6-2017

AM Happy Scale: Reliability and Validity of a Single-Item Measure of Happiness

Christina P. Moldovan

LOMA LINDA UNIVERSITY
School of Behavioral Health
in conjunction with the
Faculty of Graduate Studies

_____

AM Happy Scale: Reliability and Validity of a Single-Item Measure of Happiness

by

Christina P. Moldovan

_____

A Dissertation submitted in partial satisfaction of
the requirements for the degree
Doctor of Philosophy in Clinical Psychology

_____

June 2017

Each person whose signature appears below certifies that this dissertation in his/her opinion is adequate, in scope and quality, as a dissertation for the degree Doctor of Philosophy.

_____, Chairperson

Adam Aréchiga, Associate Professor of Psychology

_____

Lee Berk, Professor of Allied Health Studies, School of Allied Health

_____

Kendal Boyd, Associate Professor of Psychology

_____

Holly Morrell, Associate Professor of Psychology

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Dr. Aréchiga, without whom I would not have been able to start or complete this project! Thank you for giving me the flyer on the Positive Psychology seminar that changed the course of my research and helped me to gain a better understanding of myself, my life, and my purpose. Also, thank you for inviting me into your lab and entrusting me with completing the WAHA study data collection, coming up with the idea of creating a happiness scale, and nurturing my interest in happiness by encouraging me to design and test the scale. You provided me with just the right amount of autonomy and guidance to help me grow and I will be forever thankful for the opportunity to work with you in creating this project.

I would also like to thank my committee members for their brilliant ideas, their kindness, their patience, and their willingness to join me on this wonderful journey. Each of you is special to me for different reasons. Dr. Morrell, you have been there from beginning to end, always willing to help with anything I needed, whether it was emotional support, statistical advice, professional guidance, or a laugh. Knowing that I could approach you about anything meant the world to me, and I know that I could not have matched for internship without your help. Dr. Boyd, you taught the class that made me fall in love with research and you were my first clinical supervisor. Your Rogerian warmth and unconditional positive regard were pivotal in helping me gain confidence in myself as a clinician and as a person. Additionally, teaching the CBT lab helped me to hone my theoretical orientation. Dr. Berk, working on the appetite study together was a humbling experience for me as I realized how fortunate I was to be surrounded by a table of geniuses each week pooling together knowledge from their respective fields to provide

a comprehensive understanding of obesity, chocolate, and laughter. I am thankful for your enthusiasm about this project and for the wonderful ideas that you contributed.

To my friends and family, I am thankful for your endless support through this endeavor. I know it has been difficult for you not to know exactly what I am going through and I appreciate you being there for me regardless. To my parents, I would not be where I am today if it was not for you risking your lives to bring us to America. You have provided us with everything we ever needed and more, and for that I am deeply grateful. Mom, thank you for your kind words of encouragement verbally and through your thoughtful cards. Dad, thank you for your words of wisdom and motivation. Sonia, you gave me praise and admiration, things a younger sister desperately craves, and inspired me to join you in changing the world. You and Edmund also gave me one of my greatest joys in life, Kolton. Skyping with you and him has made my day countless times.

To my friends, given that this dissertation is on happiness, you have shown to be one of the strongest predictors of my happiness ($p < .001$). Mer, Annie, Kimber, Elian, and Erika, I cherish what we have because nobody else that I know has a bond quite like ours. I never laugh as much as when I am with you and nobody knows me quite as well as you all do. Arleah and Martin, you are the greatest gifts that IHS has given to me. You guys know me inside and out, and still love me. Nickol, you made college fun and exciting and we will be forever connected by our brains. Vicky, Michelle, Elena, Imari, and Aman, you guys have been my rock throughout this grad school process and I am so thankful that we shared this amazing journey together. Tina, Adriana, Suma, Jessica, and the other close friends I've made along the way have all enriched my life in unique ways.

You all constantly inspire, support, and motivate me, and I am so grateful for your friendship.

I am also grateful for my past and current professors, co-workers, bosses, supervisors, and mentors, who have helped to shape me into the human being I am today. Dr. Peters, Dr. Staack, and Dr. Testerman, the three of you have a very special place in my heart and your mentorship has been invaluable. The other people in my life, such as my neighbors, the friendly Trader Joe's and Clark's employees, my manicurist, the Post Office employee who checks in on my progress and frequently congratulates me for my hard work, and people who I have volunteered with also made the graduate school experience more amazing. It truly takes a village, and if I have forgotten to mention somebody here who has positively impacted me, I apologize and I thank you for being a part of my life.

CONTENT

TABLES

Tables Page

ABSTRACT OF THE DISSERTATION

AM Happy Scale: Reliability and Validity of a Single-Item Measure of Happiness

by

Christina P. Moldovan

Doctor of Philosophy, Graduate Program in Clinical Psychology
Loma Linda University, June 2017
Dr. Adam Aréchiga, Chairperson

Research on happiness has been abundant over the last few decades and findings have repeatedly shown that people who are happier have better life outcomes in terms of health and in nearly all other aspects of functioning. Thus, it is crucial to continue studying the construct using valid and reliable measurements. Single-item rating scales have been shown to be psychometrically sound and more convenient in comparison to multiple-item measures designed to measure certain constructs. The aim of the present study was to test the reliability and convergent and divergent validity of a single-item happiness scale, the Aréchiga-Moldovan Happy Scale (AM Happy Scale). Participants included 275 adults between the ages of 64 to 81 (M = 71.51, SD = 3.85; 63.6% female; 77.1% White, 10.5% Hispanic, 6.2% Black, and 4.7% Asian) recruited from Loma Linda, California and the surrounding communities. This population was chosen due to progressive increases in life expectancy and rates of older adults in the workforce that highlight the importance of prevention and maintenance of health. The overall range for the minimum reliability estimate of the AM Happy Scale was .27 to 1.06. The AM Happy Scale showed a convergent relationship with the Positive Affect subscale of the Positive and Negative Affect Schedule (PANAS; r_s = .57, p < .001) and the Spirituality Index of

Well-Being (r_s = .50, p < .001). The AM Happy Scale showed a divergent relationship with the Negative Affect subscale of the PANAS (r_s = -.38, p < .001) and the Patient Health Questionnaire (r_s = -.42, p < .001). Additionally, the AM Happy Scale also demonstrated a positive relationship with the 12-item Short-Form Health Survey, a measure of mental (r_s = .51, p < .001) and physical health (r_s = .26, p < .001). Findings from this study have serious clinical implications indicating that a brief measure of happiness may be a quick and efficient way to assess an individual's overall sense of well-being, which is also associated with his or her physical health.

CHAPTER ONE

INTRODUCTION

Positive psychology gained prominence when Martin Seligman presented on the topic as the theme of his inaugural address as president of the American Psychological Association in 1998. Since that time, positive psychology research has been on the rise, with multiple domains of positive psychology being studied over the last several decades. Specifically, research on happiness and related constructs, including well-being and life satisfaction has been abundant (e.g., Hooker & Siegler, 1993; Knight, Song, & Gunatilaka, 2009; Myers & Diener, 1995). Results of such research revealed that people who are happier tend to have better relationships, physical and mental health, financial success, and more effective coping strategies (Lyubomirsky, King, & Diener, 2005). Further, in a thorough review of the literature, Lyubomirsky and colleagues (2005) showed that happiness precedes these positive outcomes rather than being a result of them. Clearly, happiness is an important construct that merits further investigation; thus, the purpose of the current study is to examine the accuracy of a single-item scale of happiness, the Aréchiga-Moldovan Happy Scale (AM Happy Scale), in a population of older adults.

The current study focuses on the older adult population because this population is projected to more than double in size by the year 2060 (Colby & Ortman, 2015). As the rates of older adults increase, the life expectancy in the United States and the rates of older adults in the workforce are also rising, emphasizing the importance of physical and cognitive health maintenance in later life (Centers for Disease Control and Prevention [CDC], 2012; Colby & Ortman, 2015; U.S. Social Security Administration [SSA], 2013).

Moreover, as the older population is increasing, rates of depression in this population are also projected to increase (Jeste et al., 1999). While depression has shown to be comorbid with many medical illnesses, mental illness is still greatly stigmatized among older adults and it often goes unreported and undetected, leading to potentially fatal results (Blazer, 2003; Rodda, Walker, & Carter, 2011). Happiness has also been linked to health outcomes, and screening for happiness may eliminate some of the stigma associated with mental illness while providing valuable information about an individual's mental and physical well-being.

CHAPTER TWO

BACKGROUND

Effectiveness of Single-Item Measures

Single-item scales have been shown to be practical and psychometrically sound instruments for assessing life satisfaction (Campbell, Converse, & Rodgers, 1976; Palmore & Kivett, 1977), affect (Russel, Weiss, & Mendelsohn, 1989), subjective well-being (Sandvik, Diener, & Seidlitz, 1993), interpersonal relationships (Aron, Aron, & Smollan, 1992), attachment (Hazan & Shaver, 1987), self-esteem (Carpenter, 1996; Robins, Hendin, & Trzesniewski, 2001), and socioeconomic status (Operario, Adler, & Williams, 2004). In 1965, Cantril used a self-anchoring scale to measure happiness of residents of kibbutzim (collective utopian communities in Israel traditionally based on agriculture; Goldenberg & Wekerle, 1972) by asking participants the question: "Here is a picture of a ladder. Suppose we say that the top of the ladder represents the best possible life for you and the bottom represents the worst possible life for you. Where on the ladder do you feel you personally stand at the present time? Ten being the best possible life and zero being the worst possible life." Adaptations of Cantril's scale have been used to measure happiness and quality of life in multiple studies (e.g., Borge, Martinsen, Ruud, Watne, & Friis, 1999; Kahneman & Deaton, 2010; Sawatzky, Ratner, Johnson, Kopec, & Zumbo, 2010).

Other validated single-item measurement scales of happiness or the like include the Delighted-Terrible Scale (Andrews & Withey, 1976), Fordyce Emotion Questionnaire (Fordyce, 1988), Gurin Scale (Gurin, Veroff, & Feld, 1960), and most recently, the Single-Item Measurement of Happiness (Abdel-Khalek, 2006). The Delighted-Terrible

Scale (Andrews & Withey, 1976) asks participants how they feel about different aspects of their life, including their life as a whole and their current happiness level, and provides them with seven mood adjectives as responses, one representing "terrible" and seven representing "delighted." The Fordyce Emotion Questionnaire (Fordyce, 1988) asks how happy or unhappy an individual usually feels in general. This question has 11 response choices that are graphically anchored with a series of mood adjectives with the highest choice being "feeling extremely happy, ecstatic, joyous and fantastic" and the lowest choice being "feeling extremely unhappy, utterly depressed, completely down." The second portion of Forsythe's (1988) questionnaire asks what percentage of the time a person feels happy, unhappy, or neutral. The Gurin Scale (Gurin et al., 1960) asks participants "taking all things together, how would you say things are these days?". Response choices include "very happy," "pretty happy," and "not too happy." Lastly, the Single-Item Measurement of Happiness (Abdel-Khalek, 2006) was designed to measure happiness in the Arab culture. The scale consists of one question "do you feel happy in general?" and participants respond by circling a number from 0 to 10 on a horizontal line with instructions indicating that zero is the minimum score and 10 is the maximum score.

Although the terms "happiness" and "life satisfaction" have been used interchangeably to reflect a state of well-being, researchers have proposed that along with positive beliefs about life and positive emotions, life satisfaction is simply one of the many components encompassed by the broader term, "happiness" (Diener, Lucas, & Scollon, 2006; Lucas, Diener, & Suh, 1996). Moreover, Campbell and colleagues (1976) have noted that a term like "happiness" appears to elicit "an absolute emotional state" whereas "satisfaction" elicits a more "cognitive judgment of a situation laid against

external standards of comparison," (p. 31) suggesting that responses to "satisfaction" questions may be more relative compared to responses to "happiness" questions. Cummins (1998) explains that measures of happiness and satisfaction seem to share a maximum of 50% to 60% common variance, with considerably lower values for some population subgroups, and argues that it is useful to measure and analyze them separately.

Thus, if "happiness" and "satisfaction" are separate constructs that may elicit different responses based on the terminology or if life satisfaction is just one facet of happiness, it is possible that the term "best possible life" used by Cantril to gauge life satisfaction may not yield the same results as would using the term "happiness." In the current study, we have modified Cantril's scale in several ways, one of which is using the term "happiness" rather than "satisfaction with life" to assess an overall sense of well-being, and providing some structure to the relative aspects of the scale by asking participants to rate themselves in comparison to other people in the United States. Participants were asked to rate themselves in comparison to other people in the United States rather than to other people they know because research has shown that although individual demographics are not strong predictors of happiness, happiness does vary among nations (Myers & Diener, 1995). In the current study, the goal was to ask participants to rank themselves taking into account a broader perspective when considering their happiness levels outside of their immediate environment.

Defining Happiness

There is no concrete established definition for happiness, although the topic has

been extensively debated by philosophers and researchers alike (Diener, Scollon, & Lucas, 2003). Bradburn (1969) defined happiness in terms of the extent to which positive feelings outweigh negative feelings. Tatarkiewicz (1976) perceived happiness as having total satisfaction with life. According to Diener (2000), happiness has a cognitive and affective component and is based on subjective evaluations of one's life. Despite the subjective nature of happiness and its various definitions, Kahneman and Krueger (2006) have examined the validity of self-reported happiness and found that it correlates with numerous objective and subjective indicators of well-being.

One such indicator of well-being and established correlate of happiness is spirituality. Spirituality has been defined as the search for the sacred by Pargament and Mahoney (2002). The researchers posit that there are multiple pathways that individuals can take in attempting to discover and preserve the sacred. These pathways may include traditional organized religions, newer spirituality movements, or individualized worldviews. In order to measure spiritual well-being, researchers have created a spiritual quality-of-life measure, the Spirituality Index of Well-Being (SIWB; Daaleman, Frey, Wallace, & Studenski, 2002). The creators of the scale and other researchers have shown that general happiness is strongly linked to spiritual well-being (Ai, Tice, Peterson, & Huang, 2005; Ciarrocchi & Brelsford, 2009; Daaleman, Perera, & Studenski, 2004; Holder, Coleman, & Wallace, 2010).

For instance, Holder and colleagues (2010) found that children's spirituality, but not their religious practices (e.g., attending church, praying, and meditating), had a strong positive relationship with their levels of happiness. Several measures of spirituality and happiness were administered in this study, and depending on the measure, spirituality

accounted for 3% to 26% of the unique variance in children's happiness. Ai and colleagues (2005) found that spiritual support, a type of support derived from an individual's faith in coping, was highly correlated with optimism and had a positive impact on affect by decreasing emotional distress. Moreover, in a study examining the impact of spirituality and religion on substance coping (relying on substances to improve mood), researchers found that higher levels of spirituality were related to higher levels of positive affect and an increased satisfaction with life (Ciarrocchi & Brelsford, 2009).

In addition, spirituality has also been linked with self-reported physical health and successful ageing (Crowther, Parker, Achenbaum, Larimore, & Koenig, 2002; Koenig, George, & Titus, 2004). For instance, in a study of geriatric outpatients, researchers found that individuals who reported higher levels of spiritual well-being, measured by the SIWB, were more likely to appraise their health as good (Daaleman et al., 2004). Higher levels of spirituality continued to be associated with better self-reported health even when the researchers controlled for covariates, such as functional status and race. The researchers posit that spirituality might be useful in understanding the mechanisms behind the positive self-ratings of health. For example, older adults who are spiritual may have a more holistic perspective of health that focuses not only on symptoms, but on the role and impact of these symptoms within their greater life scheme. The SIWB was designed to measure an individual's life scheme, or the perception of one's life purpose, and was also designed to measure self-efficacy, another component that may impact how an older adult rates his or her physical health. Other research has also linked higher levels of spirituality to lower levels of mortality and hypertension among older adults (Krause et al., 2006; McCullough, Hoyt, Larson, Koenig, & Thoresen, 2000).

In a study examining traits of happy individuals, Myers and Diener (1995) reported that individuals who feel they have more control over their own lives tend to be happier. The researchers provided evidence showing that happy people who feel empowered rather than helpless tend to perform better in school, cope better with stress, and live happier lives (Campbell, 1981; Larsen, 1989). More recent studies reflect similar trends. In a sample of Muslim Arab college students, Abdel-Khalek and Lester (2017) found that participants who perceive themselves as religious are also more likely to perceive themselves as self-efficacious and to report greater levels of mental health and happiness**.** Another study using a sample of undergraduate and post-graduate students in India also showed significant positive correlations between self-efficacy and happiness (Hunagund & Hangal, 2014). In addition to predicting happiness, self-efficacy also predicts positive health behaviors and better health (Grembowski et al., 1993; McAuley et al., 2006). As previously stated, the SIWB was created to be a measure of spiritual well-being, but also touches on the domains of self-efficacy and life satisfaction, via its two subscales titled Self-Efficacy and Life Scheme.

Given that almost 95% of Americans believe in God or a higher power (Gallup & Lindsay, 1999) and the strong links between spirituality, happiness, and positive health outcomes, spirituality is an important construct to continue investigating. In fact, Gomez and Fisher (2003), who developed a different measure of spiritual well-being, argue that spiritual well-being is equally important to measure as physical, mental, and emotional well-being. Spirituality and religion are especially salient in the lives of the older adults, and particularly in medical settings where older adults may be faced with serious and life-threatening illnesses (Daaleman & VandeCreek, 2000; Koenig et al., 2004). Spirituality

and happiness appear to be intertwined; therefore, the intention of using the SIWB in the current study is to further test this relationship, as well as to examine relationships among happiness, life schemes, self-efficacy, and health in older adults.

Importance of Measuring Happiness

As previously stated, there has been an abundance of research on happiness over the past several decades due to the fact that happiness is related to positive life outcomes, making it an important topic to study globally (Diener et al., 2003). In fact, Diener (2000) found that happiness and life satisfaction were rated as well above neutral in terms of importance and regarded more important than money in a study of college students from 17 different countries. Diener (2000) also noted that even people living in relatively unhappy societies value happiness to some extent.

In terms of positive life outcomes, in a meta-analysis examining cross-sectional, experimental, and longitudinal data from 225 studies, Lyubomirsky and colleagues (2005) provided evidence showing that happy individuals are more successful across multiple life domains. The authors emphasize that this is not only because success leads to happiness, but because happy people tend to have certain traits that promote success. The meta-analysis revealed that happier people were more likely to have more fulfilling marriages and relationships, more friends and social support networks, better job opportunities, superior work performance, greater job satisfaction, higher incomes, and more community involvement.

People who expressed higher levels of positive emotions were also less likely to suffer from mental health problems, including depression, schizophrenia, anxiety, and

substance abuse issues (Lyubomirsky et al., 2005). Happier people also reported lower mortality rates, lower incidence of stroke, lower incidence of cardiovascular disease, fewer work absences, fewer hospital visits, smaller allergic reactions, and were less likely to get a cold, indicating that happiness is associated with better physical health outcomes. Further, positive emotions appeared to be beneficial in surviving crisis situations by buffering people against depression and fostering growth and resilience post-crisis. This brief summary includes only some of the benefits associated with happiness, highlighting the importance of continuing to assess this construct using psychometrically sound instruments.

The Short-Form Health Survey (SF-12; Ware, Kosinski, & Keller, 1996) is a commonly used psychometrically sound measure of physical and mental health that has been validated in multiple countries with various populations (e.g., Gandek et al., 1998; Lam, Tse, Gandek, 2005; Salyers, Bosworth, Swanson, Lamb-Pagone, & Osher, 2000; Sanderson & Andrews, 2002). The measure has been used in several studies designed to examine the impact of happiness on health (e.g., Perneger, Hudelson, & Bovier, 2004; Rowold, 2011; Takeyachi et al., 2003). In a study of Swiss college students, researchers found such a robust relationship between happiness and mental health, measured by the SF-12, that they recommended using happiness as a screening tool for mental health problems in university students (Perneger et al., 2004). The same study also found significant positive relationships among happiness, social support, and self-esteem.

A German version of the SF-12 was used in three different studies to measure psychological health and its relationship to happiness and four different types of spiritual well-being: personal, communal, environmental, and transcendental (Rowold, 2011).

These studies used a sample of convenience and two samples of university students. Results from all three studies indicated that psychological health was positively related to happiness and personal spiritual well-being. Other researchers who examined the relationships between self-reported health on the SF-12 and back pain in a Japanese sample found that higher levels of happiness correlated with better mental and physical health status, having a spouse, and being female (Takeyachi et al., 2003). Results of a study conducted with older adult participants in Alabama found a moderate positive correlation between happiness and self-reported health on the SF-12 (Angner, Ghandhi, Purvis, Amante, & Allison, 2013). Similar to these previous studies, the aim of the current study is to examine relationships among happiness and mental and physical health using the SF-12.

Older Adults and Happiness

The United States Census Bureau predicts that between 2014 and 2060, the U.S. population will increase from 319 million to 417 million (Colby & Ortman, 2015). Although the population is projected to grow more slowly in future decades than in the recent past, the percentage of the population that is aged 65 and over is estimated to more than double in size from 46 million to 98 million over this time period. The biggest increase for this population is expected to occur between 2020 to 2030, when the population aged 65 and over is projected to increase by 18 million (from 56 million to 74 million). The timing of this increase is related to the aging of the baby boom generation. The baby boomers began turning 65 in 2011 and by 2030 they all will be aged 65 and

older, with approximately 10,000 baby boomers turning 65 every day for the next 13 years (Cohn & Taylor, 2010).

Moreover, the life expectancy of the U.S. population has also been steadily increasing for adults 65 and over, with 19.3% of men surpassing the age of 65 in 2015 compared to 12.7% of men who surpassed the age of 65 in 1940 (SSA, 2013). Women also experienced an increase in life expectancy, with 21.6% of women surpassing the age of 65 in 2015 compared to 14.7% of women who surpassed the age of 65 in 1940. On average, a man reaching age 65 today can expect to live until age 84.3 and a woman turning age 65 today can expect to live until age 86.6 (SSA, 2016). Additionally, according to the SSA, approximately one of every four 65-year-olds today will live past age 90, and one of 10 will live past age 95.

As the population and life expectancy of older adults are increasing, the proportion of older adults in the work force is also on the rise. Prior to the year 2000, workers 65 years of age or older held more part-time than full-time positions (CDC, 2012). In 1995, 56% of older workers held part-time positions and 44% held full-time positions. Since then, there has been a steady increase in full-time positions for older workers, which reached a complete reversal in 2007 when 56% held full-time positions and 44% held part-time positions. The gap widened even further in 2011 when 77% of older workers were employed full-time while 23% were employed part-time.

The Bureau of Labor Statistics projects that the percent of older adults ages 65 and over in the workforce will also continue to rise from 26.8% in 2012 to over 31% in 2022 (U.S. Department of Labor, Bureau of Labor Statistics, 2013). The increase can be attributed to the increase in qualifying age to receive Social Security Benefits, changes in

eligibility for Social Security Benefits, need for health insurance, changes in the general economic climate, and the availability and design of employer-sponsored benefits, which tends to typically transfer greater responsibility to the retiree (CDC, 2012). Furthermore, older adults may continue working past retirement age to maintain their cognitive functioning in light of research findings showing that employees who retire in their early sixties have diminished cognitive ability compared with those who retire at or after age 65 (Rohwedder & Willis, 2010).

These changes in the U.S. population emphasize the importance of health maintenance in older adults. As previously mentioned, prior reports have indicated that happiness and related constructs are associated with positive health-related outcomes. For instance, individuals with higher levels of optimism or positive affect have shown to have better outcomes in terms of less pain and fewer symptoms, hospital visits, and sick days at work compared to their pessimistic counterparts (Achat, Kawachi, Spiro, DeMolles, & Sparrow, 2000; Cohen & Pressman, 2006; Levy, Lee, Bagley, & Lippman, 1988). Optimism was also associated with a lower incidence of stroke and cardiovascular disease, and a faster recovery after cardiac surgery (Kubzansky, Sparrow, Vokonas, & Kawachi, 2001; Ostir, Markides, Peek, & Goodwin, 2000; Scheier et al., 1989). Further, people who are more optimistic are more likely to attend to and remember potentially threatening health-relevant information more than those who are less optimistic (Aspinwall & Brunhart, 1996).

Steptoe and Wardle (2005) found a link between positive affect and physiological stress responses, indicating that greater happiness is associated with lower salivary cortisol, reduced fibrinogen stress responses, and lower ambulatory heart rate in men.

These effects were independent of age, socioeconomic status, smoking, body mass, and psychological distress. Findings remained consistent at a three year follow-up. In addition to protecting against disease, Ostir and colleagues (2000) found positive affect to be a significant predictor of functional independence. Further, a greater sense of well-being has been associated with increased longevity, resilience, and healthy aging (Benyamini, Idler, Leventhal, & Leventhal, 2000; Danner, Snowdon, & Friesen, 2001; Fredrickson, Tugade, Waugh, & Larkin, 2003; Levy, Slade, Kunkel, & Kasl, 2002).

Positive affect has frequently been measured by researchers using the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988). It is the second most widely used instrument in positive psychology and has been cited in the literature over 150 times (Ackerman, 2015). The scale contains adjectives endorsed by participants to reflect levels of both positive and negative affect, and has been used in multiple countries with clinical and non-clinical populations (e.g., Crawford & Henry, 2004; Reske et al., 2007). In addition to the benefits of positive affect identified above, results of studies using the PANAS to conduct happiness research have demonstrated significant relationships between positive affect and gratitude, happiness, well-being, life satisfaction, spirituality, creativity, and commitment to social justice (Cloninger & Zohar, 2011; Elam, 2000; Hervás & Vázquez, 2013; Lou et al., 2012; Pethtel & Chen, 2010; Powers, Cramer, & Grubka, 2007; Schütz et al., 2013; Watkins, Woodward, Stone, & Kolts, 2003).

In two different studies using college student populations, researchers identified a moderate to strong effect of positive affect on gratitude (Watkins et al., 2003). Additionally, these researchers found that an experiment designed to increase gratitude in

which participants were assigned to various conditions (thinking about a living person for whom they were grateful, writing an essay about such a person, or writing a letter to this person) was successful in increasing positive affect. Grateful thinking yielded the greatest increases in positive affect. Schütz and colleagues (2013) studied affective profiles created by responses on the PANAS and found that individuals with high positive affect and low negative affect (termed "self-fulfilling") had higher levels of happiness and life satisfaction and lower levels of depression compared to individuals with other profiles ("high affective" - high positive affect, high negative affect; "low affective" - low positive affect, low negative affect; and "self-destructive" - low positive affect, high negative affect). Another study examining character profiles in community residents over 40 years old in Israel found that positive affect had the strongest association with creativity (Cloninger & Zohar, 2011).

In a study of 825 Chinese elders, Lou and colleagues (2012) found that higher levels of positive affect were associated with higher levels of positive spiritual well-being, meaning of life, transcendence, and relationships with self, family, friends, others, and the environment. The relationship between positive affect and spirituality was demonstrated in another study in which the participants were African American cancer patients from Alabama (Holt et al., 2011). Results of this study indicated that higher levels of positive affect were associated with higher levels of spiritual well-being, sense of meaning, involvement in religious behaviors, and emotional health as measured by the SF-12. The same study identified positive affect as a mediator between religious behaviors and emotional functioning. The PANAS' negative affect subscale has typically been used in conjunction with the positive affect subscale and has shown to be a strong

predictor of life dissatisfaction, depression, and other negative mood states (e.g., Kercher, 1992; Pethtel & Chen, 2010; Powers et al., 2007).

In general, positive affect has shown to have positive benefits, whereas negative affect has shown to reduce an individual's functional status and health-related quality of life (Wilson & Cleary, 1995). Negative affect may also be related to chronic illness symptoms. For instance, Zautra and colleagues (1995) found that individuals with more pain and limitation from arthritis had higher levels of maladaptive coping, which was associated with lower positive affect and higher negative affect. Negative affect was also found to be a significant predictor of perceived health-related quality of life in individuals with chronic obstructive pulmonary disease (Hu & Meek, 2005). Furthermore, Koller and colleagues (1996) found that higher levels of negative affect were related to increased reports of somatic symptoms and higher levels of social stigma in cancer patients. Research has shown that health problems can provoke negative affect, which in turn may exacerbate disease and illness symptoms (Leventhal & Patrick-Miller, 2000; Watson & Pennebaker, 1989).

In addition to being associated with poorer physical health outcomes, negative affective states have been identified as a general predictor of psychiatric disorders and associated with increased anxiety and depression in patient populations (Watson, Clark, & Carey, 1988). Depressive symptoms are frequently comorbid with medical problems and have shown to increase the risk of mortality and morbidity in people with coronary heart diseases, diabetes, and other chronic illnesses (Carney, Freedland, Miller, & Jaffe, 2002; Eaton, Armenian, Gallo, Pratt, & Ford, 1996; Katon & Ciechanowski, 2002). However, patients suffering from depressive disorders often do not seek help for

psychological problems (Roness, Mykletun, & Dahl, 2005) and a World Health Organization study found that only 42% of primary care patients with major depression disorder were detected by the provider (Simon, Goldberg, Tiemens, & Ustun, 1999). Therefore, a key challenge in the health care system has been to identify depressive disorders early, which has been facilitated through the development of screening questionnaires such as the Patient Health Questionnaire (PHQ-9; Kroenke, Spitzer, & Williams, 2001).

The PHQ-9 has been studied and established as a valid and reliable measure for detecting depression in many different medical settings including primary care, hospital inpatient, and obstetrics-gynecology, and among persons diagnosed with arthritis, fibromyalgia, cancer, chronic pain, diabetes, epilepsy, substance abuse, and human immunodeficiency virus (Smarr & Keefer, 2011). The questionnaire has also been administered to postpartum women, older adults, and persons with physical and cognitive disabilities. Results of these studies have shown that the PHQ-9 is a reliable tool for predicting a diagnosis of depression in medical settings (Kroenke et al., 2001; Spitzer, Williams, Kroenke, Hornyak, McMurray, & Patient Health Questionnaire Obstetrics-Gynecology Study Group, 2000).

Researchers also found significant associations between the PHQ-9, used as a continuous variable, and multiple domains of the 20-item Short-Form Health Survey (SF-20), a longer version of the SF-12. These domains included mental health, health perceptions, social functioning, role functioning, physical functioning, and physical pain (Spitzer et al., 2000). Higher scores on the PHQ-9 predicted higher levels of functional impairment, disability days, and physician visits. In addition to being a useful screening

tool for depression in medical populations, the PHQ-9 has demonstrated good predictive validity in the general population (Martin, Rief, Klaiberg, & Braehler, 2006; Liu et al., 2016). Furthermore, the PHQ-9 has also shown to be a responsive and reliable measure of depression treatment outcomes (Löwe, Unützer, Callahan, Perkins, & Kroenke, 2004). Given its reputation as a psychometrically sound instrument, the PHQ-9 is used in the current study to measure levels of unhappiness in older adults.

Unfortunately older adults may not openly express depressive symptoms given the high levels of public stigma associated with mental illness endorsed among older adults with depression (Conner et al., 2010). This stigma also prevents older adults from seeking treatment for mental health problems (Conner et al., 2010). Hence, asking older adults to report their current level of happiness rather than depressive symptoms may reduce the stigma associated with mental health problems while providing valuable information about an individual's mental and physical health.

Clinical Implications

In addition to potentially reducing the stigma associated with reporting depressive symptoms, results from a study conducted in Switzerland led researchers to conclude that measuring happiness may help identify mental health care needs and that self-reported happiness may also be a useful outcome measure for evaluation of health interventions (Perneger et al., 2004). In a review of the research literature, Veenhoven (2008) examined longevity and happiness, based on the claim that longevity is the most objective measure of physical health. Results of the review showed that happiness is a strong predictor of longevity among healthy populations with an effect size comparable

to that of smoking, and thus serves as a protective factor against illness (Veenhoven, 2008). Veenhoven (2008) recommends that public policies aimed at increasing happiness become established to produce greater levels of happiness in a greater number of people. However, it is important to first determine proper measurements of happiness to inform the development of such policies (Stone & Mackie, 2014).

Limitations of Existing Research

To date, there are two kinds of measurements used to measure happiness and its related dimensions, multiple-item scales and single-item self-rating scales. Multiple-item scales typically contain a range of 10 to 30 items (Abdel-Khalek, 2006) and include examples such as the Authentic Happiness Inventory (Peterson, 2005), the Satisfaction with Life Scale (Diener, Emmons, Larsen, & Griffin, 1985), Subjective Happiness Scale (Lyubomirsky & Lepper, 1999), and the PANAS (Watson et al., 1988).

Single-item self-rating scales have also been used in research and may be more convenient to use than their multiple-item counterparts in terms of being less time and space consuming, and reducing the boredom, fatigue, and frustration that may result from answering similar questions repeatedly (Robins et al., 2001). Single-item self-rating scales can include a range of 2 to 200 choice points, but most commonly these single-item measures consist of a five- or seven-point Likert scale (Abdel-Khalek, 2006). Cummins and Gullone (2000) have criticized the sensitivity of the five- or seven-point Likert scales in regard to measuring subjective quality of life, which tend to produce data that are negatively skewed. The researchers argue that naming the Likert scale categories

compromises the interval nature of the derived data and makes it difficult to generate expanded choice formats.

In a review of the literature, Cummins and Gullone (2000) have found that expanding the choice-points beyond five- or seven-points increases scale sensitivity without affecting scale reliability. Consequently, they recommend that subjective quality of life be measured using ten-point (from one to ten), end-defined scales in order to maintain a continuous scoring system that is able to better detect small, clinically significant differences in comparison to a five- or seven-point bidimensional scale with a neutral mid-point. An increasing number of researchers have also used ten-point end-defined scales to measure constructs such as life satisfaction (Hooker & Siegler, 1993) and satisfaction with self (Watkins, et al., 1998).

Cummins and Gullone's (2000) argument can be applied to measuring happiness as well as quality of life given that research on happiness may also produce data that are negatively skewed (Bond & Lang, 2014) and because happiness is a construct that is perceived by people as continuous, rather than discrete (Bertrand & Mullainathan, 2001). Bond and Lang (2014) provide additional support for Cummins and Gullone's (2000) argument that ordinal scales, such as the five- or seven-point Likert scales, cannot be used to compare scores of happiness in order to draw conclusions about happiness among different groups of people given that the categories depicted in these scales are dependent on the researcher's notion of happiness, which may be arbitrary (Bertrand & Mullainathan, 2001). Thus, we have designed a ten-point, single-item, self-rating scale to measure happiness on a continuum to avoid the difficulties associated with using an ordinal scale or a five- or seven-point bidimensional scale with a neutral mid-point.

The Current Study

In summary, happiness has proven to be a valuable construct to examine given the numerous benefits associated with higher levels of well-being. Multiple-item measures can prove to be lengthy and time consuming; thus, it is beneficial to design an instrument that quickly and efficiently measures happiness. Although other single-item happiness scales have been used, to the best of our knowledge, none of them have used a ten-point end-defined scale, used the term "happiness," and added structure to the relative aspects of the scale by asking participants to rate themselves in comparison to other people in the United States.

The primary aim of the current study was to determine an overall range of reliability and to establish the convergent and divergent validity of a single-item happiness scale, the AM Happy Scale, designed to measure the happiness levels of older adults in the United States. It was hypothesized that the range of reliability of the AM Happy Scale will be between 0.6 and 0.9 given the reported reliability of other previously developed single-item measures (e.g., Abdel-Khalek, 2006; Dolbier, Webster, McCalister, Mallon, & Steinhardt, 2005). Further, it was hypothesized that the AM Happy Scale will have a convergent relationship with two established measures of happiness and will have a divergent relationship with two well-known measures of unhappiness. A secondary aim of the present study was to examine the relationship between the AM Happy Scale and a measure of mental and physical health. It was hypothesized that higher scores on the AM Happy Scale will correlate with higher levels of mental and physical health.

CHAPTER THREE

MATERIALS AND METHOD

Participants

Participants included 275 adults between the ages of 64 and 81 (*M* = 71.51, *SD* = 3.85) recruited from Loma Linda, California and the surrounding communities. The majority (*n* = 175; 63.6%) of participants were female. The sample was predominantly White (*n* = 212; 77.1%), and also consisted of 29 participants who identified as Hispanic (10.5%), 17 participants who identified as Black (6.2%), 13 participants who identified as Asian (4.7%), and 4 participants who identified as Other (1.5%). The majority of participants were married (*n* = 188; 68.4%), retired (*n* = 176; 64%), and declared themselves fully independent in terms of daily living (*n* = 222; 80.7%). Participants were highly educated (*M* = 15.65 years; *SD* = 2.46) and reported a mean of 7.03 (*SD* = 1.55) on the MacArthur Scale of Subjective Social Status designed to measure perceived socioeconomic status on a scale from 1 (lowest) to 10 (highest). Participant demographics are shown in Table 1.

The inclusion criteria for the current sample included being between ages 63 and 79 years, able to read and understand the English language, and able to come to Loma Linda University (LLU) every two months. Some participants exceed the age criterion because this project is a secondary data analysis from a larger, longitudinal randomized clinical trial on the effects of walnut consumption. The data used for the current analyses were collected during Phase 2 of the trial that took place 24 months after the initiation of the study. Exclusion criteria included having an inability to read or write, inability to undergo neuropsychological testing, and suffering from any major illnesses, including

neurodegenerative diseases (e.g., Parkinson's disease), chronic illnesses expected to shorten survival (e.g., obesity, heart failure, cancer), and psychiatric disorders (e.g., depression). A self-report health status questionnaire, brief cognitive screening tool (the Mini Mental State Examination; MMSE), and a review of candidate's medical history and recent blood work were used to determine eligibility. Participants with a MMSE score less than 24 were referred to a physician and were not eligible for the study. Additionally, individuals who were experiencing bereavement in the first year of loss were also excluded from the study.

Procedures

Participants were recruited from Loma Linda, California and surrounding areas via mailing study brochures or distributing them through the non-profit organization Institute of Aging, advertisements in the study centers, and word of mouth. Interested individuals attended an informational group meeting, completed a brief health status questionnaire, and signed an informed consent. Following this, candidates had a face-to-face interview with the study clinician, who assessed for potential compliance and reviewed the candidate's medical history, inclusion and exclusion criteria, recent blood work and use of medications or supplements, and administered the MMSE.

Eligible participants underwent neuropsychological and ophthalmologic examinations and completed several psychological measures. Participants presented to LLU every two months to be weighed and measured and to meet with the study dietitian. The study was approved by the Institutional Review Board at LLU. The participants received incentives that included a testing report summary of participants' scores on

neuropsychological measures, results of their eye exam, and varying amounts of walnuts as part of the study experimental manipulation and after the study was completed. A detailed account of the study protocol is documented elsewhere (Rajaram et al., 2017).

Scale Development

The AM Happy Scale (see Appendix A) was modeled after Cantril's self-anchoring ladder rating of life satisfaction (Cantril, 1965). However, the AM Happy Scale includes a ten-point scale rather than a ladder, and participants are asked to mark the scale to reflect their current levels of happiness in comparison to other people in the United States ranging from 1 (down in the dumps) to 10 (on top of the world). In order to make the scale more user-friendly, illustrations are included at the end-points that depict cartoon characters, one animated figure standing on a globe at the top to represent feeling on "top of the world," and on the bottom is another animated figure sticking its head out of a dumpster to reflect feeling "down in the dumps." The cartoon figure standing on the globe was found online at www.clipartof.com, and a license to use the image was purchased through the website for $15.00. The dumpster cartoon was also found online on a personal blog (http://systemconscienceme.wordpress.com/2013/06/25/my-take-on-dumpster-diving/). The author of the blog, who is also the artist of the dumpster cartoon, was contacted via electronic mail to request permission to use the image. The artist gave permission to use the image as long as a source description (the blog address) was included "somewhere within the documents that contain the illustration."

Measures

Demographic Questions

The demographic information collected from each participant included age, gender, race, marital status, employment status, socioeconomic status, religious beliefs, functional independence, and the number of years of education completed.

Positive and Negative Affect

The Positive and Negative Affect Schedule (Watson et al., 1988; see Appendix B) is a 20-item questionnaire designed to assess mood states over the past week. The measure consists of two subscales: Positive Affect (PA), which consists of ten adjectives designed to evaluate positive affect (e.g., excited, inspired, active) and Negative Affect (NA), which consists of ten adjectives designed to evaluate negative affect (e.g., guilty, hostile, afraid). Participants are asked indicate to what extent they have felt each adjective over the past week using a five-point Likert scale, ranging from *very slightly or not at all* to *extremely*. Low PA scores reflect "sadness and lethargy," whereas high PA scores reflect "high energy, full concentration, and pleasurable engagement" (Watson et al., 1988, p. 1063). Meanwhile, low NA scores describe "a state of calmness and serenity," whereas high NA scores suggest "subjective distress and unpleasurable engagement" (Watson et al., 1988, p. 1063).

Both PANAS subscales have shown to have good internal consistency, with Cronbach's alphas ranging from .86 to .90 for the PA and .84 to .87 for the NA in undergraduate college students, community-dwelling adults, and inpatient populations (Watson et al., 1988). The PANAS has also demonstrated good reliability in populations

of older adults, including older workers with a median age of 57 years (Cronbach's alphas of .87 and .89 for the PA and NA, respectively; Fletcher & Hansson, 1991); retirees over the age of 72 years (Cronbach's alphas of .75 and .81 for shortened five-item versions of the PA and NA, respectively; Kercher, 1992); older adults aged 70-100 living in Berlin, Germany (Cronbach's alphas of .78 and .81 for the PA and NA, respectively; Isaacowitz & Smith, 2003); and adults with chronic illnesses over the age of 60 (Cronbach's alphas of .86 and .83 for the PA and NA, respectively; Hu & Gruber, 2008).

In addition to being a reliable measure, studies indicate that the PANAS is also a valid measure of affect. Prior studies indicate that the scale has good discriminant validity, as demonstrated by low correlations between the PA and NA subscales in both extent and frequency response formats (*r*s = -.13 and -.28, respectively; Watson, 1988). The scale has also shown high convergent validity, as demonstrated by results of an exploratory factor analysis that included several alternative measures of positive and negative affect (Watson et al., 1988). The factor analysis revealed that the two subscales of the PANAS and the related measures clearly formed a two-factor structure that was consistent with the proposed factors of PA and NA. The PA and NA subscales were also significantly correlated with other scales measuring positive and negative affect, further indicating good convergent validity (*r*s = .76 to .92). The PANAS demonstrated good internal consistency in the present study, with Cronbach's alphas of .87 for the PA subscale and .85 for the NA subscale.

Spiritual Well-Being

The Spirituality Index of Well-Being (Daaleman et al., 2002; see Appendix C) is a 12-item questionnaire that measures an individual's perceptions of his or her spiritual quality of life. The measure consists of two subscales, Self-Efficacy (SE) and Life Scheme (LS). The SE subscale consists of items such as, "There is not much I can do to help myself," and "I can't begin to understand my problems." Sample items on the LS subscale include, "I have a lack of purpose in my life," and "There is a great void in my life at this time." Participants indicate their agreement with each statement using a five-point Likert scale, ranging from *strongly agree* to *strongly disagree*. Higher scores reflect a greater sense of spiritual well-being.

The scale and its subscales have shown good internal consistency in an adult outpatient population, with Cronbach's alphas of .86 for the SE subscale, .89 for the LE subscale, and .91 for the total scale (Daaleman et al., 2002). The scale also proved to be reliable in a sample of older, community-dwelling adults ages 65 and over (Cronbach's alpha of .87 for the total scale; Daaleman et al., 2004). Additionally, the SIWB was significantly correlated with the General Well-Being Scale ($r = 0.64$, $p < .001$) developed by McDowell and Newell (1996), indicating that it is an adequate measure of overall well-being (Daaleman & Frey, 2004). Furthermore, the scale also showed good divergent validity when correlated with the Geriatric Depression Scale ($r = -.35$) and good discriminant validity differentiating the SIWB from religiosity ($r = .12$, $p > .05$). The SIWB showed excellent internal consistency in the present study, with Cronbach's alphas of .91 and .92 for the SE and LS subscales, respectively, and .94 for the total scale.

Depression

The Patient Health Questionnaire (Kroenke et al., 2001; see Appendix D) is a nine-item criteria-based instrument for diagnosing depressive disorders. Participants indicate how often they have been bothered by problems such as, "Feeling down, depressed, or hopeless," and "Poor appetite or overeating" using a four-point Likert scale, ranging from *not at all* to *nearly every day*. Additionally, to facilitate diagnostic clarification, there is a question at the end of the scale (not included in the total score) that asks participants how much of an impact the symptoms they endorsed have on their home, work, or interpersonal functioning. The measure can be used as a categorical and a continuous measure, where higher scores reflect higher levels of depression and different ranges of scores represent different levels of severity (0-4 = minimal, 5-9 = mild, 10-14 = moderate, 15-19 = moderately severe, 20-27 = severe).

The PHQ-9 has demonstrated good diagnostic, criterion, construct, and external validity in two studies involving 3,000 patients in eight primary care clinics and 3,000 patients in seven obstetrics-gynecology clinics (Spitzer, Kroenke, Williams, & Patient Health Questionnaire Primary Care Study Group, 1999; Spitzer et al., 2000). Specifically, in terms of diagnostic validity, scores ≥ 10 had a sensitivity of 88% and a specificity of 88% for major depression (Kroenke et al., 2001). Criterion validity was established by examining sensitivity, specificity, and likelihood ratios for different PHQ-9 thresholds in 580 patients who were assessed by independent mental health professionals. The positive likelihood ratios of PHQ-9 scores of 0-4, 5-9, 10-14, 15-19, and 20-27 for major depression were 0.04, 0.5, 2.6, 8.4, and 36.8, respectively. Interpretation of these likelihood ratios means that, for example, a score in the 0-4 range is only 1/25 (0.04)

times as likely in a patient with major depression compared to a patient without major depression. Construct validity was determined by examining functional status (measured by the SF-20), disability days, symptom-related difficulty, and clinic visits over the PHQ-9 intervals using Analysis of Covariance. The PHQ-9 correlated most strongly with the mental health (.73), general health perceptions (.55), social functioning (.43), physical functioning (.37), and bodily pain (.33) subscales of the SF-20, $p < .05$. The PHQ-9 was also significantly correlated with disability days (.39), symptom-related difficulty (.55), and clinic visits (.24). Lastly, external validity was achieved by replicating findings from the 3,000 primary care patients in a second sample of 3,000 obstetrics-gynecology patients.

Within these same patient subgroups consisting of adults ages 18 and over, the PHQ-9 has also shown good test-retest reliability with a reliability coefficient of .84, and good internal consistency reliability with Cronbach's alphas of .86 (obstetrics-gynecology clinic) to .89 (primary care clinic; Kroenke et al., 2001). The PHQ-9 was also shown to have good diagnostic validity in older adults ages 65 years and over in a primary care setting (Phelan et al., 2010), and was shown to have good internal reliability in a sample of chronically ill patients over the ages of 59 years (Cronbach's alpha of .83; Lamers et al., 2008). In addition to being validated in medical settings, the PHQ-9 was also found to be a reliable instrument (Cronbach's alpha of .89) for detecting subthreshold depression in the general population (Martin et al., 2006). In the present study, the PHQ-9 also demonstrated adequate internal consistency with Cronbach's alpha of .79.

Physical and Mental Health

The Short-Form Health Survey (Ware et al., 1996; see Appendix E) is a 12-item survey of physical and mental health that provides two summary scores, one for each domain: Physical Component Summary (PCS) and Mental Component Summary (MCS). Sample items on the PCS include, "Does your health limit you in moderate activities such as moving a table, pushing a vacuum cleaner, bowling, or playing golf?" and "During the past 4 weeks, how much did pain interfere with your normal work?" The MCS includes items such as, "How much of the time during the past 4 weeks have you felt calm and peaceful?" and "How much of the time during the past 4 weeks have you felt downhearted and blue?" The SF-12 uses a combination of dichotomous questions (e.g., yes/no) and Likert scales ranging from three to six points to assess areas of physical functioning, role limitations due to physical health, bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional problems, and mental health. Higher scores indicate better health.

The SF-12 has shown adequate test-retest reliability in the general population, with reliability coefficients of .89 for the PCS and .76 for the MCS (Ware et al., 1996). Relative validity coefficients, measured with a known groups procedure, for the PCS ranged from .63 to .93 and for the MCS ranged from .03 to .11. The known groups procedure is a method of determining construct validity, and posits that test scores should discriminate across groups that theoretically are expected to be different on the trait measured (Hattie & Cooksey, 1984). In the process of validating the SF-12, comparisons were made between patient groups known to differ or to change in terms of the presence and seriousness of physical and mental conditions, acute symptoms, age and aging,

changes in health, and recovery from depression. The survey also demonstrated adequate internal consistency in a population of independent living older adults over the age of 65 years, with Cronbach's alpha of .84 for the PCS and .70 for the MCS, and can be utilized as either a predictor or an outcome measure (Resnick & Nahm, 2001). Both the PCS and MCS also demonstrated adequate to good internal consistency in the current study, with Cronbach's alphas of .81 and .73, respectively.

Data Analysis

Results of a power analysis using the G*Power 3 program (Faul, Erdfelder, Lang, & Buchner, 2007) indicated that 88 participants were needed in order to achieve sufficient power (.80) at an alpha level of .05 for the analyses. An effect size of .2 was used as recommended by Ferguson (2009) as the minimum effect size representing a "practically" significant effect for social science data for correlations (p. 533). All other analyses were performed using SPSS 20.0. Prior to conducting the main analyses, the data were tested for outliers, missing data, and violations of statistical assumptions.

Regression diagnostics were performed to evaluate the leverage, discrepancy, and influence of the data points in order to detect outliers. As a result, 28 outliers were detected and excluded from the analyses. A missing data analysis revealed that 11 participants were missing 100% of data on all variables of interest. Additionally, one participant was missing 62.5% of the data. These 12 participants were excluded from the study. Given the small percentage of missing data (less than 5%) pertaining to each variable of interest and the large sample size, listwise deletion was used to handle the remaining missing data (Allison, 2001).

Next, descriptive statistics were computed for all variables included in the study. The internal consistency of each scale was computed to assure that the scales met the minimum standards for reliability for research purposes (Cronbach's alpha of .70 or above; Furr & Bacharach, 2008), with the exception of the single-item happiness measure, for which internal consistency could not be computed. However, the minimum reliability of the AM Happy Scale was computed as shown by Dolbier and colleagues (2005) and described by Wanous, Reichers, and Hudy (1997) by using the correction for attenuation formula provided by Nunnally and Bernstein (1994): $\hat{r}_{xy} = \frac{r_{xy}}{\sqrt{r_{xx} \cdot r_{yy}}}$

In this formula, r_{xy} = the correlation between variables x and y, r_{xx} = reliability of variable x, r_{yy} = reliability of variable y, and $\hat{r}_{xy}$ = the assumed true underlying correlation between x and y if both were measured perfectly. Although the formula is typically used in situations where x and y represent different constructs, it can also be applied when both variables measure the same construct. In this situation, Nunnally (1978) wrote that the correlation between two such tests would be expected to equal the product of the terms in the denominator and consequently $\hat{r}_{xy}$ would equal 1.00…if $\hat{r}_{xy}$ were 1.00, r_{xy} would be limited only by the reliabilities of the two tests: $r_{xy} = \sqrt{r_{xx}} \cdot \sqrt{r_{yy}}$ (p. 220). Therefore, the minimum reliability of the AM Happy Scale was computed using the attenuation formula provided above, where x represented the AM Happy Scale and y represented each measure intended to converge with the AM Happy Scale (PANAS PA and SIWB Total). Knowing r_{xy} and r_{yy}, and assuming $\hat{r}_{xy}$ = 1.0 and more conservatively .90, as done by Dolbier and colleagues (2005) allowed us to solve for r_{xx} or the minimum reliability of x. The more conservative estimate of .90 for the underlying correlation between x and y provided a higher minimum reliability

estimate. We also used the sample correlation between x and y to estimate the true correlation between x and y. Additionally, we examined correlations from previous studies comparing the PANAS PA and SIWB to other happiness scales in order to determine an estimate of the strength of the relationships between these scales and the AM Happy Scale. These three methods of estimation yielded an overall range of reliability for the AM Happy Scale.

To evaluate the convergent and divergent validity of the AM Happy Scale, correlational analyses were used to correlate the scale with similar constructs that have already been established (Furr & Bacharach, 2008). Originally, Pearson's product-moment correlation coefficients were to be examined and interpreted in terms of their conceptual logic. Assumptions of Pearson's correlation state that the data must be interval in order to have an accurate measurement of the linear relationship between two variables (Field, 2009). Additionally, to determine if a correlation coefficient is significant, the data must be normally distributed. The data used were interval. To test for normality, skewness and kurtosis values were converted to z-scores by dividing them by their standard error (Field, 2009). Scores greater than 1.96 indicate that the data is significantly skewed ($p < .05$). These calculations and examination of Normal Q-Q plots revealed that the data were skewed. Results from the Kolmogorov-Smirnov test also revealed that the data were significantly non-normal, AM Happy Scale, $D(238) = .18$, $p < .001$; PANAS PA, $D(238) = .07$, $p < .01$; PANAS NA, $D(238) = .19$, $p < .001$; SIWB SE, $D(238) = .29$, $p < .001$; SIWB LS, $D(238) = .29$, $p < .001$; SIWB Total, $D(238) = .24$, $p < .001$; PHQ9, $D(238) = .23$, $p < .001$; PCS, $D(238) = .18$, $p < .001$; MCS, $D(238) = .19$, p

< .001. The majority of the variables had a negative skew and two of the variables (PANAS NA and PHQ-9) had a positive skew.

Reverse score transformations were conducted for the variables with a negative skew. Then, a log transformation, a square root transformation, and a reciprocal transformation were conducted on all the variables of interest in an attempt to correct the problems with normality (Field, 2009). Although each method was successful and resulted in a normal distribution for some of the variables, approximately half of the variables continued to have distributional problems. Thus, Spearman's correlation coefficient was used to conduct the main analyses. Spearman's coefficient does not rely on the assumptions of a parametric test, and uses ranked data to determine the significance of a relationship between two variables (Field, 2009).

Assumptions of Spearman's coefficient state that variables are measured on an ordinal**,** interval or ratio scale and that there is a monotonic relationship between the two variables that exists when either the variables increase in value together, or as one variable value increases, the other variable value decreases (Hauke & Kossowski, 2011). All variables were measured on an ordinal, interval, or ratio scale and scatterplots revealed that there were monotonic relationships between the variables. Spearman's correlation coefficient was also used to determine the relationships between the AM Happy Scale and the measure of mental and physical health. The correlations were squared, creating a coefficient of determination or R^2, in order to determine the amount of variability in one variable that is shared by another. Knowing how much variance is accounted for in one variable by another can provide useful information about the effect size of a correlation (Field, 2009). We used minimum effect size cutoffs provided by

Ferguson (2009) to determine the size of the effects (for R^2, minimum = .04, moderate = .25, strong = .64; for *r*, minimum = .2, moderate = .5, strong = .8).

CHAPTER FOUR

RESULTS

Participant demographic information is provided in Table 1. Mann-Whitney U-tests were used to test for gender differences in the study variables. Results indicated that the AM Happy Scale was the only variable that differed significantly between males (M = 8.21, SD = .94) and females (M = 8.42, SD = 1.10), U = 5382.50, z = -2.82, $p < .01$, r = .18. The effect size calculated for this analysis as shown by Rosenthal (1991) indicates a minimal effect size of these differences. Due to the multiple analyses, a Bonferroni correction was implemented using a p-value of .006 (.05/8).

Reliability

In order to determine an overall range of reliability for the AM Happy Scale, we followed a method outlined by Dolbier and colleagues (2005) and described by Wanous and colleagues (1997) by using the correction for attenuation formula provided by Nunnally and Bernstein (1994). In the attenuation formula provided above, x represented the AM Happy Scale and y represented each measure intended to converge with the AM Happy Scale (PANAS PA and SIWB). This formula allowed us to compute the minimum reliability of the AM Happy Scale (r_{xx}) by using the correlation between the AM Happy Scale and the PANAS PA and SIWB (r_{xy}), the reliability of the PANAS PA and SIWB (r_{yy}) scales, and a number that we estimated to be the true underlying correlation between the AM Happy Scale and the other measures of happiness if both were measured perfectly ($\hat{r}_{xy}$). We used four different types of values for $\hat{r}_{xy}$ in order to obtain a range of reliability estimates for the AM Happy Scale and we calculated r_{xx} eight different times (four times for each scale, PANAS PA and SIWB).

Table 1. Participant characteristics.

Variable	N (%)
Age	
64-69	102 (37.1)
70-75	123 (44.7)
76-81	50 (18.2)
Gender	
Female	175 (63.6)
Male	100 (36.4)
Race	
White	212 (77.1)
Hispanic	29 (10.5)
Black	17 (6.2)
Asian	13 (4.7)
Other	4 (1.5)
Marital Status	
Married	188 (68.4)
Widowed	30 (10.9)
Divorced	29 (10.5)
Single	10 (3.6)
Separated	1 (0.4)
Would rather not answer	1 (0.4)
Years of Education	
8-12	31 (11.3)
13-14	75 (27.3)
15-16	69 (25.1)
17-18	68 (24.7)
>19	32 (11.6)
Religious Denomination	
Yes	216 (78.5)
No	43 (15.6)

Table 1. Participant characteristics. *Continued*

Religious Belief	
Protestant	69 (25.1)
Roman Catholic	53 (19.3)
Seventh-Day Adventist	31 (11.3)
Baptist	23 (8.4)
Jew	5 (1.8)
Other	39 (14.2)
Retired	
Yes	176 (64.0)
No	83 (30.2)
Annual Combined Household Income	
Less than $15,000	5 (1.8)
$15,000 to $30,000	9 (3.3)
$30,000 to $50,000	41 (14.9)
$50,000 to $80,000	54 (19.6)
$80,000 to $120,000	40 (14.5)
$120,000 to $200,000	9 (3.3)
Would rather not answer	18 (6.5)
Functional Independence	
Completely independent	222 (80.7)
Partly independent	13 (4.7)
Provide care to dependent (full time or part time)	21 (7.6)

First, we computed minimum reliability using the value "1" for $\hat{r}_{xy}$, which assumed there was a perfect relationship between the Happy Scale and the other two scales of happiness. These calculations yielded reliability estimates for the AM Happy Scale of .37 (PANAS PA) and .27 (SIWB). Then, we used a more conservative estimate of .90 for the underlying correlation between x and y, which provided slightly higher

minimum reliability estimates of .46 (PANAS PA) and .33 (SIWB). We then rounded the values of the sample correlations calculated in the present study between the AM Happy Scale and the other two happiness scales and used these values (PANAS PA = .60 and SIWB = .50) for $\hat{r}_{xy}$. These calculations provided much higher minimum reliability estimates for the AM Happy Scale of 1.04 (PANAS PA) and 1.06 (SIWB).

Lastly, we examined correlations between the PANAS PA and SIWB and other happiness measures and calculated average correlations from these observed values to use in the correction for attenuation formula to represent $\hat{r}_{xy}$. For the PANAS PA, the average correlation between the subscale and other measures of positive affect was .83 (Watson et al., 1988). The SIWB was compared to another happiness scale in a single study in which their relationship had a correlation of .64 (Daaleman & Frey, 2004). Using these values for $\hat{r}_{xy}$ resulted in reliability estimates for the AM Happy Scale of .54 (PANAS PA) and .65 (SIWB). Overall, the eight equations revealed that the AM Happy Scale has a range of minimum reliability that is between .27 and 1.06. Results are shown in Table 2.

Table 2. Reliability estimates for AM Happy Scale.

	AM Happy Scale Reliability
PANAS PA	r_{xx} = .37 when $\hat{r}_{xy}$ = 1.00
	r_{xx} = .46 when $\hat{r}_{xy}$ = 0.90
	r_{xx} = 1.04 when $\hat{r}_{xy}$ = 0.60[a]
	r_{xx} = .54 when $\hat{r}_{xy}$ = 0.83[b]
SIWB Total	r_{xx} = .27 when $\hat{r}_{xy}$ = 1.00
	r_{xx} = .33 when $\hat{r}_{xy}$ = 0.90
	r_{xx} = 1.06 when $\hat{r}_{xy}$ = 0.50[c]
	r_{xx} = .65 when $\hat{r}_{xy}$ = 0.64[d]

Note. PANAS PA= Positive and Negative Affect Schedule Positive Affect subscale, SIWB = Spirituality Index of Well-Being. [a]Value chosen based on the sample correlation between the PANAS PA and the AM Happy Scale. [b]Value chosen based on an average value of correlation coefficients reported in other studies between the PANAS PA and other measures of happiness. [c]Value chosen based on the sample correlation between the SIWB Total score and the AM Happy Scale. [d]Value chosen based on correlation coefficient reported in another study between the SIWB and another measure of happiness.

Convergent and Divergent Validity

In order to test the convergent and divergent validity of the AM Happy Scale, Spearman's correlation coefficients were calculated, and a Bonferroni correction was implemented using a *p*-value of .006 (.05/8). Results provided support for our hypothesis indicating that the AM Happy Scale was convergent with two measures of happiness and divergent with two measures of unhappiness. Table 3 provides means and standard deviations, Mann-Whitney U-test results, possible range of scores for study variables, results of the correlations, and values for the coefficient of determination.

Specifically, the AM Happy Scale indicated a significant convergent relationship with the PANAS PA, r_s = .57; the SIWB SE, r_s = .45; the SIWB LS, r_s = .49; and the SIWB Total Scale, r_s = .50 (all *p*s < .001). Coefficients of determination (R^2) were

calculated in order to measure the amount of variability in one variable that is shared by another. Results indicate that the proportion of variance in the ranks that the PANAS PA and the AM Happy Scale share was 31.9%, which indicates a moderate effect. The proportion of variance in the ranks of the AM Happy Scale accounted for by the ranks in the SIWB SE was 20.2%, which indicates a minimal effect. The proportion of variance in the ranks of the AM Happy Scale accounted for by the ranks in the SIWB LS was 24.1%, which indicates a minimal to moderate effect. The proportion of variance in the ranks of the AM Happy Scale accounted for by the ranks in the SIWB Total Scale was 25.0%, which indicates a moderate effect.

The AM Happy Scale showed a divergent relationship with the PANAS NA (r_s = -.38) and the PHQ-9 (r_s = -.42; both *p*s < .001). The proportion of variance in the ranks of the AM Happy Scale accounted for by the ranks in the PANAS NA was 14.3%, which indicates a minimal effect. The proportion of variance in the ranks of the AM Happy Scale accounted for by the ranks in the PHQ-9 was 17.8%, which indicates a minimal effect.

Physical and Mental Health

As predicted, the AM Happy Scale showed a positive relationship with a measure of physical health, PCS (r_s = .26) and a measure of mental health, MCS (r_s = .51; both *p*s < .001). The proportion of variance in the ranks of the AM Happy Scale accounted for by the ranks in the PCS was 6.5%, which indicates a minimal effect. The proportion of variance in the ranks of the AM Happy Scale accounted for by the ranks in the MCS was 26.4%, which indicates a moderate effect.

Table 3. Means and standard deviations, Mann-Whitney U-test results, possible range of scores, correlations, and coefficients of determination.

	M (*SD*) Total Sample	*M* (*SD*) Female	*M* (*SD*) Male	*U* (*z*-score)	Possible Range of Scores	r_s AM Happy Scale	R^2
AM Happy Scale	8.34 (1.05)	8.42 (1.10)	8.21 (.94)	5382.50* (-2.82)	1-10	1.00	1.00
Convergent validity							
PANAS PA	36.84 (6.66)	37.38 (6.65)	35.94 (6.59)	5554.00 (-2.38)	0-50	.57**	.319
SIWB SE	4.57 (0.63)	4.61 (0.57)	4.50 (.70)	6155.00 (-1.34)	1-5	.45**	.202
SIWB LS	4.57 (0.66)	4.59 (.64)	4.54 (.71)	6236.50 (-1.17)	1-5	.49**	.241
SIWB Total	4.57 (0.60)	4.60 (.55)	4.52 (.67)	6120.00 (-1.35)	1-5	.50**	.250
Divergent validity							
PANAS NA	14.00 (4.64)	14.08 (4.85)	13.86 (4.29)	6358.00 (-.84)	0-50	-.38**	.143
PHQ-9	2.00 (2.62)	2.25 (2.76)	1.59 (2.31)	5991.50 (-1.60)	0-27	-.42**	.178
Relation to health							
SF-12 PCS	48.40 (8.98)	48.28 (8.98)	48.62 (9.00)	6657.00 (-.26)	0-100	.26**	.065
SF-12 MCS	53.41 (7.21)	53.08 (7.68)	54.00 (6.29)	6275.00 (-1.00)	0-100	.51**	.264

Note. PANAS = Positive and Negative Affect Schedule, PA = Positive Affect, NA = Negative Affect, SIWB = Spirituality Index of Well-Being, SE = Self-Efficacy, LS = Life Scheme, PHQ-9 = Patient Health Questionnaire, SF-12 = Short-Form Health Survey, PCS = Physical Component Summary, MCS = Mental Component Summary.
*$p < .01$. **$p < .001$.

Post-Hoc Analyses

Post-hoc analyses were conducted in order to obtain more detailed information about the relationships between the AM Happy Scale and the other variables. First, we conducted a point-biserial correlation analysis in order to determine the relationship between the AM Happy Scale and the PHQ-9 as a dichotomous variable in which we compared people with no depressive symptoms (scores of 0) to people with some depressive symptoms (scores > 0). Results indicated that the PHQ-9 was significantly related to the AM Happy Scale, $r_{pb} = -.34$, $p < .001$. However, only 11.6% of the variance in the AM Happy Scale was accounted by the dichotomous PHQ-9 variable, which indicates a minimal effect.

Next, we conducted a series of correlational analyses to examine relationships between individual items in each scale and the AM Happy Scale. Given that the data are skewed, the data were ranked and Spearman's correlation coefficient was calculated once again. Results of correlational analyses are presented in Table 4. The alpha level was set at .01 to correct for Type I error associated with running multiple correlational analyses. Given that these analyses were purely exploratory, no hypotheses were made; however, it was assumed that items on the scales measuring happiness will most likely converge with the AM Happy Scale, and items on the scales measuring unhappiness will diverge with the AM Happy Scale.

Table 4. Correlations between the AM Happy Scale and individual items.

	Correlations for the AM Happy Scale (r_s)	Coefficient of Determination (R^2)
Positive and Negative Affect Schedule		
Interested	.40**	.157
Determined	.40**	.156
Alert	.35**	.123
Attentive	.34**	.116
Excited	.30**	.092
Enthusiastic	.45**	.204
Active	.38**	.144
Inspired	.44**	.193
Strong	.37**	.139
Proud	.33**	.110
Guilty	-.24**	.055
Irritable	-.27**	.072
Distressed	-.28**	.081
Scared	-.18*	.033
Hostile	-.20*	.042
Ashamed	-.18*	.033
Jittery	-.17*	.028
Upset	-.33**	.106
Nervous	-.19*	.037
Afraid	-.15	.022
Spirituality Index of Well-Being		
1. There is not much I can do to help myself	.34**	.115
2. Often, there is no way I can complete what I started	.32**	.101
3. I can't begin to understand my problems	.30**	.087

4. I am overwhelmed when I have personal difficulties and problems	.37**	.135
5. I don't know how to begin to solve my problems	.39**	.154
6. There is not much I can do to make a difference in my life	.43**	.181
7. I haven't found my life's purpose yet	.41**	.170
8. I don't know who I am, where I came from, or where I am going	.39**	.150
9. I have a lack of purpose in my life	.40**	.160
10. In this world, I don't know where I fit in	.37**	.137
11. I am far from understanding the meaning of life	.37**	.135
12. There is a great void in my life at this time	.44**	.193
Patient Health Questionnaire		
1. Little interest or pleasure in doing things	-.36**	.128
2. Feeling down, depressed, or hopeless	-.36**	.130
3. Trouble falling or staying asleep, or sleeping too much	-.25**	.061
4. Feeling tired or having little energy	-.39**	.151
5. Poor appetite or overeating	-.24**	.056
6. Feeling bad about yourself – or that you are a failure or have let yourself or your family down	-.25**	.063
7. Trouble concentrating on things, such as reading the newspaper or watching television	-.18*	.033
8. Moving or speaking so slowly that other people could have noticed. Or the opposite – being so fidgety or restless that you have been moving around a lot more than usual	-.22**	.048
9. Thoughts that you would be better off dead, or of hurting yourself	-.10	.010
10. If you circled any problems, how difficult have these problems made it for you to do your work, take care of things at home, or get along with people?	-.23**	.055

Short-Form Health Survey		
1. How would you describe your general health?	.34**	.117
2. Does your health now limit you in moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf?	.07	.005
3. Does your health now limit you in climbing several flights of stairs?	.15	.023
4. During the past 4 weeks have you accomplished less than you would like with your work or other regular activities as a result of your physical health?	.15	.023
5. During the past 4 weeks were you limited in the kind of work or other activities as a result of your physical health?	.12	.014
6. During the past 4 weeks, were you limited in the kind of work you do or other regular activities as a result of any emotional problems (such as feeling depressed or anxious)? Accomplished less than you would like?	.17*	.030
7. During the past 4 weeks, were you limited in the kind of work you do or other regular activities as a result of any emotional problems (such as feeling depressed or anxious)? Didn't do work or other activities as carefully as usual?	.17*	.028
8. During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?	.21*	.043
9. How much of the time during the past 4 weeks have you felt calm and peaceful?	.45**	.205
10. How much of the time during the past 4 weeks did you have a lot of energy?	.43**	.183
11. How much of the time during the past 4 weeks have you felt downhearted and blue?	.49**	.237
12. During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?	.23**	.051

*$p < .01$. **$p < .001$.

Results indicated that each item on the PANAS was significantly correlated with the AM Happy Scale ($p < .001$ or $p < .01$), except for the item "afraid." As expected, items on the PA subscale had a significant convergent relationship with the AM Happy Scale, and items on the NA subscale had a divergent relationship with the AM Happy Scale. Of note, the variance accounted for by the ranks of each item in the ranks of the AM Happy Scale was fairly small, particularly in the NA subscale, and ranged from 2.2% to 10.6%, indicating minimal effect sizes. The PA subscale had slightly larger effects, and the rank of items within that scale accounted for 9.2% to 20.4% of the variance in the ranks of the AM Happy Scale, which also shows a minimal effect.

The SIWB showed a similar trend, and each item on the SIWB was significantly correlated with the AM Happy Scale ($p < .001$). There were no apparent differences in terms of the proportion of variance shared by the ranks in the AM Happy Scale and the ranks of the items in the two subscales of the SIWB, which ranged from 8.7% to 19.3% indicating a minimal effect. All of the items on the PHQ-9 also had significant divergent relationships ($p < .001$ or $p < .01$) with the AM Happy Scale, except for Question 9, which asked participants about suicidal thoughts. The variance accounted for by the ranks of the items on the PHQ-9 and the ranks of the AM Happy Scale was fairly low, and ranged from 1.0% to 15.1%, which indicates a minimal effect.

Unlike the other scales, four of the 12 items on the SF-12 did not show significant relationships with the AM Happy Scale. All of these four items were on the PCS subscale, and asked participants about physical health limitations in daily activities. The other two items on the PCS subscale that asked participants about their general health and the impact of pain in their normal work were significantly positively correlated with the

AM Happy Scale ($p < .001$ and $p < .01$, respectively). However, the variance accounted by the ranks of these two items and the ranks of the AM Happy Scale was fairly small and was 11.7% for the general health item and 4.3% for the pain item. These numbers indicate a minimal effect size. All of the items on the MCS subscale were significantly positively correlated with the AM Happy Scale. The variance accounted for by the ranks of the items on the MCS subscale and the ranks of the items on the AM Happy Scale ranged from 2.8% to 23.7%, which also represent a minimal effect.

CHAPTER FIVE

DISCUSSION

The primary aims of the current study were to determine an overall range of reliability for the Happiness Scale and to test the convergent and divergent validity of the AM Happy Scale by comparing it to other well-established measures of happiness and unhappiness. A secondary aim of the study was to compare the AM Happy Scale to a measure of physical and mental health. An overall range of reliability for the Happiness Scale was established using the correction for attenuation formula, although the range was much wider than predicted. There were significant positive associations between the AM Happy Scale and two measures of happiness, the PANAS PA subscale and the SIWB (total scale and its two subscales). There were significant negative associations between the AM Happy Scale and two measures of unhappiness, the PANAS NA subscale and the PHQ-9. Results also indicated that there were significant relationships between self-reported mental and physical health and the AM Happy Scale.

Of note, when comparing the current study to another validation study of a single-item happiness scale conducted in Kuwait (Abdel-Khalek, 2006), an interesting finding was gender differences on the one-item scale of happiness. In the current study, females reported slightly higher levels of happiness (M_d = .21) measured by the AM Happy Scale than males. On the contrary, results from Abdel-Khalek's (2006) study showed that males reported higher levels of happiness than females (M_d = .66 to .94). The differences in means varied by sample and the discrepancies were larger for younger participants than older participants. Although there have been a few studies that support gender differences in happiness (e.g., Perneger et al., 2004; Takeyachi et al., 2003), the majority of the prior

research in this area indicates that there are no gender differences in overall happiness (see Myers & Diener, 1995 for a review). Regardless, the effect size of the differences between males and females in the current study was minimal and did not meet the minimum cutoff of .2 for a significant effect as recommended by Ferguson (2009).

Moreover, the overall happiness scores for participants in the Kuwait study were lower and the standard deviations were larger (M = 5.89 to 7.36, SD = 2.01 to 2.92, depending on the study) than the scores found in the current study (M = 8.34, SD = 1.05). Of note, Abdel-Khalek used a 0 to 10 scale, rather than the 1 to 10 ranks used on the AM Happy Scale. Despite the differences in measurement and in participants, the current study revealed a trend that is consistent with findings reported in the 2016 World Happiness Report (Helliwell, Huang, & Wang, 2016). The report indicated that people in the United States endorsed higher levels of happiness (M = 7.10) than people in Kuwait (M = 6.24). Although the current study is a validation study of the AM Happy Scale and not a comparison study of happiness between nations, the trend provides further evidence for the AM Happy Scale as a valid and reliable form of measurement.

Reliability

Results did not provide support for the first hypothesis predicting that the range of reliability of the AM Happy Scale will be between 0.6 and 0.9. Rather, the range of estimated reliability in the present study was much broader producing values between .27 and 1.06. This is a wide range of values; however, the correction for attenuation formula shows that actual reliability for the AM Happy Scale cannot be higher than .37 (using the PANAS PA) or .27 (using the SIWB) because the underlying construct correlation

between the single and multiple-item happiness measures was assumed to be a perfect relationship (1.00) when these values were calculated. The formula shows that the more conservative the true underlying correlation between two variables is assumed to be, the higher the minimum reliability estimate. This pattern was evident in the current study, as the most conservative numbers we used in the correction for attenuation formula yielded the highest reliability estimates.

The range estimated in our hypothesis was determined by examining reliability coefficients of other single-item measures. For instance, Abdel-Khalek (2006) developed a single-item measure for happiness in an Arab population and found that the temporal stability of this measure was .86. This researcher did not use the correction for attenuation formula to obtain a minimum reliability estimate. Instead, the author reported the test-retest reliability of the single-item scale of happiness by administering the measure at two time points taken one week apart.

Levin and Currie (2014) also examined the test-retest reliability (using a time lag of 2 to 4 weeks) of an adapted version of Cantril's ladder in a sample of Scottish adolescents and found correlations between .58 and .70, which are slightly closer to our minimum reliability estimate values. Another study that examined the temporal reliability of Cantril's ladder in a group of community residents ages 46 to 70 over a two-year period yielded a reliability coefficient of .65 (Palmore & Kivett, 1977). The Delighted-Terrible Scale demonstrated a test-retest reliability of .65 over a 15-minute period and .40 over a six-month period (Andrews & Withey, 1976). Fordyce's Emotion Questionnaire produced test-retest coefficients of .98 for a two day period, .86 to .88 for a two week period, .81 for a one month period, and .62 and .67 for a four month period (Fordyce,

1988). These reliability coefficients indicate that reliability for single-item measures can range from .58 to .98, which is a smaller range than the reliability estimates for the AM Happy Scale found in the current study.

Robins and colleagues (2001), who developed a single-item scale designed to assess self-esteem, also reported the temporal stability of their scale based on longitudinal data. These researchers used a procedure developed by Heise (1969, Equation 9) to estimate the reliability of their single-item scale based on its patterns of autocorrelations over three time points. By using this method, the researchers calculated a mean reliability estimate of .75 for their single-item measure. The different methods of determining reliability may account for some of the discrepancy between our estimated reliability for the AM Happy Scale and the other two single-item rating scales. In future studies, researchers may wish to utilize a longitudinal design to determine the test-retest reliability of the AM Happy Scale using the Heise (1969) formula in order to obtain a better estimate of the scale's reliability.

The researchers who conducted two studies that used the correction for attenuation formula calculated reliability estimates for one-item measures of job satisfaction using only values of .9 and 1.0 for the assumed true underlying correlation between the one-item scales and other measures of job satisfaction if they were both measured perfectly (Dolbier et al., 2005; Wanous et al., 1997). This method resulted in reliability estimates of .45 and .56 in one study (Wanous et al., 1997), which are comparable to the reliability estimates calculated in the current study. In the other study, reliability estimates were calculated to be .73 and .90 (Dolbier et al., 2005), which are much higher compared to the values reported in the current study.

The discrepancies between those studies and the current study could be explained by the fact that the AM Happy Scale in the current study was intended to measure a different construct than the other two studies that measured job satisfaction. Furthermore, one of the studies (Wanous et al., 1997) was a meta-analysis of multiple studies that had used one-item measures of job satisfaction, which might explain the difference between the reliability estimates in that study compared to the reliability estimated calculated in the Dolbier and colleagues (2005) study. Results of these studies are provided to show how using the correction for attenuation formula can provide drastically different reliability estimates for the same construct.

Convergent and Divergent Validity

The convergent and divergent validity of the AM Happy Scale was established by its significant positive correlations with the PANAS PA subscale and the SIWB and its significant negative correlations with the PANAS NA and the PHQ-9. As shown in Table 3, the sizes of the relationships were minimal to moderate, but comparable to results from another study that tested the concurrent, convergent, and divergent validity of a single-item measure of happiness in an Arab cultural context in Kuwait (Abdel-Khalek, 2006). In this study, the researcher found correlations ranging from .56 to .70 between the single-item measure of happiness and the Oxford Happiness Inventory (Argyle, Martin, & Crossland, 1989) and correlations ranging from .45 to .63 between the single-item measure and the Satisfaction with Life Scale (Diener et al., 1985). The size of the correlations varied by sample, and the three samples used by this researcher included secondary school students, university undergraduates, and government employees.

Although different measures were used in both studies to determine convergent validity, the sizes of the correlations between the constructs found in Abdel-Khalek's study resemble the size of the correlations between the AM Happy Scale and the PANAS PA (.57) and SIWB (.50). The PANAS was not used in the study by Abdel-Khalek (2006); however, another measure of positive affect, the Positive Affect subscale of the Affect Balance Scale (Bradburn, 1969) was used to determine convergent validity. The correlation between this affective scale and the single-item measure was .34, a value that is lower than the association between the PANAS PA subscale and the AM Happy Scale found in the current study. Given that positive affect has shown to have a strong association with happiness and at times has been used to define happiness, the discrepancy in correlations between the single-item measures and positive affect may be due to a difference in measurement instruments or to a difference in the samples. It is also important to note that the reliability for the positive affect subscale used in Abdel-Khalek's (2006) study was fairly low (.55) and does not meet the minimum standards for reliability for research purposes (Cronbach's alpha of .70 or above) set forth by Furr and Bacharach (2008), which may have also impacted results.

To determine divergent validity, Abdel-Khalek (2006) also included the Negative Affect subscale of the Affect Balance Scale to compare with the one-item happiness measure and found a significant correlation of -.49. This correlation is slightly larger than the strength of the relationship found in the current study between the PANAS NA and the AM Happy Scale (-.38). Once again, the difference may be attributed to the different measuring instruments or different populations used in both studies. In the Abdel-Khalek (2006) study, the Affect Balance Scale was only administered to a population of

undergraduate students, whereas in the current study, the PANAS was administered to a group of older adults. Of note, the negative affect subscale in the Abdel-Khalek study (2006) did meet the minimum standards for reliability with a Cronbach's alpha of .70.

Results from the current study were also consistent with other findings showing that single-item measures of happiness had good convergent and divergent validity (Andrews & Crandall, 1975; Levin & Currie, 2014). In a validation study of Cantril's ladder using seven samples of Scottish adolescents, researchers found that the scale showed good convergent validity with other subjective well-being measures, including a life satisfaction scale and a global health-related quality-of-life measure, and good divergent validity with a measure of anxiety and depression (Levin & Currie, 2014). The sizes of the correlations between the single-item scale and other measures provided in the Scottish study were comparable to the correlations found in the current study, and consisted of .21 for the life satisfaction scale, ranged between .42 and .56 for the health-related quality-of-life measure, and ranged between -.33 and -.40 for the depression and anxiety measure (sizes of correlations varied by sample).

The validity coefficients computed for the AM Happy Scale are also consistent with convergent and divergent validity coefficients reported for the Fordyce Emotion Questionnaire (Fordyce, 1988). For example, Fordyce's questionnaire showed to be convergent with the Delighted-Terrible Scale (Andrews & Withey, 1976; r = .58), the Affectometer (Kammann & Flett, 1983; r = .69), Bradburn's positive affect score (Larsen, Diener, & Emmons, 1985; r = .53), Diener's Satisfaction with Life Scale (Larsen et al., 1985; r = .64), and Cantril's self-anchoring ladder rating of life satisfaction (Larsen et al., 1985; r = .58). The Fordyce Emotion Questionnaire was found to have a divergent

relationship with measures of depression, such as the Beck Depression Inventory (Fordyce, 1987; r = -.51), Bradburn's negative affect score (Larsen et al., 1985; r = -.33), and the Profile of Mood States depression scale (Fordyce, 1987; r = -.68).

The convergent and divergent validity coefficients calculated for the AM Happy Scale were also consistent with the results of a multimethod-multitrait analysis of several measures of happiness, including the Delighted-Terrible Scale (Andrews & Withey, 1976) and Cantril's self-anchoring ladder rating of life satisfaction (Cantril, 1965). Through the use of structural equation models, these researchers estimated that the validity of a single questionnaire or interview item used to assess perceptions of well-being fall in the range of .7 to .8 (Andrews & Crandall, 1975). The validity estimates determined by this study are slightly higher than the correlations found in the current study between the AM Happy Scale and other measures of happiness, which may be due to differences in methodology between the current study and the 1975 study. Andrews and Crandall (1975) used a different population of Americans described as closely resembling "a national sample of American adults" (p. 4); different methodology, including collecting multitrait-multimethod data from participants and from people who knew them well; and used a structural modeling approach.

Physical and Mental Health

Results of the current study confirmed the hypothesis predicting positive associations between mental and physical health and the AM Happy Scale, and are in line with findings from another study that also showed positive relationships between physical and mental health and a one-item measure of happiness (Abdel-Khalek, 2006).

The author did not use the SF-12, but instead used a one-item measure to ask participants about their general mental and physical health using a scale from 0-10. Results from Abdel-Khalek's study showed a similar trend to results from the current study, in which mental health (r = .70) had a larger correlation with happiness than physical health (r = .43).

Results from a study conducted with adolescents in Scotland found that happiness, also measured by a single-item scale, was negatively associated with subjective health complaints and anxiety and depression, measured by the General Health Questionnaire-12 (Levin & Currie, 2014). The effect sizes in this study were small; however, the trend was similar to what was found in our study. The physical health measure, subjective health complaints (r_s= -.31), exhibited smaller correlations with happiness than the measure of anxiety and depression (r_s = -.37).

The relationships between happiness, mental health, and physical health were also evident in another study using an older adult population in Alabama (Angner et al., 2013). This study, however, did not use a one-item scale to assess happiness. Instead, happiness was measured using the Subjective Happiness Scale and the SF-12 was used to measure self-reported health. The researchers found a moderate correlation between freedom from debility and happiness (.30), and they also found that unfavorable self-reported health status was associated with greater odds of being unhappy (OR = 2.90, 95% CI [1.59, 5.26]).

Another study examining happiness and health in older adult primary care patients living in Alabama also showed a moderate correlation (.37) between happiness and subjective reported health (Angner, Ray, Saag, & Allison, 2009). This study also

used the Subjective Happiness Scale to measure happiness, and a single question ("In general, would you say your health is: excellent, very good, good, fair, or poor?") was used to measure subjective health. Objective health was measured by multiple measures that consisted of detailed questions regarding health conditions, and two of these measures (debilitating pain and urinary incontinence; (OR = 6.05, 95% CI [3.38, 10.8] and (OR = 1.87, 95% CI [1.07, 3.29], respectively) were significantly associated with lower happiness levels (both *p*s < .001).

On the other hand, a study conducted in Switzerland also revealed significant positive relationships between happiness and mental health, but did not find the relationship between happiness and physical health to be significant (Perneger et al., 2004). Researchers conducting the Swiss study also used a single item to assess happiness taken from the SF-36 ("How much of the time during the past 4 weeks have you been a happy person? All of the time, most of the time, some of the time, a little of the time, none of the time") and they used the SF-12 to measure mental and physical health. Despite the findings in the Swiss study, overall, the results of the current study are consistent with the majority of the previous literature that has shown moderate relationships between happiness and mental health, but only minimal relationships between happiness and physical health.

Post-Hoc Analyses

Several post-hoc analyses were conducted in order to gain greater insight into the relationships between the AM Happy Scale and the other variables. Initially, we were interested in testing the relationship between the AM Happy Scale with the PHQ-9 as a

dichotomous variable divided into two categories of participants who were classified as depressed and non-depressed. We had intended to use a cutoff point of ≥ 10, which has shown to be the best predictive cutoff value of depression (Kroenke et al., 2001; Manea, Gilbody, & McMillan, 2012). However, only four of the 275 participants in the study endorsed values ≥ 10 on the PHQ-9. Thus, we recoded the variable into two categories of participants who scored zero and participants who scored ≥ 1, to represent people with no depression and people with some depression, respectively. This method provided a much more equal distribution of participants, with 101 participants in the non-depressed group and 163 participants in the depressed group. Results showed that the PHQ-9 was significantly related to the AM Happy Scale, although the effect was minimal. However, results provide additional support for the hypothesis that the AM Happy Scale has good divergent validity.

Furthermore, a series of correlational analyses were conducted to examine relationships between individual items in each scale and the AM Happy Scale. In regard to the PANAS, as predicted, items on the PA subscale had a significant convergent relationship with the AM Happy Scale, and items on the NA subscale had a divergent relationship with the AM Happy Scale. Each individual item was significantly correlated with the AM Happy Scale with the exception of one item, "afraid." An individual item analysis revealed that the item was highly positively skewed, with the majority of participants reporting feeling "afraid" very little or not at all. It is unclear why this variable did not have a significant negative correlation with the AM Happy Scale, especially given that the variable "scared" did have a small but significant relationship with the AM Happy Scale. The two adjectives have similar meanings and in the study,

the two variables had almost identical means and standard deviations. Researchers may wish to examine this phenomenon in future studies that include measures containing additional adjectives for fear to determine why the two words may produce different results.

The SIWB also confirmed our predictions as each item on the scale showed a significant positive relationship with the AM Happy Scale. Although the effects were minimal, it is important to note that the items on the Life Scheme subscale had slightly larger effect sizes, with an average effect size (R^2) of .16, compared to the items on the Self-Efficacy subscale that had an average effect size of .13. Although these differences are not immense, they make sense given the strong associations between life satisfaction and happiness found in multiple other studies (e.g., Diener et al., 1985; Diener et al., 2006; Lyubomirsky et al., 2005). The Life Scheme subscale does not exactly measure life satisfaction, but it contains items that are more aligned with the construct, such as "There is a great void in my life at this time," compared to the Self-Efficacy subscale.

Most of the items on the PHQ-9 were also significantly correlated with the AM Happy Scale. The relationships were negative, as was expected given that the PHQ-9 is a measure of depression. One item ("Thoughts that you would be better off dead, or of hurting yourself") on the PHQ-9 did not significantly correlate with the AM Happy Scale. An individual item analysis revealed that only two participants endorsed having thoughts of suicide on several days, whereas the remainder of the sample reported having no such thoughts. One assumption of Spearman's correlation is that relationships must be monotonic and a scatterplot revealed that the relationship between the AM Happy Scale

and the item regarding suicide was not monotonic, which may explain the non-significant finding (Goodwin & Leech, 2006; Hauke & Kossowski, 2011).

Contrary to our expectations, not all of the individual items on the SF-12 were significantly correlated with the AM Happy Scale. Specifically, four of the six items on the PCS subscale that measured physical health limitations in daily activities did not correlate with our happiness measure. The other two items on the PCS subscale that asked participants about their general health and the impact of pain in their normal work were significantly positively correlated with the AM Happy Scale, but the effect sizes were minimal. All of the items on the MCS subscale were significantly correlated with the AM Happy Scale, although the effect sizes were minimal. These results are consistent with prior findings indicating minor or non-significant relationships between happiness and measures of physical health, and minimal but significant relationships between happiness and measures of mental health (Abdel-Khalek, 2006; Angner et al., 2013; Angner et al., 2009; Levin & Currie, 2014; Perneger et al., 2004).

Strengths and Limitations

The current study has several strengths and limitations that deserve mention. With the exception of being unable to calculate an accurate measurement of internal reliability for the AM Happy Scale, the other measures used in the analyses demonstrated high internal consistency and the sample size provided sufficient power to conduct the statistical analyses. The structure and design of the AM Happy Scale as a single-item measure is also a strength given that single-item measures tend to take less space and can be more time-and cost-efficient than multiple item measures (Robins et al., 2001).

Additionally, the use of animations makes the scale user-friendly and can provide a visual representation of what the scale is intending to measure. Although the scale is specifically designed for United States, it can be modified for use in other countries and cultures.

To the best of our knowledge, the AM Happy Scale is the first ten-point end-defined scale that uses the term "happiness" versus "life satisfaction," as was done by Cantril (1965). Results show significant relationships between the AM Happy Scale, mood, life scheme, and self-efficacy with effect sizes comparable to correlations found in other studies. These results indicate that the AM Happy Scale encompasses both affective and cognitive components of happiness. Additionally, the AM Happy Scale is the first single-item scale that has added structure to the relative aspects of the scale by asking participants to rate themselves in comparison to other people in the United States.

However, asking participants to compare themselves to others in the United States may have some implications that are important to mention. The specific wording on the AM Happy Scale was chosen in order to provide some structure to the relative aspects of the scale and in hopes of eliciting a response that is coming from a broader perspective rather from comparing oneself to one's immediate social group. Given the rise of social media usage over the last few years and its reported negative impact on mood (e.g., Fardouly, Diedrichs, Vartanian, & Halliwell, 2015; Kaplan & Haenlein, 2010), we were concerned that social comparison theory may play an important role in participants' self-reports of happiness.

Social comparison theory is well-known for influencing individuals' happiness levels (e.g., Lyubomirsky & Ross, 1997; Wood, Taylor, & Lichtman, 1985; Wheeler & Miyake, 1992). The theory posits that people compare themselves to others and then

make judgments about the quality of their lives based on these observations (Festinger, 1954). Upward comparisons consist of comparing oneself to others who may appear to be better off or have desirable qualities that one is seeking (Morse & Gergen, 1970). These types of comparisons have been shown to have a negative impact on an individual's well-being and self-esteem (Morse & Gergen, 1970; Wood et al., 1985). Downward comparisons consist of comparing oneself to others who may appear to be worse off or possess traits that are undesirable. These types of comparisons have the opposite effect of upward comparisons, and result in increases in happiness and self-esteem (Wheeler & Miyake, 1992).

Of course, asking people to compare themselves to others in the United States does not guarantee the elimination of social comparison in self-reported happiness and presents its own set of challenges. For instance, some individuals may have limited exposure or knowledge about the happiness of others in the United States, and they may not be able to accurately compare themselves to others in the region. It would be interesting to learn how such individuals decided to rank their happiness (e.g., Did they use their immediate social circle as a comparison? Did they base their responses on what they had seen on social media?). In future studies, it may be helpful to ask participants this question to determine if people used the happiness levels of their friends and family group to estimate the levels of happiness of others in the U.S. Additionally, individuals may hold false assumptions, stereotypes, or biases about others in the United States, which may also impact their responses on the scale.

As mentioned, a limitation of the current study is that the internal consistency cannot be computed for a single-item measure. Although estimates were calculated using

the correction for attenuation formula, assumptions were used rather than concrete values, which may affect the accuracy of the estimates. In future studies, it would be helpful to obtain the test-retest reliability of the AM Happy Scale by administering the scale at different time points.

Another note that is important to mention in regard to measurement is the decision to use a scale with pre-determined increments resembling a Likert scale rather than using a visual analog scale (VAS). This decision was made with the intention of making the scale as user-friendly as possible. Several studies have provided evidence showing that Likert-type scales are preferred over VAS because they are easier to use for both participants and researchers and provide comparable results (e.g., Davey, Barratt, Butow, & Deeks, 2007; Jaeschke, Singer, & Guyatt, 1990; Joyce, Zutshi, Hrubes, & Mason, 1975; Laerhoven, Zaag-Loonen, & Derkx, 2004; Murphy, McDonald, Power, Unwin & MacSullivan, 1988). The use of a VAS requires participants to consider their status within a mathematical dimension, a task which may be difficult for some participants (Duncan, Bushnell, & Lavigne, 1989). In fact, some studies have even required that participants receive training to learn the correct use of the VAS (Jaeschke et al., 1990; Murphy et al., 1988).

While we are aware that there are some benefits to using a VAS over a Likert-type scale, such as having increased precision, better reproducibility, and better sensitivity to change in the assessment of symptoms (Grant et al., 1999; Paul-Dauphin, Guillemin, Virion, & Briançon, 1999), the benefits of using a Likert-type scale with pre-defined increments outweighed the costs in this particular study. Flynn, van Schaik, and

van Wersch (2004) and Hasson and Arnetz (2005) provide thorough summaries of the advantages and disadvantages of each measurement option.

The use of healthy older adults is a limitation in the current study as it limits the generalizability of the findings to other populations. The selection of "healthy" participants was not an accident, as the study was designed to exclude people who were diagnosed with major medical or mental illnesses. As previously mentioned, the current study is a part of a larger randomized controlled trial, and the goal of the larger study was to include healthy participants so that health problems would not confound the effect of the treatment on study outcome variables. Older adults were chosen for this study because the increases in life expectancy and rates of this population in the workforce make prevention and health maintenance especially important for this group (CDC, 2012; Colby & Ortman, 2015; SSA, 2013). Due to this participant selection process, the AM Happy Scale may generalize to other samples of healthy older adults, but additional future studies will need to be conducted in order to validate this instrument in other populations.

Furthermore, the sample was collected from the Loma Linda, California area. Loma Linda is designated as a "Blue Zone," defined as an area with unusually high rates of longevity (Buettner, 2015). Only five places in the world have been identified as Blue Zones, which makes our sample especially unique. People who live in Blue Zones are happier compared to people in other areas (Buettner, 2011), so it is not surprising that our sample scored high on the AM Happy Scale (M = 8.34 and SD = 1.05, on a 1-10 scale). Our findings also confirm prior research showing that happiness data for people in the United States generally tends to be negatively skewed, with the United States reporting as

one of the happiest countries in the world (Bond & Lang, 2014; Helliwell et al., 2016). Prior studies examining happiness in older adult populations also found the data to be negatively skewed (Angner et al., 2013; Angner et al., 2009).

In addition to being happy, engagement in spirituality and religion is also a common attribute of people living in Blue Zones (Buettner, 2015). The current study measured both variables, which may be conceptualized by some as separate constructs of experience, with spirituality representing the meaning that arises from life experiences (Corbett, 1990) and religiousness representing adherence to a set of organized beliefs, practices, and/or precepts of religion (Miller & Thoresen, 2003). However, Hill and Pargament (2003) argue that "most people experience spirituality within an organized religious context and fail to see the distinction between these phenomena" (p. 65). Evidence for Hill and Pargament's (2003) statement is provided in another study by Shahabi and colleagues (2002) who found that 52% of 1,422 participants from a stratified national sample of adults reported being both spiritual and religious. Only 10% of participants viewed themselves as only spiritual and another 10% of participants described themselves as being only religious. Meanwhile, 28% of participants identified themselves as neither spiritual nor religious.

In an analysis of the research on religion and spirituality, Miller and Thoresen (2003) conclude that "spirituality and religiousness may be best described as overlapping constructs, sharing some characteristics but also retaining nonshared features" (p. 28). These researchers along with Hill and Pargament (2003), who provide another review of the literature on spirituality and religion, emphasize that both constructs have been shown to have a positive impact on physical and mental health. In the present study, the majority

of participants (78.5%) in our sample endorsed belonging to a religious domination and scored extremely high on a measure of spirituality (M = 4.57 and SD = .60, on a 1-5 scale) providing additional evidence for spirituality and religion as overlapping constructs.

Furthermore, a social desirability response bias may also explain the high reported rates of happiness by participants in the current study. Social desirability refers to the tendency of participants to attribute to themselves statements that are desirable and reject those that are undesirable (Edwards, 1957). In an examination of response bias in another single item measure, Fordyce (1988) reported that most of the response bias comparisons between the Fordyce Emotion Questionnaire and several social desirability measures were non-significant. However, there were a few significant results indicating that the single item measure may be susceptible to some social desirability bias. The researcher warns that findings must be interpreted with caution but concludes that "for general research use… the HM [Happiness Measure, also known as the Emotion Questionnaire] can be considered relatively free of bias" (Fordyce, 1988, p. 372).

Initial validation studies of commonly used happiness scales (e.g., Satisfaction with Life Scale; Diener et al., 1985 and the presence of meaning subscale of the Meaning in Life Questionnaire [MLQ-P]; Steger, Frazier, Oishi,& Kaler, 2006) did not show any significant associations between these measures and a measure of social desirability (Marlowe-Crowne Social Desirability Scale; Crowne & Marlowe, 1964). Additionally, the creators of the PANAS, used in the current study, did not even include a measure of social desirability in their initial examination of the psychometric properties of the scale

(Watson et al., 1988). Overall, it appears that social desirability on well-being measures was not considered problematic in the past.

However, more recently, response bias has shown to be problematic in well-being measures. Soubelet and Salthouse (2011) found that the PANAS PA and NA subscales were significantly correlated with a measure of social desirability with correlation coefficients of .30 and -.22 ($p < .01$) for the PA and NA, respectively. The same study also found significant associations between a social desirability measure and life satisfaction and personality traits, such as agreeableness and conscientiousness. Other researchers examined response bias in measures of well-being by conducting five separate studies, each using a different method to test for the bias (Heintzelman, Trent, & King, 2015). Each of the five studies provided evidence of a consistent relationship between well-being measures, including the Satisfaction with Life Scale and the MLQ-P, and social desirability bias. The researchers attribute the changes over time in the relationship between desirability bias and well-being measures to the growth and dissemination of positive psychology research over the last few years, hypothesizing that the benefits of happiness have become so widespread that individuals may feel that it is unacceptable to report being unhappy.

Bowling, Bond, Jenkinson, and Lamping (1999) compared population norms collected from three studies conducted in Great Britain that used two different data collection methods to assess for health using the Short-Form 36 (SF-36) Health Survey questionnaire, the survey from which the SF-12 has been derived. One data collection method included face-to-face interviews and the other was via postal surveys. Results of this study revealed that participants who had face-to-face interviews scored higher in

multiple domains of the SF-36 compared to participants who submitted their responses by postal mail. The researchers concluded that the mode of questionnaire administration can affect data quality, and that data collected via face-to-face interviews may be more susceptible to the social desirability bias compared to data collected by other collection methods.

Bowling (2005) also conducted a review of the literature examining the effects of mode of questionnaire administration on data quality and found that there was a high potential for social desirability bias in data collected via face-to-face interviews and by telephone. On the contrary, self-administered surveys distributed via mail (postal and electronic) or through a computer program were the least susceptible to the bias. In the current study, there were a few participants who completed the questionnaires at the start or the end of the neuropsychological testing session, but most of the participants completed the questionnaires at home, on their own time, and returned them in person. Therefore, this method may have reduced some of the desirability response bias that can be associated with well-being measures.

Heintzelman and colleagues (2015) mention several options for reducing response bias, such as anonymity, peer reports, bio-medical markers, statistical controls, and controlling for responses on social desirability scales. However, the researchers present problems with each option and recommend two methods that may best control for desirability bias in future studies. The first is to use the bogus pipeline procedure to establish within-group desirability estimates that will provide valuable information about group norms. After desirability bias estimates are calculated within groups, researchers control for the within-group bias prior to making between-group comparisons.

The bogus pipeline procedure involves connecting participants to a device that ostensibly detects deception and has shown to increase the accuracy of scores on both socially desirable and undesirable characteristics (Roese & Jamieson, 1993). The bogus pipeline procedure may not be feasible for use in every study, and in particular those with a large number of participants and limited resources. However, including a social desirability scale and calculating and controlling for within-group desirability bias prior to calculating between-group differences is a feasible way to improve data accuracy in any study measuring characteristics that are susceptible to such bias. It would be well-advised that researchers conducting future validation studies on the AM Happy Scale also include a social desirability scale to measure such bias.

The current study used self-report measures, which may also influence participant response styles. Extreme response style (ERS) is the tendency to respond to questionnaires using extreme endpoints, high or low, on rating scales (Batchelor, Miao, & McDaniel, 2013). In a meta-analysis, Batchelor and colleagues (2013) cite research that shows that such a response style is content-irrelevant and typically viewed as stable across time and situations. The researchers also explain that ERS can be especially problematic when scales lack balance in terms of item direction. For instance, a scale designed to measure happiness is unbalanced when all the items on the scale are phrased in such a way that higher ratings always result in higher levels of happiness.

The lack of balance does not appear to be a problem for the AM Happy Scale, the SF-12, or the PANAS. The AM Happy Scale only contains one item, the SF-12 contains four items that are reverse-coded, and the PANAS does not contain items that are reverse-coded, but the scale measures two opposite constructs. The SIWB and the PHQ-9

used in the current study are unbalanced, as all of the items on each scale are worded in such a way that higher ratings result in higher levels of spirituality and depression, respectively. These scales may be more susceptible to error from ERS, and may provide another explanation for the extremely high scores of spirituality and low scores of depression reported in the current sample. ERS may still be problematic in spite of or in addition to the imbalanced scale problem due to the fact that it is construct-irrelevant, which increases within-group variance that in turn decreases statistical power and the magnitude of relationships among the variables.

Batchelor and colleagues (2013) also explored correlates of ERS in their meta-analysis and identified several participant and scale characteristics, including race, intelligence, acquiescence, education, age, and number of points on a scale that impacted the likelihood of ERS. Specifically, the researchers found that Black and Hispanic participants were more likely to engage in ERS compared to Whites, although the effect size for Hispanics was fairly small (d = -.09) when using values of .2, .5, and .8 to determine values that qualify as small, moderate, and large effect sizes, respectively, as recommended by Cohen (1969). In turn, Whites were more likely to engage in ERS than Asians (d = -.16). Given that the majority of our sample is White, it is possible that race played a role on the impact of ERS on results, but the small effect size of the racial difference in the Batchelor and colleagues (2013) study makes it unlikely.

Moreover, it was found that females were more likely to engage in ERS compared to males. The fact that the majority of our participants were female may be problematic in terms of ERS. However, once again, the effect size of the gender differences in the Batchelor and colleagues (2013) study was quite small (d = .09), indicating that gender

may not be such a strong determinant of ERS. On the other hand, ERS may explain why females scored higher than males on the AM Happy Scale. It would be advised to test this hypothesis in future studies with larger sample sizes.

Lower levels of intelligence also predicted a higher likelihood of ERS, but these results must also be interpreted with caution due to the fact that only two studies with a low number of total participants (N = 231) were included in the meta-analysis. High variance in education levels also increased the chances of ERS, and the researchers concluded that when education levels are heterogeneous, less educated samples would produce higher ERS. The standard deviation for years of education in our sample was 2.46 and the education level of our participants was high with 88.7% of participants reporting at least one year of college education or more. Therefore, given the high levels of education and moderate standard deviation in our sample, it is likely that education levels did not heavily impact ERS in the current study.

Acquiescence was also shown to impact ERS; however, there was no measurement of acquiescence included in the current study, so it is not possible to estimate its potential impact on ERS and on our results. A vector correlation revealed that younger age was also positively associated with ERS. Specifically, results indicated that ERS tends to increase until the age of 20 at which point it begins to decrease. The ages of participants in our sample were well above 20 years, so age was not likely an impacting factor on ERS in the present study. Lastly, the meta-analysis showed that ERS increased as the number of points and the number of items on a scale increased. The authors did not make recommendations about how many points to include in an ideal scale in order to minimize ERS, but the scales include in our study were fairly short (< 12 items) and

included a maximum of six response points, with the exception of the AM Happy Scale that had 10. Thus, it is uncertain but unlikely that the length of the surveys negatively impacted ERS in the current study.

Given that our sample consisted of older, highly educated adults and taking into account the small effect sizes for the other predictors of ERS, it is not likely that ERS has a large impact on our study. Implementing better selection processes that include short and balanced scales and assessment for ERS in future studies may reduce ERS, thereby improving the accuracy of results.

Over- and under-reporting are other forms of extreme responding, in which individuals either consciously or unconsciously provide inaccurate responses that are either higher (over-reporting) or lower (under-reporting) than their true responses (De Jong, Fox, & Steenkamp, 2015). This type of response bias is commonly seen on dietary surveys, where individuals often under-report their food intake (Black & Cole, 2011). In an effort to determine whether biased over-reporting or under-reporting is a characteristic of certain individuals or if it occurs randomly, Black and Cole (2011) analyzed data from seven longitudinal studies using multiple measures, including biological markers, in order to detect over-or under-reporters. The researchers found that over-and under-reporting is a characteristic of certain individuals and that those individuals who tended to over- or under-report on one measure were likely to over-or under-report on other measures as well, and that these patterns persisted over time.

Over- and under-reporting bias is not just limited to dietary studies. Happiness researchers have reported that participants in such studies also tend to respond in this biased manner (Veenhoven, 2000). Veenhoven (2000) proposes that “People who are

actually dissatisfied with their lives say that they are contented" and attributes ego protection and social appearances to be the cause of such distortions (p. 9). He proposes several hypotheses to explain the phenomenon. First, he suggests that people may not be over-reporting in happiness studies, but may truly be happy with life, which may be a legitimate response if their living conditions are good. Next, he suggests that people may underestimate the happiness of other people given that misery is more apparent than prosperity. Third, although he initially claims that psychosomatic complaints may be a sign of over-reporting in happy people, he suggests that sometimes a headache is just a headache and may not be indicative of bias. To the best of our knowledge and per his report, Veenhoven's hypotheses have not yet been tested.

De Jong and colleagues (2015) review previous methods used to detect over- and under-reporting, such as the objective criterion approach that depends on objective measures rather than self-report and the subjective criterion approach that depends on self-other criteria or social consensus criteria. The experimental approach is another method of preventing this bias, and is executed by comparing one group's answers to another group's answers that were incentivized to tell the truth. The researchers pose problems with each of these methods and propose an integrated "Bayesian item response theory model" that works by comparing answers obtained under direct questioning and randomized response. However, this method is designed to be used in marketing research, and a scientific method designed to detect and control for over-or under-reporting in psychological studies has not yet been developed. Over- or under-reporting may be a problem in the current study given that happiness scores were quite high and

depression scores were quite low; however, this may just be due to the fact that the study design and criteria was intended to only include healthy participants.

Lastly, non-response bias is another type of bias that is important to consider when analyzing results. In an analysis of the non-response bias phenomenon, Berg (2005) defines the concept as "the mistake one expects to make in estimating a population characteristic based on a sample of survey data in which, due to non-response, certain types of survey respondents are under-represented" (p. 3). In other words, non-response bias occurs when people who do not respond to surveys bias the results because they differ in some way from people who do respond to surveys (Hill, Roberts, Ewings, & Gunnell, 1997). Berg (2005) explains that when participants are systematically omitted from a particular sample because they have not responded to questionnaires, the sample can no longer be called "random." Thus, any patterns found in a non-random sample prevent results from being generalized to the entire population.

In order to prevent non-response bias, Berg (2005) recommends that researchers consult with the literature on study design in order to take preventative measures against non-response bias prior to starting the study. If a researcher is not involved in the data collection stage, Berg (2005) recommends analyzing the missing data using techniques from the statistical and econometric literature under the heading, "measurement error." Special terms can be used to describe the missing data, such as "volunteer bias" if the missing data was due to participants being volunteers.

Berg (2005) recommends testing for non-response bias through a method called validation that involves comparing two different samples drawn from the same population. Unfortunately, this method will not work in the current study given that we

only have one sample. Berg (2005) also outlines techniques, including imputation and weighting as ways to deal with missing data, concluding that the maximum-likelihood approach is the best way to correct for non-response bias. However, given that only 12 participants were deleted from the dataset for having missing data and 11 of those participants were missing 100% of the data, we did not believe that it was necessary to use this method of estimation.

Furthermore, compared to other studies that typically show average rate of 20% for non-responders (e.g., Hill et al., 1997; Whitehead, Groothuis, & Blomquist, 1993), the current study had a small percentage of non-responders (4.18%). The small percentage of non-responders in the current study may be due to the study design, which included face-to-face appointments, a method least likely to result in non-responders (Berg, 2005). Non-responders tend to be most prevalent and most problematic in studies using telephone reports (Berg, 2005; Hill et al., 1997). Furthermore, the small percentage of non-responders in the current study is consistent with findings from the happiness literature indicating that non-response bias in studies measuring happiness tends to be fairly small, typically ranging between 0% to 2%, with the exception of Japan, which has a 12% average non-response bias (Ouweneel & Veenhoven, 1991; Veenhoven, 2012).

In summary, a major limitation of the current study is that it did not take into account four different types of response biases, the social desirability bias, ERS, over- and under-reporting, and the non-response bias. Given the evidence presented above, it does not seem likely that the current study was heavily influenced by any of these response biases. However, it is possible that bias may partly explain the high levels of self-reported happiness and low levels of self-reported depression in our sample.

Alternatively, these scores may simply be a product of the study design to include only healthy participants. Future validation studies conducted on the AM Happy Scale should include methods to measure and control for these biases.

Implications

The current study has several implications for researchers and clinicians. Results show that the AM Happy Scale has adequate validity to measure happiness in older adults. Wanous and colleagues (1997) stipulate that a single-item measure may be an acceptable option to use under certain circumstances, including situational constraints, when time and space are limited, and when the research or assessment question implies the use of an overall measure of a certain construct. Although the stipulations for the use of single-item measures made by Wanous and colleagues (1997) were intended for job satisfaction measures, they can be applied to other single-item measures, such as happiness, as well. Research has shown that there are two main components of happiness, cognitive and affective (Diener, 2000), and the AM Happy Scale provides an estimate of an overall happiness level that appears to encompass both cognitive and affective dimensions of the construct.

The ability to quickly and easily measure an individual's level of happiness is of value to primary care physicians, who are often presented with patients with mental health problems that are often overlooked (Simon et al., 1999). The stigma associated with mental illness may prevent patients from openly discussing their mental health problems or endorsing symptoms on mental health screening tools, such as the PHQ-9 (Conner et al., 2010). However, using the term "happiness" rather than using the term

"depression" to screen for mental health problems may reduce some of the stigma associated with mental illness and allow individuals to respond more openly to such questions. Furthermore, because the AM Happy Scale was also a small, but albeit significant, predictor for physical health problems, it may also be of interest to health practitioners who wish to be proactive in encouraging patients to engage in happiness interventions to benefit both mental and physical health.

The links found in this study and in previous studies between happiness, mental health, and physical health may also encourage policymakers to take action as recommended by Veenhoven (2008) in order to make changes in perceptions of health at a societal level. Several researchers have provided compelling arguments for the importance of happiness in public policy (e.g., Dolan & Peasgood, 2008; Dolan & White, 2007; Frey & Stutzer, 2012; Helliwell, 2006), and policymakers have already taken steps to promote the well-being of the public. For instance, in 2011, the United Nations made a decision to start a movement that places a higher value on happiness in determining how to achieve and measure social and economic development ("Happiness in Development Policy," 2011).

The ability to screen for happiness levels is also of value to individuals, who when asked to reflect on their happiness compared to others, may realize that their score is lower in than they would like it to be. This awareness provides individuals with options to pursue interventions in order to raise their levels of happiness to not only improve their affect, but to also benefit from all of the positive outcomes associated with having higher levels of happiness. Future validation studies in other populations that include methods of controlling for different types of response biases are warranted.

Conclusion

In summary, the correction for attenuation formula produced a wide range of estimate reliabilities for the AM Happy Scale, with a maximum reliability estimate of .37. This value indicates that the AM Happy Scale is not suitable for research purposes. However, using the correction for attenuation formula is not a common practice in scale reliability testing; thus, it is advised that future studies measure the AM Happy Scale at different time points in order to obtain a test-retest reliability coefficient.

The AM Happy Scale showed adequate convergent and divergent validity when compared with two measures of happiness and unhappiness. Although effect sizes were small, the size of the correlations between the AM Happy scale, PANAS, SIWB, and PHQ-9 were comparable with the size of correlations from other studies that also investigated relationships between single-item happiness scales and other well-known measures of happiness. Some of these studies were conducted in other countries and used different populations, which indicates that the AM Happy Scale might be generalizable to other populations rather than just older adults in the United States. However, future studies will need to be conducted in order to test this hypothesis.

Lastly, the AM Happy Scale showed a significant relationship between physical and mental health. Although the effect sizes were small, the results obtained in the current study are consistent with results from prior studies examining similar relationships. The effect on mental health was larger than the effect on physical health, which also reflects previous findings in the literature. These findings have implications for individuals, healthcare providers, and policy makers, highlighting the significant relationship between happiness and health.

Overall, the AM Happy Scale appears to be an adequate measure of happiness in older adults. This study provides insight into how using the term "happiness" rather than "life satisfaction" in an adapted version of Cantril's (1969) original self-anchoring scale can produce notable results. Findings from the current study also confirm relationships identified in previous studies among happiness, mood, spirituality, life satisfaction, self-efficacy, and mental and physical health. Although more research is needed to confirm these findings and to validate the scale in other populations, current study findings suggest that the AM Happy Scale may be a reasonable option of assessing happiness levels for individuals, providers, and researchers.

REFERENCES

Abdel-Khalek, A. M. (2006). Measuring happiness with a single-item scale. *Social Behavior and Personality: An International Journal, 34*(2), 139-150.

Abdel-Khalek, A. M., & Lester, D. (2017). The association between religiosity, generalized self-efficacy, mental health, and happiness in Arab college students. *Personality and Individual Differences, 109,* 12-16. doi:10.1016/j.paid.2016.12.010

Achat, H., Kawachi, I., Spiro, A., III, DeMolles, D. A., & Sparrow, D. (2000). Optimism and depression as predictors of physical and mental health functioning: The Normative Aging Study. *Annals of Behavioral Medicine, 22*(2), 127–130.

Ackerman, C. E. (2015, April). *Measures in positive psychology.* Presented at the 95th annual meeting of the Western Psychological Association, Las Vegas, NV.

Ai, A. L., Tice, T. N., Peterson, C., & Huang, B. (2005). Prayers, spiritual support, and positive attitudes in coping with the September 11 national crisis. *Journal of Personality, 73*(3), 763-792.

Allison, P. D. (2001). *Missing data*. Thousand Oaks, CA: Sage Publications, Inc.

Andrews, F. M., & Crandall, R. (1976). The validity of measures of self-reported well-being. *Social Indicators Research, 3*(1), 1-19.

Andrews, F. M., & Withey, S. B. (1976). *Social indicators of well-being: Americans' perceptions of life quality*. New York: Plenum Press.

Angner, E., Ghandhi, J., Purvis, K. W., Amante, D., & Allison, J. (2013). Daily functioning, health status, and happiness in older adults. *Journal of Happiness Studies, 14*(5), 1563-1574.

Angner, E., Ray, M. N., Saag, K. G., & Allison, J. J. (2009). Health and happiness among older adults: A community-based study. *Journal of Health Psychology, 14*(4), 503-512.

Argyle, M., Martin, M., & Crossland, J. (1989). Happiness as a function of personality and social encounters. In J. P. Forgas & J. M. Innes (Eds.), *Recent advances in social psychology: An international perspective* (pp. 189–203). Amsterdam, The Netherlands: Elsevier.

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology, 63*(4), 596-612.

Aspinwall, L. G., & Brunhart, S. M. (1996). Distinguishing optimism from denial: Optimistic beliefs predict attention to health threats. *Personality and Social Psychology Bulletin, 22*(10), 993-1003.

Batchelor, J. H., Miao, C., & McDaniel, M. A. (2013, April). *Extreme response style: A meta-analysis*. Presented at the 28th Annual Conference of the Society for Industrial and Organizational Psychology, Houston, TX.

Benyamini, Y., Idler, E. L., Leventhal, H., & Leventhal, E. A. (2000). Positive affect and function as influences on self-assessments of health expanding our view beyond illness and disability. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *55*(2), P107-P116.

Berg, N. (2005). Non-response bias. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (Fall 2010 ed.). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1691967

Bertrand, M., & Mullainathan, S. (2001). Do people mean what they say? Implications for subjective survey data. *The American Economic Review, 91*(2), 67-72.

Black, A. E., & Cole, T. J. (2001). Biased over-or under-reporting is characteristic of individuals whether over time or by different assessment methods. *Journal of the American Dietetic Association, 101*(1), 70-80.

Blazer, D. G. (2003). Depression in late life: Review and commentary. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *58*(3), M249-M265.

Bond, T. N., & Lang, K. (2014). *The sad truth about happiness scales* (No. w19950). National Bureau of Economic Research.

Borge, L., Martinsen, E. W., Ruud, T., Watne, Ø., & Friis, S. (1999). Quality of life, loneliness, and social contact among long-term psychiatric patients. *Psychiatric Services, 50*(1), 81-84.

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health*, *27*(3), 281-291.

Bowling, A., Bond, M., Jenkinson, C., & Lamping, D. L. (1999). Short-Form 36 (SF-36) Health Survey questionnaire: which normative data should be used? Comparisons between the norms provided by the Omnibus Survey in Britain, the Health Survey for England and the Oxford Healthy Life Survey. *Journal of Public Health*, *21*(3), 255-270.

Bradburn, N. M. (1969). *The structure of psychological well-being*. Oxford, England.

Buettner, D. (2015). *The blue zones solution: Eating and living like the world's healthiest people*. National Geographic Books.

Buettner, D. (2011). *Thrive: Finding happiness the blue zones way*. National Geographic Books.

Campbell, A. (1981). *The sense of well-being in America*. New York: McGraw- Hill.

Campbell, A., Converse, P. E., & Rodgers, W. L. (1976). *The quality of American life: Perceptions, evaluations, and satisfactions*. New York: Russell Sage.

Cantril, H. (1965). *The pattern of human concerns*. New Brunswick, NJ: Rutgers University Press.

Carney, R. M., Freedland, K. E., Miller, G. E., & Jaffe, A. S. (2002). Depression as a risk factor for cardiac mortality and morbidity: A review of potential mechanisms. *Journal of Psychosomatic Research, 53*(4), 897-902.

Carpenter, J. S. (1996). Applying the Cantril methodology to study self-esteem: Psychometrics of the Self-Anchoring Self-Esteem Scale. *Journal of Nursing Measurement, 4*(2), 171-189.

Centers for Disease Control and Prevention. (2012, July). *Older employees in the work place* (Issue Brief No. 1). Atlanta, GA: National Center for Chronic Disease Prevention and Health Promotion, Division of Population Health.

Ciarrocchi, J. W., & Brelsford, G. M. (2009). Spirituality, religion, and substance coping as regulators of emotions and meaning making: Different effects on pain and joy. *Journal of Addictions & Offender Counseling, 30*(1), 24-36.

Cloninger, C. R., & Zohar, A. H. (2011). Personality and the perception of health and happiness. *Journal of Affective Disorders, 128*(1), 24-32.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Cohen, S., & Pressman, S. D. (2006). Positive affect and health. *Current Directions in Psychological Science, 15*(3), 122-125.

Cohn, D. & Taylor, P. (2010). *Baby boomers approach 65 – glumly*. Retrieved from http://www.pewsocialtrends.org/2010/12/20/baby-boomers-approach-65-glumly/

Colby, S. L. & Ortman, J. M. (2015). *Projections of the Size and Composition of the U.S. Population: 2014 to 2060* (Report No. P25-1143). Washington, DC: U.S. Census Bureau.

Conner, K. O., Copeland, V. C., Grote, N. K., Koeske, G., Rosen, D., Reynolds, C. F., & Brown, C. (2010). Mental health treatment seeking among older adults with depression: The impact of stigma and race. *The American Journal of Geriatric Psychiatry, 18*(6), 531-543.

Corbett, J. M. (1990). *Religion in America*. Englewood Cliffs, NJ: Prentice Hall.

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, *43*(3), 245-265.

Crowne, D. P., & Marlowe, D. A. (1964). *The approval motive*. New York, NY: Wiley.

Crowther, M. R., Parker, M. W., Achenbaum, W. A., Larimore, W. L., & Koenig, H. G. (2002). Rowe and Kahn's model of successful aging revisited positive spirituality—the forgotten factor. *The Gerontologist, 42*(5), 613-620.

Cummins, R. A. (1998). The second approximation to an international standard for life satisfaction. *Social Indicators Research, 43*(3), 307-334.

Cummins, R. A. & Gullone, E. (2000). *Why we should not use 5-point Likert scales: The case for subjective quality of life measurement*. Proceedings, Second International Conference on Quality of Life in Cities (pp.74-93). Singapore: National University of Singapore.

Daaleman, T. P., & Frey, B. B. (2004). The Spirituality Index of Well-Being: A new instrument for health-related quality-of-life research. *The Annals of Family Medicine, 2*(5), 499-503.

Daaleman, T. P., Frey, B. B., Wallace, D., & Studenski, S. A. (2002). Spirituality Index of Well-Being Scale: Development and testing of a new measure. *The Journal of Family Practice, 51*(11), 952-952.

Daaleman, T. P., Perera, S., & Studenski, S. A. (2004). Religion, spirituality, and health status in geriatric outpatients. *The Annals of Family Medicine, 2*(1), 49-53.

Daaleman, T. P., & VandeCreek, L. (2000). Placing religion and spirituality in end-of-life care. *JAMA, 284*(19), 2514-2517.

Danner, D. D., Snowdon, D. A., & Friesen, W. V. (2001). Positive emotions in early life and longevity: Findings from the nun study. *Journal of Personality and Social Psychology*, *80*(5), 804-813.

Davey, H. M., Barratt, A. L., Butow, P. N., & Deeks, J. J. (2007). A one-item question with a Likert or Visual Analog Scale adequately measured current anxiety. *Journal of Clinical Epidemiology, 60*(4), 356-360.

De Jong, M. G., Fox, J. P., & Steenkamp, J. B. E. (2015). Quantifying under-and over-reporting in surveys through a dual-questioning-technique design. *Journal of Marketing Research, 52*(6), 737-753.

Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist, 55*(1), 34-43.

Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment, 49*(1), 71-75.

Diener, E., Lucas, R. E., & Scollon, C. N. (2006). Beyond the hedonic treadmill: Revising the adaptation theory of well-being. *American Psychologist, 61*(4), 305-314.

Diener, E., Scollon, C. N., & Lucas, R. E. (2003). The evolving concept of subjective well-being: The multifaceted nature of happiness. *Advances in Cell Aging and Gerontology, 15*(2), 187-219.

Dolan, P., & Peasgood, T. (2008). Measuring well-being for public policy: Preferences or experiences. *Journal of Legal Studies, 37*(S2), S5–S31.

Dolan, P., & White, M. (2007). How can measures of subjective well-being be used to inform public policy? *Perspectives on Psychological Science, 2*(1), 71–85.

Dolbier, C. L., Webster, J. A., McCalister, K. T., Mallon, M. W., & Steinhardt, M. A. (2005). Reliability and validity of a single-item measure of job satisfaction. *American Journal of Health Promotion, 19*(3), 194-198.

Duncan, G. H., Bushnell, M. C., & Lavigne, G. J. (1989). Comparison of verbal and visual analogue scales for measuring the intensity and unpleasantness of experimental pain. *Pain, 37*(3), 295-303.

Eaton, W. W., Armenian, H., Gallo, J., Pratt, L., & Ford, D. E. (1996). Depression and risk for onset of type II diabetes: A prospective population-based study. *Diabetes Care, 19*(10), 1097-1102.

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. Ft Worth, TX: Dryden Press.

Elam, D. A. (2000). *An exploration of the relationship between spirituality and emotional well-being* (Doctoral dissertation, Southern Illinois University at Carbondale).

Fardouly, J., Diedrichs, P. C., Vartanian, L. R., & Halliwell, E. (2015). Social comparisons on social media: The impact of Facebook on young women's body image concerns and mood. *Body Image, 13*, 38-45.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40*(5), 532-538.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*(2), 117-140.

Field, A. (2009). *Discovering statistics using SPSS*. London, England: Sage Publications.

Fletcher, W. L., & Hansson, R. O. (1991). Assessing the social components of retirement anxiety. *Psychology and Aging, 6*(1), 76-85.

Flynn, D., van Schaik, P., & van Wersch, A. (2004). A comparison of multi-item Likert and visual analogue scales for the assessment of transactionally defined coping function. *European Journal of Psychological Assessment, 20*(1), 49-58.

Fordyce, M. (1987). *Research and tabular supplement for the Happiness Measures*. Fort Myers, FL: Cypress Lake Media.

Fordyce, M. W. (1988). A review of research on the happiness measures: A sixty second index of happiness and mental health. *Social Indicators Research, 20*(4), 355-381.

Fredrickson, B. L., Tugade, M. M., Waugh, C. E., & Larkin, G. R. (2003). What good are positive emotions in crisis? A prospective study of resilience and emotions following the terrorist attacks on the United States on September 11th, 2001. *Journal of Personality and Social Psychology, 84*(2), 365-376.

Frey, B., & Stutzer, A. (2012). The use of happiness research for public policy. *Social Choice and Welfare, 38*(4), 659–674.

Furr, R. M. & Bacharach, V. R., (2008). *Psychometrics: An introduction.* Thousand Oaks, CA: Sage Publications, Inc.

Gallup, G., Jr., & Lindsay, D. M. (1999). *Surveying the religious landscape: Trends in U.S. beliefs*. Harrisburg, PA: Morehourse.

Gandek, B., Ware, J. E., Aaronson, N. K., Apolone, G., Bjorner, J. B., Brazier, J. E., ... & Sullivan, M. (1998). Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: Results from the IQOLA Project. *Journal of Clinical Epidemiology, 51*(11), 1171-1178.

Goldenberg, S., & Wekerle, G. R. (1972). From utopia to total institution in a single generation: The Kibbutz and the Bruderhof. *International Review of Modern Sociology, 2*(2), 224-232.

Gomez, R., & Fisher, J. W. (2003). Domains of spiritual well-being and development and validation of the spiritual well-being questionnaire. *Personality and Individual Differences, 35*, 1975–1991.

Goodwin, L. D., & Leech, N. L. (2006). Understanding correlation: Factors that affect the size of r. *The Journal of Experimental Education*, *74*(3), 249-266.

Grant, S., Aitchison, T., Henderson, E., Christie, J., Zare, S., McMurray, J., & Dargie, H. (1999). A comparison of the reproducibility and the sensitivity to change of visual analogue scales, Borg scales, and Likert scales in normal subjects during submaximal exercise. *Chest Journal*, *116*(5), 1208-1217.

Grembowski, D., Patrick, D., Diehr, P., Durham, M., Beresford, S., Kay, E., & Hecht, J. (1993). Self-efficacy and health behavior among older adults. *Journal of Health and Social Behavior, 34*(2), 89-104.

Gurin, G., Veroff, J., & Feld, S. (1960). *Americans view their health*. New York, NY: Basic Books.

Happiness should have greater role in development policy – UN Member States. (2011, July 19). *UN News Centre*. Retrieved from http://www.un.org/apps/news/story.asp?NewsID=39084#.WP7ZPsb_rIU

Harker, L., & Keltner, D. (2001). Expressions of positive emotion in women's college yearbook pictures and their relationship to personality and life outcomes across adulthood. *Journal of Personality and Social Psychology*, *80*(1), 112-124.

Hasson, D., & Arnetz, B. B. (2005). Validation and findings comparing VAS vs. Likert scales for psychosocial measurements. *International Electronic Journal of Health Education, 8*, 178-192.

Hauke, J., & Kossowski, T. (2011). Comparison of values of Person's and Spearman's correlation coefficients on the same sets of data. *Quaestiones Geographicae, 30*(2), 87-93.

Hazan, C., & Shaver, P. (1987). Romantic love conceptualized as an attachment process. *Journal of Personality and Social Psychology, 52*(3), 511-524.

Heintzelman, S. J., Trent, J., & King, L. A. (2015). Revisiting desirable response bias in well-being reports. *The Journal of Positive Psychology*, *10*(2), 167-178.

Heise, D. R. (1969). Separating reliability and stability in test-retest correlation. *American Sociological Review, 34*(1), 93-101.

Helliwell, J. (2006). Well-being, social capital, and public policy: What's new? *Economic Journal, 116*(510), C34–C45.

Helliwell, J. F., Huang, H., & Wang, S. (2016). The distribution of world happiness. *World happiness report 2016: Update*, 9-49.

Hervás, G., & Vázquez, C. (2013). Construction and validation of a measure of integrative well-being in seven languages: The Pemberton Happiness Index. *Health and quality of life outcomes, 11*(1), 1-13.

Hill, P. C., & Pargament, K. I. (2003). Advances in the conceptualization and measurement of religion and spirituality: Implications for physical and mental health research. *American Psychologist, 58*(1), 64-74.

Hill, A., Roberts, J., Ewings, P., & Gunnell, D. (1997). Non-response bias in a lifestyle survey. *Journal of Public Health, 19*(2), 203-207.

Holder, M. D., Coleman, B., & Wallace, J. M. (2010). Spirituality, religiousness, and happiness in children aged 8–12 years. *Journal of Happiness Studies, 11*(2), 131-150.

Holt, C. L., Wang, M. Q., Caplan, L., Schulz, E., Blake, V., & Southward, V. L. (2011). Role of religious involvement and spirituality in functioning among African Americans with cancer: Testing a mediational model. *Journal of Behavioral Medicine*, *34*(6), 437-448.

Hooker, K., & Siegler, I. C. (1993). Life goals, satisfaction, and self-rated health: Preliminary findings. *Experimental Aging Research*, *19*(1), 97-110.

Hu, J., & Gruber, K. J. (2008). Positive and negative affect and health functioning indicators among older adults with chronic illnesses. *Issues in Mental Health Nursing*, *29*(8), 895-911.

Hu, J., & Meek, P. (2005). Health-related quality of life in individuals with chronic obstructive pulmonary disease. *Heart & Lung: The Journal of Acute and Critical Care*, *34*(6), 415-422.

Hunagund, D. L., & Hangal, S. J. (2014). Self-efficacy and happiness in youth. *Journal of the Indian Academy of Applied Psychology, 40*(1), 70-73.

Isaacowitz, D. M., & Smith, J. (2003). Positive and negative affect in very old age. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 58*(3), P143-P152.

Jaeschke, R., Singer, J., & Guyatt, G. H. (1990). A comparison of seven-point and visual analogue scales: Data from a randomized trial. *Controlled Clinical Trials*, *11*(1), 43-51.

Jeste, D. V., Alexopoulos, G. S., Bartels, S. J., Cummings, J. L., Gallo, J. J., Gottlieb, G. L., ... & Lebowitz, B. D. (1999). Consensus statement on the upcoming crisis in geriatric mental health: Research agenda for the next 2 decades. *Archives of General Psychiatry*, *56*(9), 848-853.

Jonas, B. S., & Lando, J. F. (2000). Negative affect as a prospective risk factor for hypertension. *Psychosomatic Medicine, 62*(2), 188-196.

Joyce, C. R. B., Zutshi, D. W., Hrubes, V., & Mason, R. M. (1975). Comparison of fixed interval and visual analogue scales for rating chronic pain. *European Journal of Clinical Pharmacology, 8*(6), 415-420.

Kahneman, D., & Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences, 107*(38), 16489-16493.

Kahneman, D., & Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *The Journal of Economic Perspectives, 20*(1), 3-24.

Kammann, R., & Flett, R. (1983). Affectometer 2: A scale to measure current level of general happiness. *Australian Journal of Psychology, 35*(2), 259-265.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, *53*(1), 59-68.

Katon, W., & Ciechanowski, P. (2002). Impact of major depression on chronic medical illness. *Journal of Psychosomatic Research*, *53*(4), 859-863.

Kercher, K. (1992). Assessing subjective well-being in the old-old: The PANAS as a measure of orthogonal dimensions of positive and negative affect. *Research on Aging*, *14*(2), 131-168.

Kivimäki, M., Vahtera, J., Elovainio, M., Helenius, H., Singh-Manoux, A., & Pentti, J. (2005). Optimism and pessimism as predictors of change in health after death or onset of severe illness in family. *Health Psychology, 24*(4), 413-421.

Knight, J., Song, L., & Gunatilaka, R. (2009). Subjective well-being and its determinants in rural China. *China Economic Review, 20*(4), 635-649.

Koenig, H. G., George, L. K., & Titus, P. (2004). Religion, spirituality, and health in medically ill hospitalized older patients. *Journal of the American Geriatrics Society, 52*(4), 554-562.

Koller, M., Kussman, J., Lorenz, W., Jenkins, M., Voss, M., Arens, E., ... & Rothmund, M. (1996). Symptom reporting in cancer patients: The role of negative affect and experienced social stigma. *Cancer*, *77*(5), 983-995.

Kozma, A., & Stones, M. J. (1987). Social desirability in measures of subjective well-being: A systematic evaluation. *Journal of Gerontology, 42*(1), 56-59.

Krause, N., Liang, J., Shaw, B. A., Sugisawa, H., Kim, H. K., & Sugihara, Y. (2002). Religion, death of a loved one, and hypertension among older adults in Japan. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 57*(2), S96-S107.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9):606-613.

Kubzansky, L. D., Sparrow, D., Vokonas, P., & Kawachi, I. (2001). Is the glass half empty or half full? A prospective study of optimism and coronary heart disease in the normative aging study. *Psychosomatic Medicine*, *63*(6), 910-916.

Laerhoven, H. V., Zaag-Loonen, H. V. D., & Derkx, B. H. (2004). A comparison of Likert scale and visual analogue scales as response options in children's questionnaires. *Acta Paediatrica, 93*(6), 830-835.

Lam, C. L., Tse, E. Y., & Gandek, B. (2005). Is the standard SF-12 health survey valid and equivalent for a Chinese population?. *Quality of Life Research, 14*(2), 539-547.

Lamers, F., Jonkers, C. C., Bosma, H., Penninx, B. W., Knottnerus, J. A., & van Eijk, J. T. M. (2008). Summed score of the Patient Health Questionnaire-9 was a reliable and valid method for depression screening in chronically ill elderly patients. *Journal of Clinical Epidemiology, 61*(7), 679-687.

Larsen, R. (1989). Is feeling "in control" related to happiness in daily life?. *Psychological Reports, 64*(3), 775-784.

Larsen, R. J., Diener, E. D., & Emmons, R. A. (1985). An evaluation of subjective well-being measures. *Social Indicators Research, 17*(1), 1-17.

Leventhal, H., & Patrick-Miller, L. (2000). Emotions and physical illness: Causes and indicators of vulnerability. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 523-537). New York: Guildford Press.

Levin, K. A., & Currie, C. (2014). Reliability and validity of an adapted version of the Cantril Ladder for use with adolescent samples. *Social Indicators Research, 119*(2), 1047-1063.

Levy, S. M., Lee, J., Bagley, C., & Lippman, M. (1988). Survival hazards analysis in first recurrent breast cancer patients: Seven-year follow-up. *Psychosomatic Medicine, 50*(5), 520-528.

Levy, B. R., Slade, M. D., Kunkel, S. R., & Kasl, S. V. (2002). Longevity increased by positive self-perceptions of aging. *Journal of Personality and Social Psychology, 83*(2), 261-270.

Liu, Z. W., Yu, Y., Hu, M., Liu, H. M., Zhou, L., & Xiao, S. Y. (2016). PHQ-9 and PHQ-2 for screening depression in Chinese rural elderly. *PloS one, 11*(3), e0151042. doi:10.1371/journal.pone.0151042

Lou, V. W., Chan, C. L., Leung, D. Y., Fok, S. S., Tsui, A. K., Gui, S. X., & Shek, D. T. (2012). Conceptualizing and assessing spirituality among Chinese elders: Towards evidence-based practice. *Generations Review, 22*(1). 15-16

Löwe, B., Unützer, J., Callahan, C. M., Perkins, A. J., & Kroenke, K. (2004). Monitoring depression treatment outcomes with the patient health questionnaire-9. *Medical Care, 42*(12), 1194-1201.

Lucas, R. E., Diener, E., & Suh, E. (1996). Discriminant validity of well-being measures. *Journal of Personality and Social Psychology, 71*(3), 616-628.

Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: Does happiness lead to success?. *Psychological Bulletin, 131*(6), 803-855.

Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research, 46*(2), 137-155.

Lyubomirsky, S., & Ross, L. (1997). Hedonic consequences of social comparison: A contrast of happy and unhappy people. *Journal of Personality and Social Psychology, 73*(6), 1141-1157.

MacVean, M. (2015, July 11). Why Loma Linda residents live longer than the rest of us: They treat the body like a temple. The Los Angeles Times. Retrieved from http://www.latimes.com/health/la-he-blue-zone-loma-linda-20150711-story.html

Manea, L., Gilbody, S., & McMillan, D. (2012). Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): A meta-analysis. *Canadian Medical Association Journal, 184*(3), E191-E196.

Martin, A., Rief, W., Klaiberg, A., & Braehler, E. (2006). Validity of the brief patient health questionnaire mood scale (PHQ-9) in the general population. *General Hospital Psychiatry, 28*(1), 71-77.

McAuley, E., Konopack, J. F., Motl, R. W., Morris, K. S., Doerksen, S. E., & Rosengren, K. R. (2006). Physical activity and quality of life in older adults: Influence of health status and self-efficacy. *Annals of Behavioral Medicine, 31*(1), 99-103.

McCullough, M. E., Hoyt, W. T., Larson, D. B., Koenig, H. G., & Thoresen, C. (2000). Religious involvement and mortality: A meta-analytic review. *Health Psychology, 19*(3), 211-222.

McDowell, I. & Newell, C. (1996). *Measuring health, a guide to rating scales and questionnaires*. New York, NY: Oxford University Press.

Miller, W. R., & Thoresen, C. E. (2003). Spirituality, religion, and health: An emerging research field. *American Psychologist, 58*(1), 24-35.

Morse, S., & Gergen, K. J. (1970). Social comparison, self-consistency, and the concept of self. *Journal of Personality and Social Psychology, 16*(1), 148-156.

Murphy, D. F., McDonald, A., Power, C., Unwin, A., & MacSullivan, R. (1987). Measurement of pain: A comparison of the visual analogue with a nonvisual analogue scale. *The Clinical Journal of Pain, 3*(4), 197-200.

Myers, D. G., & Diener, E. (1995). Who is happy?. *Psychological Science, 6*(1), 10-19.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Operario, D., Adler, N. E., & Williams, D. R. (2004). Subjective social status: Reliability and predictive utility for global health. *Psychology and Health, 19*(2), 237-246.

Ostir, G. V., Markides, K. S., Black, S. A., & Goodwin, J. S. (2000). Emotional well-being predicts subsequent functional independence and survival. *Journal of the American Geriatrics Society, 48*(5), 473-478.

Ouweneel, P., & Veenhoven, R. (1991). Cross-national differences in happiness: Cultural bias or societal quality?. In N. Bleichrodt & P. J. Drenth (Eds.), *Contemporary issues in cross-cultural psychology* (pp. 168-184). Amsterdam, The Netherlands: Swets & Zeitlinger.

Palmore, E., & Kivett, V. (1977). Change in life satisfaction: A longitudinal study of persons aged 46–70. *Journal of Gerontology, 32*(3), 311-316. doi:10.1093/geronj/32.3.311

Pargament, .K. I., & Mahoney, A. (2002). Spirituality: Discovering and conserving the sacred. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 646-659). New York: Oxford University Press.

Paul-Dauphin, A., Guillemin, F., Virion, J. M., & Briançon, S. (1999). Bias and precision in visual analogue scales: A randomized controlled trial. *American Journal of Epidemiology, 150*(10), 1117-1127.

Perneger, T. V., Hudelson, P. M., & Bovier, P. A. (2004). Health and happiness in young Swiss adults. *Quality of Life Research, 13*(1), 171-178.

Peterson, C. (2005). Authentic Happiness Inventory Questionnaire. Measures Overall Happiness. Michigan: University of Michigan.

Pethtel, O., & Chen, Y. (2010). Cross-cultural aging in cognitive and affective components of subjective well-being. *Psychology and Aging*, *25*(3), 725-729.

Phelan, E., Williams, B., Meeker, K., Bonn, K., Frederick, J., LoGerfo, J., & Snowden, M. (2010). A study of the diagnostic accuracy of the PHQ-9 in primary care elderly. *BMC Family Practice, 11*(1), 1-9.

Powers, D. V., Cramer, R. J., & Grubka, J. M. (2007). Spirituality, life stress, and affective well-being. *Journal of Psychology and Theology, 35*(3), 235-243.

Rajaram, S., Valls-Pedret, C., Cofán, M., Sabaté, J., Serra-Mir, M., Pérez-Heras, A. M., ... & Doménech, M. (2017). The Walnuts and Healthy Aging Study (WAHA): Protocol for a nutritional intervention trial with walnuts on brain aging. *Frontiers in Aging Neuroscience, 8*, 1-12. doi:10.3389/fnagi.2016.00333

Reske, M., Kellermann, T., Habel, U., Shah, N. J., Backes, V., von Wilmsdorff, M., ... & Schneider, F. (2007). Stability of emotional dysfunctions? A long-term fMRI study in first-episode schizophrenia. *Journal of Psychiatric Research, 41*(11), 918-927.

Resnick, B., & Nahm, E. S. (2001). Reliability and validity testing of the revised 12-item Short-Form Health Survey in older adults. *Journal of Nursing Measurement, 9*(2), 151-161.

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin, 27*(2), 151-161.

Rodda, J., Walker, Z., & Carter, J. (2011). Depression in older adults. *BMJ, 343,* 683-687.

Roese, N. J., & Jamieson, D. W. (1993). Twenty years of bogus pipeline research: A critical review and meta-analysis. *Psychological Bulletin, 114*(2), 363-375.

Rohwedder, S. & Willis, R. J. (2010). Mental retirement. *Journal of Economic Perspectives, 24*(1), 119-138.

Roness, A., Mykletun, A., & Dahl, A. A. (2005). Help-seeking behaviour in patients with anxiety disorder and depression. *Acta Psychiatrica Scandinavica, 111*(1), 51-58.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (2nd ed.). Newbury Park, CA: Sage.

Rowold, J. (2011). Effects of spiritual well-being on subsequent happiness, psychological well-being, and stress. *Journal of Religion and Health, 50*(4), 950-963.

Russel, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology, 57*(3), 493-502.

Salyers, M. P., Bosworth, H. B., Swanson, J. W., Lamb-Pagone, J., & Osher, F. C. (2000). Reliability and validity of the SF-12 health survey among people with severe mental illness. *Medical Care, 38*(11), 1141-1150.

Sanderson, K., & Andrews, G. (2002). The SF-12 in the Australian population: Cross-validation of item selection. *Australian and New Zealand Journal of Public Health, 26*(4), 343-345.

Sandvik, E., Diener, E., & Seidlitz, L. (1993). Subjective well-being: The convergence and stability of self-report and non-self-report measures. *Journal of Personality, 61*(3), 317-342.

Sawatzky, R., Ratner, P. A., Johnson, J. L., Kopec, J. A., & Zumbo, B. D. (2010). Self-reported physical and mental health status and quality of life in adolescents: A latent variable mediation model. *Health and Quality of Life Outcomes, 8*(17), 1-11. doi:10.1186/1477-7525-8-17

Scheier, M. F., Matthews, K. A., Owens, J. F., Magovern, G. J., Lefebvre, R. C., Abbott, R. A., & Carver, C. S. (1989). Dispositional optimism and recovery from coronary artery bypass surgery: The beneficial effects on physical and psychological well-being. *Journal of Personality and Social Psychology, 57*(6), 1024-1040.

Schütz, E., Sailer, U., Al Nima, A., Rosenberg, P., Arntén, A. C. A., Archer, T., & Garcia, D. (2013). The affective profiles in the USA: Happiness, depression, life satisfaction, and happiness-increasing strategies. *PeerJ, 1,* e156. doi:10.7717/peerj.156

Shahabi, L., Powell, L. H., Musick, M. A., Pargament, K. I., Thoresen, C. E., Williams, D., ... & Ory, M. A. (2002). Correlates of self-perceptions of spirituality in American adults. *Annals of Behavioral Medicine*, *24*(1), 59-68.

Simon, G. E., Goldberg, D., Tiemens, B. G., & Ustun, T. B. (1999). Outcomes of recognized and unrecognized depression in an international primary care study. *General Hospital Psychiatry*, *21*(2), 97-105.

Smarr, K. L., & Keefer, A. L. (2011). Measures of depression and depressive symptoms: Beck Depression Inventory-II (BDI-II), Center for Epidemiologic Studies Depression Scale (CES-D), Geriatric Depression Scale (GDS), Hospital Anxiety and Depression Scale (HADS), and Patient Health Questionnaire-9 (PHQ-9). *Arthritis Care & Research*, *63*(S11), S454–S466.

Soubelet, A., & Salthouse, T. A. (2011). Influence of social desirability on age differences in self-reports of mood and personality. *Journal of Personality*, *79*(4), 741-762.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Patient Health Questionnaire Primary Care Study Group. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *JAMA*, *282*(18), 1737-1744.

Spitzer, R. L., Williams, J. B., Kroenke, K., Hornyak, R., McMurray, J., & Patient Health Questionnaire Obstetrics-Gynecology Study Group. (2000). Validity and utility of the PRIME-MD patient health questionnaire in assessment of 3000 obstetric-gynecologic patients: The PRIME-MD Patient Health Questionnaire Obstetrics-Gynecology Study. *American Journal of Obstetrics and Gynecology*, *183*(3), 759-769.

Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The meaning in life questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology*, *53*(1), 80-93.

Steptoe, A., & Wardle, J. (2005). Positive affect and biological function in everyday life. *Neurobiology of Aging*, *26*(1), 108-112.

Stone, A. A., & Mackie, C. (Eds.). (2014). *Subjective well-being: Measuring happiness, suffering, and other dimensions of experience*. Washington, DC: National Academies Press.

Takeyachi, Y., Konno, S. I., Otani, K., Yamauchi, K., Takahashi, I., Suzukamo, Y., & Kikuchi, S. I. (2003). Correlation of low back pain with functional status, general health perception, social participation, subjective happiness, and patient satisfaction. *Spine, 28*(13), 1461-1466.

Tatarkiewicz, W. (1976). *Analysis of happiness*. The Hague, The Netherlands: Martinus Nijhoff.

U.S. Department of Labor, Bureau of Labor Statistics. (2013). *Labor force projections to 2022: The labor force participation rate continues to fall*. Retrieved from http://www.bls.gov/opub/mlr/2013/article/labor-force-projections-to-2022-the-labor-force-participation-rate-continues-to-fall.htm

U.S. Social Security Administration. (2013). *The annual report of the board of trustees of the federal old-age and survivors insurance and federal disability insurance trust funds*. Washington, DC: U.S. Government Printing Office.

U.S. Social Security Administration. (2016). Calculators: Life Expectancy. Retrieved from https://www.ssa.gov/planners/lifeexpectancy.html

Veenhoven, R. (2000). Freedom and happiness: A comparative study of 46 nations in the early 1990s. *Subjective wellbeing across culture.* MIT Press: Massachusetts.

Veenhoven, R. (2012, August). Social development and happiness in nations. Presented at the Presentation at the International Institute of Social Studies conference 'Taking stock: Measuring social development', The Hague, The Netherlands.

Veenhoven, R. (2008). Healthy happiness: Effects of happiness on physical health and the consequences for preventive health care. *Journal of Happiness Studies, 9*(3), 449-469.

Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: How good are single-item measures?. *Journal of Applied Psychology, 82*(2), 247-252.

Ware Jr, J. E., Kosinski, M., & Keller, S. D. (1996). A 12-Item Short-Form Health Survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care, 34*(3), 220-233.

Watkins, D., Akande, A., Fleming, J., Ismail, M., Lefner, K., Regmi, M., & ... Wondimu, H. (1998). Cultural dimensions, gender, and the nature of self-concept: A fourteen-country study. *International Journal of Psychology, 33*(1), 17-31. doi:10.1080/002075998400583

Watkins, P. C., Woodward, K., Stone, T., & Kolts, R. L. (2003). Gratitude and happiness: Development of a measure of gratitude, and relationships with subjective well-being. *Social Behavior and Personality: An International Journal, 31*(5), 431-451.

Watson, D. (1988). The vicissitudes of mood measurement: Effects of varying descriptors, time frames, and response formats on measures of positive and negative affect. *Journal of Personality and Social Psychology, 55*(1), 128-141.

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology, 97*(3), 346-353.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070.

Watson, D., & Pennebaker, J. W. (1989). Health complaints, stress, and distress: Exploring the central role of negative affectivity. *Psychological Review, 96*(2), 234-254.

Wheeler, L., & Miyake, K. (1992). Social comparison in everyday life. *Journal of Personality and Social Psychology, 62*(5), 760-773.

Whitehead, J. C., Groothuis, P. A., & Blomquist, G. C. (1993). Testing for non-response and sample selection bias in contingent valuation: Analysis of a combination phone/mail survey. *Economics Letters, 41*(2), 215-220.

Wilson, I. B., & Cleary, P. D. (1995). Linking clinical variables with health-related quality of life: A conceptual model of patient outcomes. *JAMA, 273*(1), 59-65.

Wood, J. V., Taylor, S. E., & Lichtman, R. R. (1985). Social comparison in adjustment to breast cancer. *Journal of Personality and Social Psychology, 49*(5), 1169-1183.

Zautra, A. J., Burleson, M. H., Smith, C. A., Blalock, S. J., Wallston, K. A., DeVellis, R. F., ... & Smith, T. W. (1995). Arthritis and perceptions of quality of life: An examination of positive and negative affect in rheumatoid arthritis patients. *Health Psychology, 14*(5), 399-408.

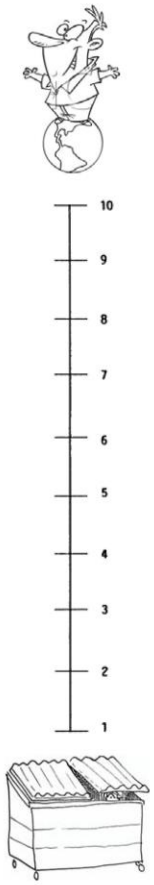
APPENDIX A

THE AM HAPPY SCALE

This scale is designed to measure happiness levels of people in the United States. At the top of the scale are the people who are the most happy and feeling as if they are on top of the world. At the bottom of the scale are the people who are most unhappy and feeling as though they are down in the dumps.

Using this scale, what is your current level of happiness? Please mark the scale to reflect how happy you are in general in comparison to other people in the United States.

1 = down in the dumps, and 10 = on top of the world



APPENDIX B

THE POSITIVE AND NEGATIVE AFFECT SCHEDULE

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Indicate to what extent you have felt this way during the past week.

1	2	3	4	5
very slightly or not at all	a little	moderately	quite a bit	extremely

__ interested
__ distressed
__ excited
__ upset
__ strong
__ guilty
__ scared
__ hostile
__ enthusiastic
__ proud
__ irritable
__ alert
__ ashamed
__ inspired
__ nervous
__ determined
__ attentive
__ jittery
__ active
__ afraid

APPENDIX C

SPIRITUALITY INDEX OF WELL-BEING

Which statement best describes your feelings and choices? Indicate how you feel about each statement by circling the appropriate number.

Strongly Agree 1	Agree 2	Neither Agree nor Disagree 3	Disagree 4	Strongly Disagree 5

1. There is not much I can do to help myself	1	2	3	4	5
2. Often, there is no way I can complete what I started	1	2	3	4	5
3. I can't begin to understand my problems	1	2	3	4	5
4. I am overwhelmed when I have personal difficulties and problems	1	2	3	4	5
5. I don't know how to begin to solve my problems	1	2	3	4	5
6. There is not much I can do to make a difference in my life	1	2	3	4	5
7. I haven't found my life's purpose yet	1	2	3	4	5
8. I don't know who I am, where I came from, or where I am going	1	2	3	4	5
9. I have a lack of purpose in my life	1	2	3	4	5
10. In this world, I don't know where I fit in	1	2	3	4	5
11. I am far from understanding the meaning of life	1	2	3	4	5
12. There is a great void in my life at this time	1	2	3	4	5

APPENDIX D

PATIENT HEALTH QUESTIONNAIRE

Over the last two weeks how often have you been bothered by any of the following problems?

(Please circle the numbers to indicate your answers)	Not at all	Several days	More than half the days	Nearly every day
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed, or hopeless	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself – or that you are a failure or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed. Or the opposite – being so fidgety or restless that you have been moving around a lot more than usual	0	1	2	3
9. Thoughts that you would be better off dead, or of hurting yourself	0	1	2	3
10. If you circled any problems, how difficult have these problems made it for you to do your work, take care of things at home, or get along with people?	Not difficult at all _____ Somewhat difficult _____ Very difficult _____ Extremely difficult _____			

APPENDIX E

SHORT-FORM HEALTH SURVEY

This information will help your doctors keep track of how you feel and how well you are able to do your usual activities. Answer every question by placing a check mark on the line in front of the appropriate answer. If you are unsure about how to answer a question, please give the best answer you can and make a written comment beside your answer.

1. In general, would you say your health is:

_____ Excellent (1)

_____ Very Good (2)

_____ Good (3)

_____ Fair (4)

_____ Poor (5)

The following two questions are about activities you might do during a typical day. Does YOUR HEALTH NOW LIMIT YOU in these activities? If so, how much?

2. MODERATE ACTIVITIES, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf:

_____ Yes, Limited A Lot (1)

_____ Yes, Limited A Little (2)

_____ No, Not Limited At All (3)

3. Climbing SEVERAL flights of stairs:

_____ Yes, Limited A Lot (1)

_____ Yes, Limited A Little (2)

_____ No, Not Limited At All (3)

During the PAST 4 WEEKS have you had any of the following problems with your work or other regular activities AS A RESULT OF YOUR PHYSICAL HEALTH?

4. ACCOMPLISHED LESS than you would like:

_____ Yes (1)

_____ No (2)

5. Were limited in the KIND of work or other activities:

_____ Yes (1)

_____ No (2)

During the PAST 4 WEEKS, were you limited in the kind of work you do or other regular activities AS A RESULT OF ANY EMOTIONAL PROBLEMS (such as feeling depressed or anxious)?

6. ACCOMPLISHED LESS than you would like:

_____ Yes (1)

_____ No (2)

7. Didn't do work or other activities as CAREFULLY as usual:

_____ Yes (1)

_____ No (2)

8. During the PAST 4 WEEKS, how much did PAIN interfere with your normal work (including both work outside the home and housework)?

_____ Not At All (1)

_____ A Little Bit (2)

_____ Moderately (3)

_____ Quite A Bit (4)

_____ Extremely (5)

The next three questions are about how you feel and how things have been DURING THE PAST 4 WEEKS. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the PAST 4 WEEKS –

9. Have you felt calm and peaceful?

_____ All of the Time (1) _____ Some of the Time (4)

_____ Most of the Time (2) _____ A Little of the Time (5)

_____ A Good Bit of the Time (3) _____ None of the Time (6)

10. Did you have a lot of energy?

_____ All of the Time (1) _____ Some of the Time (4)

_____ Most of the Time (2) _____ A Little of the Time (5)

_____ A Good Bit of the Time (3) _____ None of the Time (6)

11. Have you felt downhearted and blue?

_____ All of the Time (1) _____ Some of the Time (4)

_____ Most of the Time (2) _____ A Little of the Time (5)

_____ A Good Bit of the Time (3) _____ None of the Time (6)

12. During the PAST 4 WEEKS, how much of the time has your PHYSICAL HEALTH OR EMOTIONAL PROBLEMS interfered with your social activities (like visiting with friends, relatives, etc.)?

_____ All of the Time (1) _____ Some of the Time (4)

_____ Most of the Time (2) _____ A Little of the Time (5)

_____ A Good Bit of the Time (3) _____ None of the Time (6)