statistically significant for PA, $p < .001$, $F (4, 323) = 12.81$ and SWL, $p < .001$, $F (4, 323) = 125.46$, but not significant for NA, $p = .051$, $F (4, 323) = 2.39$.

Effort and Activity. The first four groups were divided by two dimensions: Effort (yes try and no try) and Activity (yes activity and no activity). The outcomes for these four groups were entered into a 2 (effort type) x 2 (activity type) complex-design MANCOVA with pre-test SWL score as the co-variate and the three parts of subjective well-being (described above) as the three dependent variables. See tables 3.2, 3.3, and 3.4 for means and marginal means by DV and by group.

### TABLE 3.2

<table>
<thead>
<tr>
<th></th>
<th>Try</th>
<th>No Try</th>
<th>Activity</th>
<th>Try</th>
<th>No Try</th>
<th>No Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Activity</td>
<td>26.24</td>
<td>25.91</td>
<td>26.08</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SATISFACTION WITH LIFE
MEANS AND MARGINAL MEANS AT POST-TEST
TABLE 3.3

POSITIVE AFFECT
MEANS AND MARGINAL MEANS AT POST-TEST

<table>
<thead>
<tr>
<th>Activity</th>
<th>Try</th>
<th>No Try</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Try</td>
<td>34.43</td>
<td>32.76</td>
<td>33.60</td>
</tr>
<tr>
<td>No Activity</td>
<td>33.64</td>
<td>30.17</td>
<td>31.91</td>
</tr>
<tr>
<td>No Activity</td>
<td>34.04</td>
<td>31.47</td>
<td>32.76</td>
</tr>
</tbody>
</table>

TABLE 3.4

NEGATIVE AFFECT
MEANS AND MARGINAL MEANS AT POST-TEST

<table>
<thead>
<tr>
<th>Activity</th>
<th>Try</th>
<th>No Try</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Try</td>
<td>20.23</td>
<td>19.62</td>
<td>19.91</td>
</tr>
<tr>
<td>No Activity</td>
<td>20.57</td>
<td>19.80</td>
<td>20.18</td>
</tr>
<tr>
<td>No Activity</td>
<td>20.90</td>
<td>19.97</td>
<td>20.44</td>
</tr>
</tbody>
</table>

The overall interaction effect of activity x effort was not statistically significant, $F(3, 323) = .562, p = .64$. The main effect for activity across levels of effort was significant $F(3, 323) = 2.767, p = .042$, partial eta squared = .029. The main effect for effort groups across levels of activity was also significant $F(3, 323) = 5.269, p = .001$, partial eta squared = .049. Though SWL at posttest was significant in the overall model, it failed to
show significant differences in any of the main effect or interaction effect tests because it was so highly correlated with the covariate of SWL at pretest. SWL at the pretest was correlated with each of the dependent variables and is therefore an appropriate covariate.

The Spearman rho correlations with SWL at pretest were as follows: N = 412 for each SWL at time 2 rho = .80, p < .001, PA rho = .31, p < .001, NA rho = -.16, p < .001. A non-parametric correlation was chosen because it was not assumed that the distributions were normal. Pearson’s correlation coefficients were very similar, however. This similarity is likely due to the large sample size for which assumptions of normality as more robust to violation. Each of the measures demonstrated acceptable reliability at posttest: SWL $\alpha = .85$, PA $\alpha = .88$, NA $\alpha = .88$.

Samples sizes were not equal in each cell so the assumption of the equality of covariance matrices was tested using Box’s M test. Box’s M (18, 370454) = 13.6, $p = .77$. Therefore, the null hypothesis that the covariance matrices are equal fails to be rejected and the assumption of equality holds.

Because the interaction of activity x effort was not significant, it is possible to interpret the main effects of activity and time separately. Further comparisons were made to determine between-subjects effects for each dependent variable. To maintain an overall p value of .05 for multiple contrasts, the Bonferroni adjustment was used and .05 was divided by the number of contrasts, 6. This gave a new p-value cut-off of .0083. For the effect of type of activity across levels of effort, the yes activity group was higher on PA ($M = 33.7, SD = 6.4$) than those in the no activity group ($M = 31.7, SD = 6.4$). See figure 3.1. This is a statistically significant difference, $p = .006$, F (1, 329) = 7.43. The yes activity group was, on average, slightly higher than the no activity group on SWL and
slightly lower than the no activity group on NA, though these differences failed to reach significance. See Figures 3.2 and 3.3. Table 3.5 lists the means, standard errors, and p values for each of the three dependent measures.

Covariates appearing in the model are evaluated at the following values: SWLTOTAL1 = 25.17

Figure 3.1. Positive affect at post-test
Figure 3.2. Satisfaction with life at post-test
Figure 3.3. Negative affect at post-test

Covariates appearing in the model are evaluated at the following values: SWLTOTAL1 = 25.17
TABLE 3.5

MEAN DIFFERENCES, STANDARD ERRORS, P VALUES, 
AND CONFIDENCE INTERVAL FOR CONTRASTS

TESTING MAIN EFFECT OF ACTIVITY ACROSS LEVELS OF EFFORT

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Mean Difference</th>
<th>Std. Error</th>
<th>p value</th>
<th>95% Confidence Interval for Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWL</td>
<td>-.227</td>
<td>.675</td>
<td>.736</td>
<td>-1.554</td>
</tr>
<tr>
<td>PA</td>
<td>-2.048</td>
<td>.737</td>
<td>.006</td>
<td>-3.499</td>
</tr>
<tr>
<td>NA</td>
<td>.178</td>
<td>.789</td>
<td>.822</td>
<td>-1.375</td>
</tr>
</tbody>
</table>

For the effect of level of effort across activity, those who were instructed to try to be happy had a PA mean of 34.1 while those who were not instructed had a mean of 31.4. This difference was significant ($p < .001$, $F(1, 329) = 13.57$). The yes try group was, on average, slightly higher than the no try group on SWL and also slightly higher than the no try group on NA, though these differences failed to reach significance ($p < .05$). See table 3.6 for means, standard errors and p values each of the three dependent measures.

Examining individual cells to make simple main effect comparisons showed where further differences between groups exist in PA. Within the no try group, those in the yes activity group scored higher on PA ($M = 32.9$, $SD = 6.7$) than those in the no activity group ($M = 29.9$, $SD = 6.6$), $t(165) = 2.91$, $p = .004$, $d = .45$. Similarly, for participants in the no activity group, those in the yes try group scored higher ($M = 33.5$, $SD = 7$) than those in the no try group ($M = 29.9$, $SD = 7$) $t(157) = 3.24$, $p < .001$, $d = .51$. 

55
In other words, type of activity was only important for those who were not trying to be happy, and whether someone was trying to be happy was only important if they were doing a control activity. Though SWL at time 2 was significant in the overall model, it failed to show significant differences in any of the main effect or interaction effect tests because it was so highly correlated with the covariate of SWL at time 1, Pearson’s $r = .786$, $p < .001$.

Likewise, doing a bona fide happiness-boosting activity (as opposed to a control activity) results in increased PA regardless of whether one is trying to be happy or not. The effect sizes for effort and activity are modest and similar. It should be noted that though the differences failed to reach significance, the general pattern of effort and activity increasing components of happiness (by decreasing NA and increasing SWL) was present and consistent. Neither trying to be happy, using a happiness-boosting activity

### TABLE 3.6

MEAN DIFFERENCES, STANDARD ERRORS, P VALUES, AND CONFIDENCE INTERVAL FOR CONTRASTS TESTING MAIN EFFECT OF EFFORT ACROSS LEVELS OF ACTIVITY

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Mean Difference</th>
<th>Std. Error</th>
<th>p value</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWL</td>
<td>-.261</td>
<td>.675</td>
<td>.699</td>
<td>-1.588</td>
<td>1.066</td>
</tr>
<tr>
<td>PA</td>
<td>-2.630</td>
<td>.737</td>
<td>.000</td>
<td>-4.081</td>
<td>-1.179</td>
</tr>
<tr>
<td>NA</td>
<td>-1.033</td>
<td>.789</td>
<td>.192</td>
<td>-2.586</td>
<td>.520</td>
</tr>
</tbody>
</table>
activity, nor a combination results in a statistically significant change in either direction for the other two measures of happiness, NA and SWL.

Though it was not a central research question, the question of time was examined to test whether the effects found became more pronounced, less pronounced, or stayed the same over time. Using a MANCOVA analysis, time was introduced as a three level within-subject factor with the levels end of activity, one week follow-up, and one month follow-up. The following effects were evaluated: Main effect of time and interactions of time x activity, time x effort, and time x activity x effort. The results of this test show whether the components of happiness changed after the one-week activity was over, after controlling for pre-test SWL score. See Table 3.7 for complete results. Neither the main effect for time nor the interactions of time x activity, time x effort, and time x activity x effort were significant at \( p = .05 \). Time x activity F (6, 320) = .34, \( p = .91 \), time x effort F (6, 320) = 1.1, \( p = .37 \), time x activity x effort F(6, 320) = 1.3, \( p = .26 \). This shows that participants stayed at essentially the same levels of SWL, PA, and NA from the post-test to the one week and five weeks follow-up assessments. Using MANCOVAs to follow-up on this finding, it was found that there was a significant main effect of effort, a significant main effect of activity, and no significant interaction effects for effort x activity at both the one week and the five week follow-ups. There were no significant effects for either of the other two DVs at either time point in either direction. See Tables 3.8 and 3.9 for results of tests of between-subject effects at the one week and five week follow-up, respectively. In summary, the small effect that effort and the small effect that activity had on PA were persistent, and the non-effect of effort and activity on NA and SWL also did not change with time.
### TABLE 3.7

BETWEEN AND WITHIN SUBJECT EFFECTS FOR FACTORS TIME, ACTIVITY, TRY 
AND INTERACTIONS TIME * ACTIVITY, TIME * TRY, AND TIME * ACTIVITY * TRY.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Value</th>
<th>F</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>p value</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWL</td>
<td>.324</td>
<td>225.078</td>
<td>3</td>
<td>323</td>
<td>.676</td>
<td></td>
</tr>
<tr>
<td>Activity</td>
<td>.971</td>
<td>3.213</td>
<td>3</td>
<td>323</td>
<td>.023</td>
<td>.029</td>
</tr>
<tr>
<td>try</td>
<td>.951</td>
<td>5.603</td>
<td>3</td>
<td>323</td>
<td>.001</td>
<td>.049</td>
</tr>
<tr>
<td>activity * try</td>
<td>.998</td>
<td>.237</td>
<td>3</td>
<td>323</td>
<td>.870</td>
<td>.002</td>
</tr>
<tr>
<td>Within Subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>time</td>
<td>.987</td>
<td>.679</td>
<td>6</td>
<td>320</td>
<td>.667</td>
<td>.013</td>
</tr>
<tr>
<td>time * SWL</td>
<td>.989</td>
<td>.571</td>
<td>6</td>
<td>320</td>
<td>.754</td>
<td>.011</td>
</tr>
<tr>
<td>time * activity</td>
<td>.994</td>
<td>.342</td>
<td>6</td>
<td>320</td>
<td>.914</td>
<td>.006</td>
</tr>
<tr>
<td>time * try</td>
<td>.980</td>
<td>1.091</td>
<td>6</td>
<td>320</td>
<td>.367</td>
<td>.020</td>
</tr>
<tr>
<td>time * activity * try</td>
<td>.976</td>
<td>1.300</td>
<td>6</td>
<td>320</td>
<td>.257</td>
<td>.024</td>
</tr>
</tbody>
</table>
TABLE 3.8

BETWEEN SUBJECT EFFECTS AT ONE WEEK FOLLOW-UP

<table>
<thead>
<tr>
<th>Source</th>
<th>DV</th>
<th>Type III SS</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p value</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWL Total</td>
<td>SWL</td>
<td>7641.019</td>
<td>1</td>
<td>7641.019</td>
<td>571.382</td>
<td>.000</td>
<td>.637</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>1379.206</td>
<td>1</td>
<td>1379.206</td>
<td>32.103</td>
<td>.000</td>
<td>.090</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>226.109</td>
<td>1</td>
<td>226.109</td>
<td>4.766</td>
<td>.030</td>
<td>.014</td>
</tr>
<tr>
<td>Activity</td>
<td>SWL</td>
<td>.988</td>
<td>1</td>
<td>.988</td>
<td>.074</td>
<td>.786</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>252.291</td>
<td>1</td>
<td>252.291</td>
<td>5.872</td>
<td>.016</td>
<td>.018</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>40.590</td>
<td>1</td>
<td>40.590</td>
<td>.856</td>
<td>.356</td>
<td>.003</td>
</tr>
<tr>
<td>Try</td>
<td>SWL</td>
<td>1.116</td>
<td>1</td>
<td>1.116</td>
<td>.083</td>
<td>.773</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>172.064</td>
<td>1</td>
<td>172.064</td>
<td>4.005</td>
<td>.046</td>
<td>.012</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>144.889</td>
<td>1</td>
<td>144.889</td>
<td>3.054</td>
<td>.081</td>
<td>.009</td>
</tr>
<tr>
<td>Activity *</td>
<td>SWL</td>
<td>14.077</td>
<td>1</td>
<td>14.077</td>
<td>1.053</td>
<td>.306</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>1.624</td>
<td>1</td>
<td>1.624</td>
<td>.038</td>
<td>.846</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>58.146</td>
<td>1</td>
<td>58.146</td>
<td>1.226</td>
<td>.269</td>
<td>.004</td>
</tr>
<tr>
<td>Error</td>
<td>SWL</td>
<td>4346.180</td>
<td>325</td>
<td>13.373</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>13962.522</td>
<td>325</td>
<td>42.962</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 3.9

BETWEEN SUBJECT EFFECTS AT ONE MONTH FOLLOW-UP

<table>
<thead>
<tr>
<th>Source</th>
<th>DV</th>
<th>Type III SS</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p value</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWL Total</td>
<td>SWL</td>
<td>7438.144</td>
<td>1</td>
<td>7438.144</td>
<td>451.971</td>
<td>.000</td>
<td>.582</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>775.534</td>
<td>1</td>
<td>775.534</td>
<td>15.120</td>
<td>.000</td>
<td>.044</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>366.710</td>
<td>1</td>
<td>366.710</td>
<td>7.620</td>
<td>.006</td>
<td>.023</td>
</tr>
<tr>
<td>Activity</td>
<td>SWL</td>
<td>.085</td>
<td>1</td>
<td>.085</td>
<td>.005</td>
<td>.943</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>194.829</td>
<td>1</td>
<td>194.829</td>
<td>3.799</td>
<td>.052</td>
<td>.012</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>.079</td>
<td>1</td>
<td>.079</td>
<td>.002</td>
<td>.968</td>
<td>.000</td>
</tr>
<tr>
<td>Try</td>
<td>SWL</td>
<td>3.667</td>
<td>1</td>
<td>3.667</td>
<td>.223</td>
<td>.637</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>714.979</td>
<td>1</td>
<td>714.979</td>
<td>13.940</td>
<td>.000</td>
<td>.041</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>20.809</td>
<td>1</td>
<td>20.809</td>
<td>.432</td>
<td>.511</td>
<td>.001</td>
</tr>
<tr>
<td>Activity * Try</td>
<td>SWL</td>
<td>15.640</td>
<td>1</td>
<td>15.640</td>
<td>.950</td>
<td>.330</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>.437</td>
<td>1</td>
<td>.437</td>
<td>.009</td>
<td>.927</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td>27.130</td>
<td>1</td>
<td>27.130</td>
<td>.564</td>
<td>.453</td>
<td>.002</td>
</tr>
<tr>
<td>Error</td>
<td>SWL</td>
<td>5348.566</td>
<td>325</td>
<td>16.457</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>16669.488</td>
<td>325</td>
<td>51.291</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Just trying to be happy. The next analysis was conducted to determine whether trying to be happy had a differential effect on each of three groups: bona fide activity aka yes activity, control activity aka no activity (both included in the previous analysis) and a pure try condition, which received only the instruction to try to be happy at the post-test. The pure try group is the only novel group in this analysis. As has already been established above, there is no statistically significant difference between the yes activity and no activity when they are both in the yes try condition. Therefore the only contrasts of interest are whether the pure try group differs from either of the other two groups on any of the three outcome measures. See table 3.10 for descriptive statistics of all outcome measures at post-test. Controlling for pre-test SWL score, there was a significant overall effect amongst the three groups for NA, F (2, 240) = 3.63, p = .028, partial eta squared = .029. There were no significant differences amongst the three groups for SWL, F(2, 240) = .20, p = .816, nor for PA, F(2, 240) = .45, p = .641). This was followed up with two pairwise comparisons (pure try vs. each of the other two groups compared on NA since these are now the only relationships of interest.) To control p-value inflation, the Scheffe adjustment was made to the p values and confidence intervals. The pure try group was significantly lower M = 17.94, SD = 5.5, meaning they reported less NA, than the no activity group M = 20.9, SD = 7.6 p = .03, 95% CI (-5.69, -.23). The pure try group was lower than the yes activity group but not significantly so, M = 20.23, SD = 7.4, p = .17, 95% CI (-5.00, .42). See Figure 3.4. This means that simply trying to be happy with no additional activity results in lower NA than trying to be happy with a placebo activity.

Using a MANCOVA analysis, time was introduced as a new factor with three levels: immediately following the intervention, or posttest (time 1), one week after the
### TABLE 3.10

**DESCRIPTIVE STATISTICS AT POST-TEST**

**ON EACH DEPENDENT MEASURE**

<table>
<thead>
<tr>
<th>DV</th>
<th>Try Category</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>yes try</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWL</td>
<td>yes activity</td>
<td>26.22</td>
<td>6.434</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>yes try</td>
<td>26.26</td>
<td>6.152</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>no activity</td>
<td>27.07</td>
<td>5.147</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>pure try</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>26.52</td>
<td>5.929</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td>yes try</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>yes activity</td>
<td>34.43</td>
<td>6.294</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>yes try</td>
<td>33.64</td>
<td>7.018</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>no activity</td>
<td>33.94</td>
<td>7.486</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>pure try</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>34.01</td>
<td>6.923</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td>yes try</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>yes activity</td>
<td>20.23</td>
<td>7.739</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>yes try</td>
<td>20.90</td>
<td>7.630</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>no activity</td>
<td>17.94</td>
<td>5.494</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>pure try</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>19.69</td>
<td>7.119</td>
<td>244</td>
</tr>
</tbody>
</table>
I tested whether the within-subject effect of time and the between subject effects of activity performed differentially across the three measurements of subjective well being. Pre-test SWL score was included as a covariate on analyses of main and interaction effects. Descriptive statistics at the one week follow-up are presented in Table 3.11. Casually examining the data for trends at one week follow-up, it is notable that the pure try group is the highest of the three groups in SWL and PA and the lowest in NA at each time point. In other words, based on means, the pure try group is the happiest of the three groups.

The main effect for time was not significant, $F (6, 235) = 1.01, p = .42$. This means that across measures and activity groups, participants did not significantly change
TABLE 3.11

DESCRIPTIVE STATISTICS OF BETWEEN-SUBJECT EFFECTS AT FOLLOW UP 1.

<table>
<thead>
<tr>
<th>DV</th>
<th>Try Category</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWL</td>
<td>yes try yes activity</td>
<td>26.34</td>
<td>6.294</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>yes try no activity</td>
<td>26.91</td>
<td>5.894</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>pure try</td>
<td>27.49</td>
<td>5.158</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>26.91</td>
<td>5.801</td>
<td>244</td>
</tr>
<tr>
<td>PA</td>
<td>yes try yes activity</td>
<td>33.71</td>
<td>6.252</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>yes try no activity</td>
<td>32.11</td>
<td>6.977</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>pure try</td>
<td>34.00</td>
<td>6.637</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>33.28</td>
<td>6.649</td>
<td>244</td>
</tr>
<tr>
<td>NA</td>
<td>yes try yes activity</td>
<td>20.19</td>
<td>6.606</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>yes try no activity</td>
<td>20.58</td>
<td>7.142</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>pure try</td>
<td>18.37</td>
<td>5.834</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>19.71</td>
<td>6.591</td>
<td>244</td>
</tr>
</tbody>
</table>

on the three subjective well-being measures following the intervention. The time by activity interaction was also not significant, $F(12, 486) = 1.14$, $p = .36$. This means that the members of the three activity groups (bona fide, control, pure try) were not differentially affected by time across the subjective well-being measures. For the between and within subject contrasts, the assumption of sphericity was violated according to Bartlett’s test of sphericity (chi square $= 129.4$ between, $256.2$ within, $df = 5$, $p < .001$). Because there were no significant contrasts, no adjustment was necessary to make the tests more conservative.
The statistically significant difference that existed between groups at post-test on NA is no longer significant by the one-week follow-up, $F(2, 241) = 2.62, p = .08$. SWL and PA continued to show no group differences, SWL $F(2, 241) = .81, p = .45$; PA $(2, 241) = 1.90, p = .152$. See Table 3.12 for M and SD for each group at follow up 2 (the five week follow-up.) At the five-week follow-up, neither NA, SWL, nor PA showed significant difference between groups. NA $F(2, 241) = .051, p = .95$; SWL $F(2, 241) = .193, p = .82$; PA $(2, 241) = .714, p = .49$.

**Experiment 2: Does the assessment of happiness undermine happiness?**

The purpose of experiment 2 was to evaluate the impact of measuring subjective well-being on levels of subjective well-being. Of the 204 participants who completed the baseline measure, 174 (85%) completed the post-test and one week follow up. Of the 174 participants who completed all measures and follow-ups, 91 were in the experimental (happiness measures) group and 82 were in the control (control measures) group. Those who did not drop did not differ systematically with regard to group membership, goodness of fit $X^2(1, N = 173) = .468, p = .49$. Only those who completed all measures were included in this analysis. The sample is mostly US residents (91%) who are single (61%) female (67%) with mean age: 28.1 years, (SD = 11.9).

As a manipulation check, records were kept of how many people took each assessment each day. Overall, each assessment was taken by nearly every participant. Each participant took five assessments, for a total of 1020 possible assessment taken during the week. Of those possible 1020 assessments, 998, or 97.8%, were taken. In this sense, the manipulation was successful. Those assigned to the experimental (happiness
## Table 3.12

### Descriptive Statistics of Between-Subject Effects at Follow Up 2.

<table>
<thead>
<tr>
<th>DV</th>
<th>Try Category</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWL</td>
<td>yes try yes activity</td>
<td>26.43</td>
<td>6.606</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>yes try no activity</td>
<td>26.90</td>
<td>6.223</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>pure try</td>
<td>26.98</td>
<td>5.227</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>26.77</td>
<td>6.031</td>
<td>244</td>
</tr>
<tr>
<td>PA</td>
<td>yes try yes activity</td>
<td>34.67</td>
<td>6.103</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>yes try no activity</td>
<td>33.43</td>
<td>7.055</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>pure try</td>
<td>33.58</td>
<td>8.563</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>33.90</td>
<td>7.294</td>
<td>244</td>
</tr>
<tr>
<td>NA</td>
<td>yes try yes activity</td>
<td>19.99</td>
<td>7.363</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>yes try no activity</td>
<td>19.95</td>
<td>7.312</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>pure try</td>
<td>19.65</td>
<td>7.145</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>19.86</td>
<td>7.246</td>
<td>244</td>
</tr>
</tbody>
</table>

measures) group took happiness assessments nearly every day and those assigned to the control (control measures) group took control assessments nearly every day. Reliability was acceptable for each measure at posttest, PA $\square = .90$, NA $\square = .84$, SWL $\square = .88$.

The key research question is whether the assessment of happiness results in decreased happiness. If the results show no difference between groups or show that those who assessed their own happiness became happier than those who did not, this is evidence to support the assertion that the assessment of happiness does not result in decreased happiness. If the results show that the happiness measures group is
significantly lower at the post-test or follow-up, that is evidence to support the assertion that the assessment of happiness results in decreased happiness.

Descriptive analysis shows general trends in the results. Means, standard errors, and 95% confidence intervals are presented in Table 3.13. It is notable that for each measure at both the posttest and the follow-up, the mean score of the happiness measures group is lower than the control group only once, and only by a very small amount with SWL at follow up. For every other measure at both posttest and follow-up, the mean of the experimental group is higher.

Using a 2x3 repeated measures MANOVA design with group (with levels control & experimental) and time (with levels pre, post & one week follow up) as the factors, there is a significant time x group interaction F(6, 167)= 2.3, \( p = .035 \), partial eta squared = .077; a significant time main effect F(6, 167) = 7.5, \( p < .001 \), partial eta squared = .212; and a non-significant group main effect F(3, 170) = 1.0, \( p = .401 \). Because there is a significant interaction effect, straightforward interpretation of main effects is difficult or impossible. Based on the significant interaction effect, simple main effects were be analyzed to find where and how differences took place.

Simple main effect comparisons for group and measure at each time point are presented in Table 3.14. Notable comparisons are that for no measure did the experimental and control group differ at the pretest, \( p > .05 \), and that by the follow-up, the groups did not differ on any measure. Thus, it is reasonable to conclude that any effects of the manipulation were indeed a result of the manipulation (given the equal pre-test scores and the use of random assignment) and that any effects from the manipulation were short-lived (given the equal follow-up scores).
### TABLE 3.13

**DESCRIPTIVE STATISTICS OF EACH GROUP ON EACH MEASURE AT EACH TIME POINT**

<table>
<thead>
<tr>
<th>DV</th>
<th>Group</th>
<th>Time</th>
<th>Mean</th>
<th>Std. Error</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWL</td>
<td>control measures group</td>
<td>1</td>
<td>25.807</td>
<td>.698</td>
<td>24.429</td>
<td>27.186</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>25.325</td>
<td>.727</td>
<td>23.890</td>
<td>26.761</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>25.795</td>
<td>.706</td>
<td>24.402</td>
<td>27.189</td>
</tr>
<tr>
<td></td>
<td>happiness measures group</td>
<td>1</td>
<td>26.374</td>
<td>.667</td>
<td>25.057</td>
<td>27.690</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>25.901</td>
<td>.695</td>
<td>24.530</td>
<td>27.272</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>25.989</td>
<td>.674</td>
<td>24.658</td>
<td>27.320</td>
</tr>
<tr>
<td>PA</td>
<td>control measures group</td>
<td>1</td>
<td>32.145</td>
<td>.814</td>
<td>30.539</td>
<td>33.751</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>28.699</td>
<td>.781</td>
<td>27.158</td>
<td>30.240</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>28.590</td>
<td>.814</td>
<td>26.983</td>
<td>30.198</td>
</tr>
<tr>
<td></td>
<td>happiness measures group</td>
<td>1</td>
<td>32.110</td>
<td>.777</td>
<td>30.576</td>
<td>33.644</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>31.143</td>
<td>.746</td>
<td>29.671</td>
<td>32.615</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>30.516</td>
<td>.778</td>
<td>28.981</td>
<td>32.052</td>
</tr>
<tr>
<td>NA</td>
<td>control measures group</td>
<td>1</td>
<td>19.386</td>
<td>.670</td>
<td>18.063</td>
<td>20.708</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>20.193</td>
<td>.708</td>
<td>18.796</td>
<td>21.589</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>19.084</td>
<td>.692</td>
<td>17.718</td>
<td>20.451</td>
</tr>
<tr>
<td></td>
<td>happiness measures group</td>
<td>1</td>
<td>19.286</td>
<td>.640</td>
<td>18.023</td>
<td>20.549</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>18.495</td>
<td>.676</td>
<td>17.161</td>
<td>19.828</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>17.989</td>
<td>.661</td>
<td>16.684</td>
<td>19.294</td>
</tr>
</tbody>
</table>
TABLE 3.14

PAIRWISE COMPARISONS BETWEEN EXPERIMENTAL AND CONTROL GROUPS FOR EACH MEASURE AND AT EACH TIME POINT WITH BONFERRONI ADJUSTED CI.

<table>
<thead>
<tr>
<th>DV</th>
<th>Time</th>
<th>Mean Difference</th>
<th>Std. Error</th>
<th>p value</th>
<th>95% Confidence Interval for Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
</tr>
<tr>
<td>SWL</td>
<td>1</td>
<td>-.566</td>
<td>.966</td>
<td>.558</td>
<td>-2.473</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-.576</td>
<td>1.006</td>
<td>.568</td>
<td>-2.561</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-.194</td>
<td>.976</td>
<td>.843</td>
<td>-2.121</td>
</tr>
<tr>
<td>PA</td>
<td>1</td>
<td>.035</td>
<td>1.125</td>
<td>.975</td>
<td>-2.186</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-2.444</td>
<td>1.080</td>
<td>.025</td>
<td>-4.575</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-1.926</td>
<td>1.126</td>
<td>.089</td>
<td>-4.149</td>
</tr>
<tr>
<td>NA</td>
<td>1</td>
<td>.100</td>
<td>.926</td>
<td>.914</td>
<td>-1.726</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.698</td>
<td>.978</td>
<td>.084</td>
<td>-.233</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.095</td>
<td>.957</td>
<td>.254</td>
<td>-.794</td>
</tr>
</tbody>
</table>

Amongst the three pairwise comparisons of potential interest (each of the three measures at posttest), the only test of significance is PA. The mean difference between experimental and control groups is 2.44. With Bonferroni adjustment, this is a statistically significant difference, \( p = .025 \), 95% CI (0.313, 4.58). See Figures 3.5, 3.6, and 3.7. In other words, at the post-test the happiness measures group scored significantly higher on PA than the control measures group.
Figure 3.5. Positive affect in project 2, happiness measures compared to control measures
Figure 3.6. Negative affect in project 2, happiness measures compared to control measures.
Figure 3.7. Satisfaction With Life in project 2, happiness measures compared to control measures.
The worst way to achieve certain goals is by directly pursuing them. The primary aim of this study was to test whether achieving happiness is one such goal. This was done in two ways. Experiment 1 tested willful effort directly as an isolated factor. Effort alone was compared to a traditional happiness intervention in a randomized controlled trial looking at effort and the intervention alone and in combination with one another. If the pursuit of happiness is self-defeating, then the assessment of happiness is a potential mechanism for this relationship. It is possible that those who consciously pursue happiness assess their own happiness more often than those who do not. For this reason, the assessment of happiness was the construct of interest in experiment 2. Experiment 2 tested the hedonic impact of the assessment of happiness over the course of one week. The purpose of this study was to examine whether assessment of happiness decreased happiness. The results of each study are discussed below. To conclude, the studies will be discussed together in light of current research on happiness and effort, limitations of the current studies will be outlined, and future research directions will be explored.

Experiment 1

The data consistently support the notion that the pursuit of happiness does not undermine subjective well-being. The most important consideration in making this claim
is the direction of the differences between groups. There is not a single comparison between those who received instructions to try to be happy and those who did not in which those in the try group did worse than their counterparts. On the contrary, trying to be happy appears to have a minor beneficial effect on PA and no effect on NA and SWL. The two affective components of subjective well-being, positive and NA, were sensitive to experimental manipulations in certain conditions. The most stubborn aspect of happiness is cognitive appraisal. This is consistent with theoretical notions of cognitive appraisal, as this is meant to be a global assessment of one’s satisfaction with one’s life including history, current circumstances, and the projected future. It is not surprising that this is the variable with the least movement of the three components of subjective well-being. Even given its stubborn nature, however, these results are not entirely discouraging with regard to improving cognitive appraisal. Increased PA can, over time, lead to improved cognitive SWL (Cohn, Fredrickson, Brown, Mikels, & Conway, 2009).

One possible mechanism for the effectiveness of effort and activity is the frequency of actually doing the activity. Three pieces of data support this possibility. First, the most frequent actors were those participants in the yes try yes activity group, who performed the activity a full one day more, on average, than the lowest group, yes try no activity. Second, across groups there was a positive correlation between number days the activity was performed and SWL and NA at posttest. SWL and NA significantly correlated with number of days at $r = .10, p = .03$ and $r = -.12, p = .012$, respectively. This hypothesis is further supported by the non-significant correlation between pre-test SWL and number of days the activity was performed, $r = .08, p = .083$. One might observe that correlations of .08 and .10 are very similar, though one is significant and one is not.
Though each piece of data individually is not strong support for the importance of the frequency of activity in happiness, the data together begin to provide the basis of a possible mechanism.

Most importantly, the results of experiment 1 support the assertion that the pursuit of happiness does not undermine happiness. Those participants who were instructed to pursue happiness scored as well or better than their counterparts in all measures. Within those who received instruction to try to be happy, a control activity was no better than a bona fide happiness increasing activity. Having no additional activity was slightly better than both of the activities in regard to NA. It is possible those who saw a purpose in doing the activity, and were given a credible activity related to the stated purpose of the study were more motivated to participate in the study daily.

Those who dropped out before the study’s completion were somewhat lower on the SWL scale than those who completed the study. This represents a challenge in interpreting the results, but it is not clear how the results might be adjusted to reflect the dropout bias. Because the dropouts occurred evenly across the groups, it is reasonable to assume that the dropouts did not bias the between group comparisons. However, it is possible that results would have been different if dropouts had not varied systematically from non-dropouts. Again, it is not clear if the results are too conservative or too liberal. Because there was likely some ceiling effect for some participants, it is possible that the dropouts created an overly conservative bias with regard to the effect of effort and the intervention on happiness. Because they started out somewhat lower on SWL than those who completed the study, they had more room to improve during the week of effort and intervention. There is evidence for this interpretation. Those who got lower scores on the
post-test than the pretest for the SWL averaged a score of 24.65 on the pretest ($SD = 6.1$) whereas those who stayed the same or got better averaged 27.05 ($SD = 5.69$). This is a statistically significant difference, $t(410) = 3.69$, $p < .001$. In other words, those who got worse were more likely to have started out higher on SWL than those who stayed the same or got better. By eliminating some of the slightly lower scoring participants because they dropped out, it is possible that the results are overly conservative and the effect size of the improvement as a result of the intervention or trying to be happy is underestimated. The dropouts occurred essentially equally in all groups, so again, this did not bias the results of the between group comparisons in a clearly correctable way.

Is it possible to increase happiness using an intervention? These results reflect the ambivalence in the happiness intervention literature. The supposedly effective happiness activity rarely outperformed a control activity and never outperformed simply trying to be happy. Is it possible to will yourself happier? Again, the results are mixed. It can be safely said that trying to be happy appears to do no harm.

These results are inconsistent with the findings and conclusions of Schooler et al (2003) and are somewhat more consistent with the conclusions of Sin and Lyubomirsky (2009) and other happiness intervention researchers. That is to say, in this experiment the pursuit of happiness does not appear to be self-defeating and happiness interventions appear to be generally successful compared to a control activity. Another finding is that PA appears to be a separate construct from NA, given that the constructs behaved in different (though not opposite) ways in the experiment. This is consistent with the conceptualization of separate negative and positive affect outlined by Watson, Clark, & Tellegen (1988).
Experiment 2

For all aspects of SWB, there is a negligible effect of taking happiness assessments compared to control assessments. Taking happiness assessments may offer brief protection from a drop in PA. In other words, based on this experiment, the assessment of happiness is not self-defeating.

Limitations

Some limitations of this study include a reliance on self-report measures. Self-report is not a problem itself, as the construct of interest is subjective and self-conscious well-being is literally impossible to measure without using self-report. However, multiple types of measures can provide powerful converging evidence for a given phenomena (Cole, Howard, & Maxwell, 1981). Future studies might consider including supplemental measures like the report of a friend or partner, or a behavioral measure or checklist. Another limitation of this study is that some contrasts were under-powered and so though the direction of the effect was as predicted, the contrast failed to meet criteria significance at $p = .05$. Any study with dropouts becomes biased in potentially unknown ways. Dropouts in this study did not differ systematically with regard to group membership, but did differ systematically with regard to pre-test SWL score. On average, dropouts scored one point lower on the pre-test measure. This may have biased the conclusions in either direction. It is possible that slightly less satisfied people dropping out created a bias toward an increase in happiness overall in the outcome, because SWL at pretest and SWL at post-test were correlated, Spearman’s $r = .79$, $p < .001$. However, it is also possible that this created a decrease in happiness, because SWL at pretest is also significantly
negatively correlated with amount of change between pre and post test, Spearman’s $r = - .34, p < .001$. Therefore, it is unknown how the dropouts biased the final results and it appears that dropouts did not occur completely at random. Finally, the samples for each of the studies were heterogeneous but non-representative of any particular population. Concretely, the sample was biased toward those who have computer access and who check their email daily. It is possible that results could have differed with some other method of reminders and data collection.

Another, perhaps more important, limitation is some uncertainty regarding whether the manipulation worked, particularly in experiment 2 that examined the impact of the assessment of happiness on happiness itself. Although participants were remarkably compliant with the study protocol in the sense that nearly all participants took each assigned assessment nearly every day, it is unknown whether this manipulation achieved the actual goal of meaningfully increasing the frequency of happiness assessment for the experimental group. It is possible, for example, that we all assess our own happiness many times a day and that one formal assessment of happiness does not do much to change our overall habit of continually assessing how happy we are. If this is the case, the effect of the manipulation would be minimal because both the experimental and the control group would be more or less equivalent in how often they assess their own happiness. However, there is some evidence that we are not continually assessing our own happiness. Schooler et al. (2003) demonstrated that increasing the assessment of happiness can have deleterious effects on happiness in the lab, from which we can infer that we are not continually assessing our own happiness. However, it is possible that being asked about happiness in a lab by a researcher is not identical to self-assessing
happiness in the privacy of your own thoughts. Research using rigorous and validated dependent measures remains to be done on this phenomenon.

**General Discussion**

The results of both experiments can be summed up in these two key findings. The first key finding is that doing a bona fide activity or trying to be happy for a week has a significant and similar effect on positive affect, but combining the two provides no additional benefit. This finding can be interpreted in at least two ways. First, it is possible that receiving the instruction to try to be happy is itself a legitimate intervention for increasing positive affect. This is an important finding in and of itself. It implies that we are not all constantly maximizing our happiness, as many social science researchers implicitly assume (Slife & Williams, 1995). If we can increase our happiness by simply trying to do so, we may conclude that we are not constantly trying to increase our happiness as a default or baseline activity. This is true because to instruct a person who is already maximizing their own happiness to be happier ought not to increase their level of happiness. In this experiment, that instruction did impact positive affect. Therefore, those who were impacted were not already (prior to the experiment) maximizing their positive affect. This finding is also important because instructing others to try to be happier is an extraordinarily easy intervention to implement for those who wish to help others increase their happiness. If there was a measurement problem (for example, the ceiling effect already discussed above), then it is possible that the intervention to try to be happier may be more effective than is reflected in this experiment. Even if this turns out not to be the
case, it is reasonable to believe the new hypothesis that the pursuit of happiness is at least not self-defeating.

The second interpretation of this finding is that those in the bona fide activity group experienced genuine improvement while the change in the try group was a result of demand effect. This interpretation would undermine the first interpretation, but should be considered. Demand effects are pervasive in all research involving human participants and are difficult to detect. In this case, the purpose of the experiment was disguised from participants and this disguise worked. When asked to guess at the purpose of the experiment prior to taking the follow-up measures post-intervention, less than 10% of the participants in the non-try groups guessed that the purpose was related to happiness. Therefore only those in the try groups are likely to have systematically responded to demand effects to try to be happier. In this case, demands effects are negligible because whether the participants were trying to be happier based on the experimenter’s demands or whether the participants were trying to be happier for some other reason is irrelevant, assuming the validity of the self-report. In an applied situation, someone may try to be happier because a therapist, coach, or friend suggested it (demand effects) or someone may try to be happier for some other reason. Either way, the person is likely to be successful in increasing their positive affect. Furthermore, the dependent variables all behaved somewhat differently—if participants were merely trying to show what they suspected the experimenter was looking for, it is dubious that they would collectively decide to manifest this while taking one half of one measure (PA), but not on the other half of the measure (NA) nor on the other measure (SWL).
The second key finding is that assessment of happiness does not reduce happiness. One possible complication in the interpretation of these results is that there was significant overlap between the subject pool in experiment 1 and experiment 2. 139 out of the initial 210 participants in experiment 2, or 66%, had already participated in experiment 1. This is a possible problem because depending on the condition in experiment 1, they may have already been in the habit of assessing their own happiness in an online survey which may impact their reported happiness for the project. However, because the participants were randomized in both studies, there is no evidence that systematic error was present. Therefore, this is not a confound, merely a consideration when interpreting the results. It is possible some people in the control group were in the habit of regularly assessing their own happiness, thus deflating the difference found between the experimental and control group in this study. Another difficulty comes in interpreting a null result. The emphasis on this finding, therefore, is not on the existence on a null result but on the direction of the data in aggregate. There is a lack of evidence to support the assertion that the assessment of happiness causes decreases in happiness, but there could be further tests of this assertion if doubt remains. One such test is proposed below.

The two experiments, taken together, point toward a possible mechanism which links effort and increased positive affect. A new theory could explain the findings of the two experiments: those who are trying to be happy assess their own happiness more frequently. This theory would be consistent with mindfulness approaches to happiness, such as those discussed by Brown and Ryan (2003). This mechanism could be tested by studying the pursuit and assessment of happiness simultaneously in a factorial design.
This could be accomplished by conducting a four-group study in which the factors were pursuit of happiness with two levels (pursue and do not pursue) and assessment of happiness (assess happiness and do not assess happiness) and were completely crossed. Those assigned to pursue happiness would be given instructions identical to the “pure try” group in study 1. Those assigned to the “do not assess” condition would be given no instructions about pursuit of happiness. Those assigned to assess happiness would receive daily happiness assessment. Those assigned to not assess happiness would receive daily control assessments. This proposed experiment would extend the findings of the present study by pulling apart the relationship between the pursuit and assessment of happiness. If the pure pursue and the pure assessment groups have similar results, that would provide further evidence that those who are pursuing happiness assess their happiness more frequently and that the assessment of happiness is a mechanism for increased happiness.

Another future experiment could test the relationship between effort other related constructs in addition to subjective well-being. This is appropriate given that one prominent subjective well-being researcher has recently called for a de-emphasis on “happiness” as a key construct in research and a greater emphasis on other constructs such as meaning, relationships, and accomplishment (Seligman, 2011.) (It should be noted that not all positive psychology researchers agree with the de-emphasis on happiness as a central construct, notably Lyubomirsky, 2011.) The relationship between effort and meaning, positive relationships or accomplishment could be investigated in an experimental framework. Accomplishment is the least likely construct of these to be paradoxically related to willful effort, but there have not yet been studies to isolate willful
effort in a study of meaningfulness or positive relationships. Meaningfulness is most closely related to the cognitive appraisal aspect of subjective well-being. In this one-week study, cognitive appraisal was not affected by the bona fide happiness intervention. To test the relationship between effort and meaningfulness, a longer-term intervention that specifically targets meaningfulness should be included in an experiment with a similar structure to study 1. Efforts to increase meaning are certainly nothing new in psychology. Therapeutic approaches like logotherapy and existential psychotherapy address the issue of meaning directly (Frankl, 1963, Yalom, 1980.) It could be argued that all of therapy has an underlying goal to increase meaning for clients. Even more broadly, this may be a primary goal of all of the humanities and liberal arts. This proposed experiment differs from these broad enterprises because it would isolate effort as a variable in increasing meaningfulness by using the logic of a randomized and controlled scientific experiment.

A third future experiment could target a clinically depressed population as opposed to a general convenience sample. This approach would advance the field in two ways: first, it is possible that there is something of a ceiling effect in subjective well-being, thus under-estimating the effect of either effort or the intervention. If there is an effective ceiling on subjective well-being in a general sample, that would be consistent with the modest effect sizes in this study. By using a depressed population, this ceiling effect would be, presumably, eliminated. Second, a clinical sample would show whether the happiness interventions and pure effort could be useful in clinical practice. This is useful because though symptoms of depression have been shown to be reactive to happiness interventions, studies using a clinically depressed sample have rarely been published.
In a quote sometimes attributed to Nathaniel Hawthorne, an author once wrote that “Happiness is a butterfly, which pursued is always just beyond your grasp, but which, if you will sit down quietly, may alight upon you.” Further studies remain to be done, but in the light of the results of this study, it appears that the butterfly is somewhat more graspable than once thought.
APPENDIX A
NO EFFORT/CONTROL ACTIVITY GROUP INSTRUCTIONS

Consider for a moment your earliest memories. Out of all the experiences of a lifetime, we only hold onto a few in the form of early memories. A careful consideration of our earliest memories may help us better understand who we are today. Your assignment is as follows: Every night for one week, starting tonight, set aside 10 minutes before you go to bed. Use that time to think of an early memory and write it down in as much detail as possible. Try to remember what you were doing, what you were feeling, and the other people who were with you. (If you cannot remember some of these details, that is OK. Just write down what you can remember.) You may use a journal or your computer to write about the events, but please make sure you actually write (or type) your memories. After seven nights of doing this exercise, look back over your collection of memories. Notice any similarities or patterns across the memories. Then log back on the website so that we can learn how this exercise affected you.

To review, here are the steps of this exercise: Every night for one week, starting tonight, write down an early memory in as much detail as possible. On the seventh night, look for similarities or patterns in your memories. We will send you a link on the eighth day which will bring you back to our site to complete the study. We look forward to hearing from you.
APPENDIX B
WILLFUL EFFORT/CONTROL ACTIVITY GROUP INSTRUCTIONS

There are many things which impact our happiness. One of those things is willful effort. For this project, you will try to be happier than you currently are. We think too much about the present and the future and not enough about the past. Of course, it makes sense for us to think about the present and future. However, it also might be helpful to think more about the past. Most of us are not nearly as good at analyzing past events as we are at considering present and future, so this is a skill that needs practice. As you become better at focusing on the past, you will become more grateful for what you have, more hopeful about the future, and happier overall. So let’s get started.

Consider for a moment your earliest memories. Out of all the experiences of a lifetime, we only hold onto a few in the form of early memories. A careful consideration of our earliest memories may help us better understand who we are today. Your assignment is as follows: Every night for one week, set aside 10 minutes before you go to bed. Remember that the purpose of this exercise is to increase your happiness. Use that time to think of an early memory and write it down in as much detail as possible. Try to remember what you were doing, what you were feeling, and the other people who were with you. (If you cannot remember some of these details, that is OK. Just write down what you can remember.) You may use a journal or your computer to write about the events, but please
make sure you actually write (or type) your memories. After seven nights of doing this exercise, look back over your collection of memories. Notice any similarities or patterns across the memories. Writing about past events in your life may seem awkward at first, but please stick with it for one week. It will get easier. After seven nights of trying to increase your happiness by doing this exercise, we will send you a link so you can log back on to the website and tell us about your experience.

To review, here are the steps of this exercise:

1. Every night for one week, starting tonight, try to be happier by writing down an early memory in as much detail as possible. We will send you one reminder to work on your project each day.
2. On the seventh night, look for similarities or patterns in your memories.
3. On the eighth day we will send you a link so that you can return to the website and we can learn about your experience. We look forward to hearing back from you soon!
APPENDIX C

NO WILLFUL EFFORT/THREE GOOD THINGS GROUP INSTRUCTIONS

Every night for one week, starting tonight, set aside 10 minutes before you go to bed. Use that time to write down three things that went really well on that day and why they went well. You may use a journal or your computer to write about these events, but it is important that you have a physical record of what you wrote. It is not enough to do this exercise in your head. The three things you list can be relatively small in importance (“My husband picked up my favorite ice cream for dessert on the way home from work today”) or relatively large in importance (“My sister just gave birth to a healthy baby boy”). Next to each positive event in your list, answer the question, “Why did this good thing happen?” For example, someone might write that her husband picked up ice cream “because my husband is really thoughtful sometimes” or “because I remembered to call him from work and remind him to stop by the grocery store.” When asked why her sister gave birth to a healthy baby boy, someone might write that “God was looking out for her” or “She did everything right during her pregnancy.”

Writing about “why” the positive events in your life happened may seem awkward at first, but please stick with it for one week. It will get easier. After seven nights of doing this exercise, please log back on the website so that we can learn how this exercise affected you.
To review, here are the steps of this exercise:

1. Every night before bed for one week, think about three good things that went well that day.
2. Write down the three things that went well.
3. Then write down why each thing went well.
4. We will send you one reminder to work on your project each day. On the eighth day we will send you a link so that you can return to the website and we can learn about your experience. We look forward to hearing back from you soon!
APPENDIX D

WILLFUL EFFORT/THREE GOOD THINGS GROUP INSTRUCTIONS

There are many things which impact our happiness. One of those things is willful effort. For this assignment, you will try to be happier than you currently are. Psychologists have discovered a simple new way that we can make ourselves happier. We think too much about what goes wrong and not enough about what goes right in our lives. Of course, sometimes it makes sense for us to analyze bad events so that we can learn from them and avoid them in the future. However, people tend to spend more time thinking about what is bad in life than is helpful. Worse, this tendency to focus on bad events sets us up for anxiety and depression. One way to keep this from happening is to develop our ability to think about the good in life. Most of us are not nearly as good at analyzing good events as we are at analyzing bad events, so this is a skill that needs practice. As you become better at focusing on the good in your life, you will become more grateful for what you have, more hopeful about the future, and happier overall. So let’s get started.

Every night for one week, set aside 10 minutes before you go to bed. Use that time to write down three things that went really well on that day and why they went well. Remember that the purpose of this exercise is to increase your happiness. You may use a journal or your computer to write about the events, but it is important that you have a physical record of what you wrote. It is not enough to do this exercise in your head. The
three things you list can be relatively small in importance (“My husband picked up my favorite ice cream for dessert on the way home from work today”) or relatively large in importance (“My sister just gave birth to a healthy baby boy”). Next to each positive event in your list, answer the question, “Why did this good thing happen?” For example, someone might write that her husband picked up ice cream “because my husband is really thoughtful sometimes” or “because I remembered to call him from work and remind him to stop by the grocery store.” When asked why her sister gave birth to a healthy baby boy, someone might write that “God was looking out for her” or “She did everything right during her pregnancy.” Writing about “why” the positive events in your life happened may seem awkward at first, but please stick with it for one week. It will get easier. After seven nights of trying to increase your happiness by doing this exercise, please log back on the website so that we can learn how this exercise affected you.

To review, here are the steps of this exercise:
1. Every night before bed for one week, try to be happier by thinking about three good things that went well that day.
2. Write down the three things that went well.
3. Then write down why each thing went well.
4. We will send you one reminder to work on your project each day. On the eighth day we will send you a link so that you can return to the website and we can learn about your experience. We look forward to hearing back from you soon!
APPENDIX E
THE WILLFUL EFFORT/NO ACTIVITY GROUP INSTRUCTIONS

There are many things which impact our happiness. One of those things is willful effort. For this assignment, you will try to be happier than you currently are. We have a limited amount of activities we can do each day and it is possible that we only do a few that actually make us happier. A careful consideration of what we do to be happier may help us better understand how the process of happiness change works.

Your assignment is as follows: For one week, make a conscious effort to be happier. Do this through any safe means you deem appropriate. Pay attention to what you do to try to be happier. After seven days, log back on the website so that we can learn what you did and how this exercise affected you.

To review, here are the steps of this exercise: First, for one week starting today, try to be as happy as possible. We will send you one reminder to work on your project each day. On the eighth day, we will send you a link back to this site so that we can learn about your experience. We look forward to hearing back from you soon!
REFERENCES


Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the


