

Florida Institute of Technology

Scholarship Repository @ Florida Tech

Theses and Dissertations

1-2018

An Examination of the Scale of Accurate Personality Prediction's (SAPP) Test-Retest Reliability Across a Four-Week Period

Taylor Catherine Anderson

Follow this and additional works at: <https://repository.fit.edu/etd>



Part of the [Clinical Psychology Commons](#)

An Examination of the Scale of Accurate
Personality Prediction's (SAPP) Test-Retest Reliability
Across a Four-Week Period

By

Taylor Catherine Anderson

Bachelor of Arts
Psychology
St. Joseph's University
2014

Master of Science
In Clinical Psychology
Florida Institute of Technology
2016

A doctoral research project
submitted to School of Psychology at
Florida Institute of Technology
in partial fulfillment of the requirements
for the degree of

Doctor of Psychology
in
Clinical Psychology

Melbourne, Florida
January 2018

We the undersigned committee hereby approve the attached doctoral level project,
“An Examination of the Scale of Accurate Personality Prediction’s (SAPP) Test-
Retest Reliability Across a Four-Week Period,” by Taylor Catherine Anderson.

Philip D. Farber, Ph.D.
Associate Professor
College of Psychology and Liberal Arts
Committee Chair

Richard T. Elmore, Ph.D.
Associate Professor
College of Psychology and Liberal Arts
Committee Member

John Frongillo, Ph.D.
Assistant Professor
School of Arts and Communication
Committee Member

Mary Beth Kenkel, Ph.D.
Professor, Dean
College of Psychology and Liberal Arts

ABSTRACT

TITLE: An Examination of the Scale of Accurate Personality Prediction's (SAPP)

Test-Retest Reliability Across a Four-Week Period

AUTHOR: Taylor Catherine Anderson, M.S.

MAJOR ADVISOR: Philip D. Farber, Ph.D.

The questions surrounding how well one can truly know him/herself dates as far back as Ancient Greek times, and answering the question has been a difficult task given the complexity of human beings, and a somewhat subjective term such as the “self.” As science and technology continue to advance, so do the attempts to accurately define self-knowledge. Miller (2000) took defining self-knowledge a step further, and developed a test derived from the Sixteen Personality Prediction Questionnaire (16PF) to measure individuals’ level of self-knowledge how well they know themselves, based on their ability to accurately predict their scores on the 16PF. Several reliability and validity studies have been conducted, and the present study serves as an additional assessment to further explore the stability of the SAPP regarding test-retest reliability. To assess the test-retest reliability of the SAPP, the SAPP scores of 29 participants were derived during initial testing trials, and again four weeks later. A Pearson correlation, which was used to determine if there is a significant correlation between the two trials of SAPP scores, revealed a significant strong correlation ($r^2 = .584$, $p < .01$). The implications of these results as well as limitations of the study are discussed.

TABLE OF CONTENTS

Abstract.....	iii
List of Tables.....	v
Literature Review.....	1
Statement of Purpose.....	34
Method.....	35
Results.....	38
Discussion.....	41
References.....	46
Tables & Appendices.....	51

LIST OF TABLES

TABLE

1. Demographic Percentages
2. Test-Retest Reliability Data
3. SAPP Ranges, Means, and Standard Deviations

ACKNOWLEDGEMENTS

I would like to start off by thanking Dr. Farber for his support and guidance throughout this process. His guidance and support extends well beyond this project, as he has been instrumental in my overall profession development, and has continuously provided me emotional support as well. Through his encouragement, support, and genuineness, I felt capable of overcoming the ever-changing stressors of graduate school. He was always willing to give open and honest advice, while simultaneously supportively challenging me to discover the answers on my own. I am truly thankful for his presence and encouragement throughout my graduate career. I would also like to thank Dr. Elmore and Dr. Frongillo for taking time to serve on my committee, and provide feedback and support. I would like to thank each one of my family and friends who have supported me throughout my graduate career with endless encouragement and love. Especially Sean Cunningham, Lisa Stewart, Lauren Dilullo, Damla Til Ogut, and my Mom and Dad. Each of you have been a listening ear, a shoulder to cry on, and most importantly there for me when I needed it most, unconditionally. There are not enough words for my gratitude. I would also like to thank the staff at PSI/PAN, and particularly to Michael Stowers, Psy.D., who made this research possible by providing grant in-kind funding for it, in the form of on-line administrations of the 16PF, as well as on-line scoring, and access to the necessary data files.

LITERATURE REVIEW

WHAT IS PERSONALITY?

The term personality is derived from the Latin root *persona*, meaning a mask, specifically a theatrical mask (Bishop, 2009). The concept of personality has vacillated and been explained differently from theorist to theorist, and continues to do so presently. It is likely that a contributing factor into the difficulty of describing personality is that one cannot easily and objectively quantify, and define such a broad, and often subjective, term.

Psychoanalytic Conceptualizations of Personality

The traditional psychoanalytic perspective of personality development has been heavily reliant on Sigmund Freud's contributions to the field. The theory is constructed of various segments, including the genetic or biological side, the psychological side, and society's influence on personality. Freud identified the id, the ego, and the super ego as the basic processes of the mind. The id is the most basic system, composed of physical drives and primitive psychological needs, and operates on the Freudian coined phrase, the pleasure principle. The ego is able to regulate the id's impulsive and instinctual needs by delaying gratification through self-control. The super ego is society's influential component, and it is the center ground between balancing the id and the ego through the morality principle, incorporating society's norms and values into behavior (Allen, 2003).

Freud indicated that personality in adulthood is formed largely by factors experienced during childhood. He postulated infants are born with a sexual energy, or libido, which evolves via the psychosexual stages, and is embedded in a visceral process of humans. The psychosexual stages are consecutive, and Freud identified them as the oral, anal, phallic, and genital stage (Hjelle & Ziegler, 1992).

The oral stage is the narcissistic stage beginning at birth through 18 months of age. This stage encompasses fulfilling needs surrounding the child's mouth, such as sucking, biting, and chewing. Freud argued there are two personality types in adulthood which represent inadequate development of this stage. The oral-receptive personality type, which is characterized by the person who is gullible or willing to "swallow" anything, and who is likely to show dependency characteristics. The oral-aggressive personality type is displayed in individuals who are overly aggressive verbally. Following the oral stage, the anal stage begins and the child begins the process of toilet training, and as a result, the two personality types emerging in this stage are the anal-retentive individuals, who are overly controlled and delay satisfaction, and the anal-expulsive individuals who are less inhibited, and display a disregard for orderliness and rules. The phallic stage begins around three years of age, and its satisfaction is centered around the genital area, with the child identifying with the same sex parent. Freud suggested the adult personality type associated within this stage is a promiscuous individual, with low ego-strength, and one who may tend to not adhere to their biological gender norms. The genital stage is Freud's final stage, where mature sexual love is developed

beginning at puberty, and it involves a less “me focused” approach, and instead directs attention to another person (Allen, 2003).

Freud proposed as individuals move through the psychosexual stages, their sexual energy move to more mature objects. However, a person can become stagnant in any stage where they have experienced trauma. This stagnancy would be evident when their sexual energy is representative of an earlier stage of development, and when they become fixated on a less mature or acceptable object. Freud offered his personality typology according to this idea, stating people who display a disparate amount of sexual energy subsumed in one particular stage will display personality traits of this age (Abrams, Ellis, & Abrams, 2009).

Behavioral Conceptualizations of Personality

In contrast to Freud’s “within the person” theoretical concept of personality, behaviorists view personality as a broad term that mostly consists of a set of learned skills acquired over the course of a life time. Within this learning model, it is the environment that shapes an individual, rather than internal mental events. The idea that individual’s behavior and personality is contingent upon their environmental factors is most notably exemplified by Ivan Pavlov and B.F. Skinner (Hjelle & Ziegler, 1992).

Pavlov, who was not a psychologist, and his learning contributions to the field of personality psychology were a by-product of his work with dogs, and helped develop the concept of classical conditioning. Pavlov first discovered dogs being fed through a feeding tube would salivate during feedings, and he identified

this salivating response as an unconditioned response. He further discovered that by pairing a neutral stimulus, such as a sound, with feeding over time would lead the dog to salivate to the sound alone. Pavlov coined this salivation to the sound as a conditioned reflex, or response (Abrams et al., 2009).

Skinner's research focused on respondent and operant behavior and the environmental consequences to it, which has become to be known as operant conditioning. In a manner similar to Pavlov, Skinner suggested that behavior is determined, predictable, and environmentally controlled. Therefore, Skinner would likely describe the study of personality as examining distinct patterns of links between various learned behavior and their reinforcing consequences. Skinner's contributions to the field of personality and psychology are particularly noted because of his empirically based research, as he has provided extensive experimental data to support his theory of how most human behavior is established and modified. Skinner's work varied from Pavlov's by focusing on the stimulus that governs behavior, specifically that some behaviors are not reflexes, and instead are influenced by a cueing stimulus, and the rewarding or punishing response to it (Hjelle & Ziegler, 1992). Therefore, strict behaviorists would render free will and consciousness as meaningless constructions as a result of their deterministic view of behavior.

The Social Cognitive Conceptualizations of Personality

The social cognitive model followed the behaviorist tradition, and has postulated that personality cannot be considered separate from the social

environment from which it develops, and it emphasizes individuals' cognition regarding their social environment. Social learning suggests interactions with people and elements in the environment assist individuals in acquiring useful information. Specifically, social cognitive theorists have emphasized expectancies (Allen, 2003).

Julian Rotter and Albert Bandura are some of the first among many individuals to study social learning theory. Rotter greatly influenced the latter part of the 20th century with his work emphasizing personality as a true construct, in contrast to the behaviorists view. Rotter identified individual's behavior as having significant complexity, which could not be fully explained by a just stimulus control and subsequent reinforcement contingencies. Rotter's model instead heavily emphasizes expectancies and anticipation of being rewarded, which then directly influences the direction of behavior. His model allows for individuals to have a higher level of freedom and choice in behavior. Furthermore, he has identified behaviors not just from the individual's successes and failures, but from observing those of others, nulling the reinforcement component of behavior (Abrams et al., 2009).

In a somewhat similar manner, Bandura depicted individuals as agents, or originators, of experience. He describes agency as the capacity to act and make things happen, occurring within a larger segment of sociostructurally bidirectional influences, both producers and products of their environment. Bandura discovered and coined the triadic reciprocal causation model, suggesting human behavior is an

interlocking network of behavior, cognitive, and environmental factors (Engler, 2005). His model relies heavily on the concept of observational learning, along with expected as well as actual reinforcement contingencies.

Humanistic and Existential Conceptualizations of Personality

Humanistic and existential (E/H) psychologists attempted to steer away from behaviorism, which they viewed as reducing individuals only to their observable and measurable behaviors, and psychoanalysis, which they viewed as studying primarily the non-observable and unconscious elements of individuals. E/H thought was developed to study what it means to be human, with an emphasis on conscious processes. E/H approaches within psychology are largely seen as an alternative to the reductionistic and mechanistic approaches inherent within the experimental and scientific models that have been so successful in understanding the world of the natural sciences. Rather than having their birth within an objective/scientific template, E/H theories had their beginnings with philosophical thought, particularly within a philosophy of being. For an existentialist, what individuals choose to do in the present moment is not driven by unconscious processes or learning contingencies only, but also by free and independent choices at any given moment or experience. No doubt, there are limitations to human choices, but what is most important in understanding human behavior are the choices that humans make or do not make (Abrams et al., 2009).

Abraham Maslow, considered a humanistic psychologist by many, was widely known for his contribution of the hierarchy of needs. The hierarchy of needs postulates that needs must be met in a specific order. First and foremost, is biological needs including food, water and air. Once biological needs are met, an individual must fulfill safety needs including physical security, followed by needs social attachments and social approval. The next stage includes the needs for achievement and self-esteem, and lastly for self-actualization. According to Maslow, self-actualization is an intrinsic tendency for individuals to enhance their potential assuming all of the basic needs preceding it are fulfilled. Maslow's concept of self-actualization became a central component to the humanistic movement (Abrams et al., 2009).

Rollo May, a prominent existential psychologist, focused much of his work on introducing powerlessness, anxiety, and the loss of values. May believed a focal point to many problems is a feeling of powerlessness, whether in the face of social, cultural, or economic problems, and the inability to be effective in the face of the aforementioned problems. Anxiety, as proposed by May is "the apprehension cued off by a threat to some value that the individual holds essential to his or her existence as a person" (as cited in Engler, 2005, p. 356), and is an inevitable construct held by every individual. May goes on to describe a loss of values as a prominent source of problems faced by all individuals. Specifically, May describes a loss of values as more recently problematic compared to earlier times, such as the Renaissance, where personal values were held in competitive prestige measures in

terms of work and financial success. May disputes these values as relevant today, as a result of the shift in the modern world where one works with others in order to thrive. Therefore, individual competition now creates problems where it previously did not. In regards to personality research, May discusses confronting paradoxes. For example, the concept of love, has been seen as the answer to many problems, but has now become the problem because people are unable to love. May offers a solution of experiencing and rediscovering these initial paradoxes as a means to solve them, such as rediscovering care (Engler, 2005).

THE SELF

The concept of the self cannot be solely encompassed by a single definition. For every definition, there is an equally conflicting explanation. The definitions are not always compatible, nor related. Is the self concrete or abstract, material or immaterial, permanent or temporary, natural or human constructed (Olson, 1998)? Given the immense conflicting theories and definitions, one may consider the self as a combination of facets such as self-awareness, self-esteem, identity, self-concept, and self-control. One component each of the aforementioned facets share in common is the ability to self-reflect, which can be considered a necessity for countless human behaviors wherein many of these behaviors would be unable

to occur without it. On the other hand, many human behaviors do not require self-reflection, such as those that occur nonconsciously or automatically. There are behaviors which can occur without self-reflection, however when self-reflection is added to the same behaviors, they are substantially altered, such as mating, cooperation, interpersonal communication, nonsocial emotions such as sadness and fear. These later behaviors do not necessitate self-reflection, but are influenced greatly by it (Leary & Tangney, 2005).

The self can be considered malleable, as this can be seen when examining choices an individual makes which may feel identity congruent in one situation, yet may not feel congruent in another situation to the same individual. This flexibility is an important component of the self, because by nature, individuals are reasoners, aware of contextual allowances and restrictions, and able to make choices based on situational nuances. Given all of the grey areas, intricacies, and conflicting definitions of the self, it may be hard to identify reasons to study the self, yet making sense of oneself, learning who one is, were, may become, or how to choose a path to take, is a core “self-project” believed to be possessed by all or most individuals. Learning about oneself, can help individuals make sense of their environment (Oyserman, Elmore, & Smith, 2012).

Work surrounding the study of the self date back further than Plato in 428-347 B.C.E., with much of the earlier works examining questions about the self, reflexive consciousness, and identity. However, for the next roughly two thousand years, theories of the self were discussed within religious and theologically

constraints. During the enlightenment, many significant philosophers discussed problems of the self. It was not until 1890 that William James offered a detailed psychological discussion of the self. James laid a strong conceptual foundation of the self, and emphasized the importance of the self in understanding human behaviors. Nonetheless, it would still be many years before the self would be considered an important component of studying human behavior, because the behaviorism movement that followed James did not delve into the “invisible internal entities.” There were a handful of influential theorists who emphasized the importance of the self during the early and mid-20th century. Among sociologists, Charles Horton Cooley brought the self to attention in 1902, which was later advanced by George Herbert Mead in 1934. Furthermore, Ellsworth Faris and Herbert Blumer in 1937 advocated for the study of the self. In 1959, Erving Goffman’s work on self-presentation advanced another surge of curiosity in the self. At this time, neo-Freudians began to explore the idea of the self and offer alternative theories (Leary & Tangney, 2005).

Three advancements converged to increase focus given to the self by the second half of the 20th century. The first development grew in the context of self-esteem in the 1950’s and 1960’s. At that time, those psychologists studying the self, began developing self-report measures designed to measure various aspects of the self. The cognitive revolution in psychology brought about the second advancement in studying the self, which reintroduced the study of thoughts and internal control processes, by providing new models of how individuals attend to

and process information from a cognitive level. The third advancement came about through several publications providing measures of dispositional attributes related to the self. By the 1980's, the advancement of self-reported measures made the construct of the self somewhat more tangible, and allowed the self to emerge as a salient topic of examination and study (Leary & Tangney, 2005).

Theories and models of the self vary greatly, and tend to be quite variable even cross-culturally. Specifically, Western cultures such as the United States, Canada, and Western European countries tend to view the self individually, focusing on the personal self. Eastern cultures such as Japan, China, and India, on the other hand, tend to hold a collectivistic view point of the self, deemphasizing the individual self compared to Western cultures. However, it is important to note that one should not overgeneralize or overemphasize these differences, as there is greater within group variability than between groups (Robins, Norem, & Cheek, 1999).

Additionally, Leary and Tangney (2005), described five different types of the self, including the self as the total person, the self as personality, the self as experiencing subject (the self as "I"), the self as beliefs about oneself (the self as "me"), and the self as executive agent. The self as beliefs about oneself (the self as "me"), will be the main focus of the Scale of Accurate Personality Prediction (SAPP).

SELF-KNOWLEDGE

Self-knowledge followed a similar path throughout history as the study of personality, and the study of the self. The study of self-knowledge has been around since ancient Greece, and has been largely emphasized within religious and theological spheres. Within the field of psychology, self-knowledge seems to have been largely neglected as a core component of study up until only recently. The recent influx of studying self-knowledge can be seen across multiple psychological disciplines including sociology, personality, cognitive psychology, developmental psychology, clinical psychology, and neuroscience. Dual-process theories appear to be a commonality among all disciplines and continue to gain popularity in the study of self-knowledge (Vazire & Wilson, 2012).

According to the Stanford Encyclopedia of Philosophy, self-knowledge is “knowledge of one’s particular mental states, including one’s beliefs, desires and sensation”. Similarly, the Merriam-Webster online dictionary depicts self-knowledge as “the knowledge or understanding of one’s own capabilities, character, feelings and motivation” (Merriam Webster Online, n.d.). Often times, the goal in psychotherapy is to increase self-knowledge, and can reasonably be considered a valuable function in a person’s psychological functioning, as it allows for better self-awareness of one’s strengths and weaknesses, and allows individuals to realistically set goals for themselves.

If self-knowledge is often a central part of psychotherapy, how does one measure self-knowledge? The vast majority of research on self-knowledge has focused on the processes involved in gaining or having self-knowledge, rather than

the precision or measurement of an individual's self-knowledge. However, Vogt and Colvin (2005) developed a sophisticated method of measuring the accuracy of an individual's self-knowledge, and discussed the implications that such a method might play for mental health patients. Their assessment method relied on obtaining information from the individual themselves, from knowledgeable others of the individual, and from the observations of behavior within a laboratory setting. Participants described their individual personality characteristics using the California Adult Q-sort (CAQ), and through completion of the Revised NEO Personality Inventory (NEO-PI-R). Information on the individuals were also obtained by parents and friends, through the use of modified versions of the CAQ and the NEO-PI-R, which were designed to measure aspects of a target individual, and not themselves. The participant's behavior was coded using a 64-item Behavioral Q-sort by a team of trained coders. Lastly, individuals rated each other's behavior after watching a video-taped interaction. The authors provided psychometric evidence for this multifactorial assessment procedure, including internal consistency reliability, convergent and discriminate validity, and criterion-related validity. The authors determined accurate self descriptions of personality should predict behavior, and agree with personality ratings provided by knowledgeable others. Therefore, two types of measures were utilized to assess the accuracy, including correspondence between self-descriptions and behavioral ratings, and correspondence between self-descriptions and parent's ratings (Vogt & Colvin, 2005). The authors measure met satisfactory internal consistency reliability

and convergent validity, and some initial evidence was observed for discriminate and criterion related validity.

While this method is quite sophisticated for the study of self-knowledge, it appears that its main limitation is its limited feasibility to a clinical population. In a somewhat paradoxical manner, its strength in utilizing multiple access points renders it somewhat cumbersome and not easily transmutable to use with individuals. If true, then what is needed for the clinical arena is a measurement, or scale, that is reliable, valid, and above all, doable and practical. It was awareness of this need that originally led to the development of the Scale of Accurate Personality Prediction (SAPP), which is based on the well-known personality instrument, the 16PF, and which is the focus of this present work.

THE 16PF

The Sixteen Personality Factor Questionnaire (16PF) was created by Raymond Cattell as an assessment measure of personality. The 16 PF, in its fifth and most recent version, includes 187 questions for participants to answer on a likert scale, and is designed to measure the principal human personality factors. For example, questions were designed to highlight the way an individual perceives, interprets, thinks, and acts over a length of time, across a variety of circumstances. Through a series of obliquely-rotated, factor analyses, Cattell identified 15 personality traits (a 16th was later added to give a rough estimate of overall cognitive functioning), which are referred to as primary factors. Further factor

analyses of these 16 factors yielded five global factors, and then three validity scales were added. In its present 5th edition form, the 16PF's primary factor scales include Warmth, Reasoning, Emotional Stability (Ego Strength), Dominance (Assertiveness), Liveliness, Rule-Consciousness, Social Boldness, Sensitivity, Vigilance, Abstractedness (Impracticality), Privateness, and Apprehension, Openness to Change, Self-Reliance, Perfectionism (Compulsivity) and Tension. factors. The five global factors are Extraversion, Anxiety, Tough-mindedness, Independence, and Self-control. See Appendix A for the 16PF profile scale, which includes these 21 Factors. The three validity scales include Impression Management, Acquiescence, and Infrequency. Across the 21 Factors, raw scores are converted to Sten scores, which have a mean range of scores from 1 – 10, with all having a mean of 5.5 and a standard deviation of 3. What follows is a description of all the scales just mentioned. Lower end scores generally refer to Sten score of 1-3, and higher end scores to Sten scores of 7-10.

Primary Factors

For Factor A, the Warmth factor, descriptors on the low end (A-) include reserved, impersonal, and distant, versus the high end (A+), which describes those who are warm, outgoing, and attentive to others. Typically, people who score high on Warmth are those whom others gravitate toward, and who in return gravitate toward them. Individuals high on warmth are likely easygoing, adaptable, warmhearted, and attentive to others. Low scorers on Warmth typically prefer alone

time, and are often seen as reclusive or aloof. Individuals scoring low on this factor may be less inclined to be concerned about how their actions may affect others (Karson, Karson, & O'Dell, 1997).

The Reasoning factor's (Factor B) best descriptor includes on the low scoring end, concreteness, and on the high scoring end, abstractness. This factor is affected by an individual's intellectual capacity in the verbal, problem-solving sphere. Cattell's intent for this scale was for it to be a middle ground between most personality tests which do not include a measure of overall intelligence, and those intellectual tests which are quite time consuming. This scale is not solely affected by one's native intellectual capacity, but also by the individual's exposure to intellectually stimulating environments and academic problem-solving, as well as levels of concentration. High scores on this scale indicate superior reasoning ability and verbal facility. Low scores on this scale are not necessarily reflective of only low intellectual capacity, but also of motivation or fatigue issues (Karson, Karson, & O'Dell, 1997).

For Factor C, the Emotional Stability (Ego Strength) factor, descriptors on the high end include emotional stability, adaptiveness, maturity, and on the low end, reactivity, and emotional changeability. Low scorers on this scale can be associated with a wide variety of psychopathologies, symptoms, and adjustment problems. Low scores also can commonly suggest an individual with a desire to look bad, an attempt to cry for help, or genuine adjustment problems. Higher scores on this scale suggest an individual who is well adjusted, however, because it may

also indicate an individual trying to present him or herself as well adjusted, even if he or she is not, this scale should be interpreted along with the Impression Management validity scale (Karson, Karson, & O'Dell, 1997).

For, Factor E, the Dominance (Assertiveness) factor, descriptors on the low scoring end include a tendency to be deferential, cooperative, and avoidant of conflicts, and on the high scoring end suggest dominance, forcefulness, and assertiveness. For many, this scale has more to do with assertiveness than dominance. High scores on this scale suggest an individual who likes to let others know what he/she thinks, and one who may favor a need to have control over his/her space, work, and plans. Individuals high on this score do not hesitate to express themselves or stand up for their opinions. They tend to appear confident and competent. Low scorers tend to be more humble and submissive. Their lack of assertiveness may cause them to view others as disinterested in their needs (Karson, Karson, & O'Dell, 1997).

For Factor F, the Liveliness factor, descriptors on the low end include seriousness, restraint, and excessively carefulness, and on the high end include being lively, animated, and spontaneous. Cattell suggested that this scale relates to an individual's history of punishment, such that rewarding environments breed optimism, and punishing environments lead individuals to be more cautious in their approach to the world. As a result, scores on the low end suggest an individual who is considered to not necessarily be fearful of negative consequences, but rather

pessimistic about effort producing rewards. Scores on the higher ends could also suggest immaturity and impulsivity (Karson, Karson, & O'Dell, 1997).

Low scores on Factor G, the Rule-Consciousness factor, suggest expediency and nonconformity, and on the high end suggest rule-consciousness, and dutifulness. High scorers are also seen as being rule conscious, moralistic, and respectful of others. Significantly low scores on this scale suggest the individual does not appreciate any of the uses of rules in society, and is often seen commonly in adolescents who are in the process of differentiating from their parents. Low scorers closer to the center may be more likely to “cut their own path”, or rely on their experience to justify their behaviors (Karson, Karson, & O'Dell, 1997).

Factor H, the Social Boldness factor, high scores typically depict an individual who is socially bold, venturesome, and thick-skinned. Those scoring on the low end are often seen as shy, threat-sensitive, and timid. Individuals scoring on the lower end of this scale also tend to react to threats by withdrawing into shyness, and typically approach social interactions more cautiously. Individuals on the higher end of this scale tend to seek out experiences in the social environment, and are not deterred by threat-arousal (Karson, Karson & O'Dell, 1997; Russell & Karol, 1994).

High scores on the Sensitivity factor (Factor I) suggests an individual who tends to be sensitive, aesthetic, and sentimental. These individuals tend to be considered “tender”, using empathy as a means to relate to the world, in comparison to lower score individuals, who tend to be considered “tough minded”.

Lower scores in this scale suggest a person who may prefer to be more objective, and rely less on emotional experience when making decisions. It has been suggested that this scale could be labeled as the stereotypic “feminine” scale, however there appears to be a negative connotation with using the term as a label for this scale (Karson, Karson & O’Dell, 1997; Russell & Karol, 1994).

Factor L, known as the Vigilance scale, describes individuals on the high scoring end as vigilant, suspicious, skeptical, and wary. They may experience interpersonal tension resulting from their tendency to feel skeptical of the motives of others. They may often believe others take advantage of them, and as a result, these individuals can often appear hostile and angry toward others. Lower scorers tend to be seen as more trusting and accepting of others. They are unsuspecting, and tend to see the good in others (Karson, Karson & O’Dell, 1997; Russell & Karol, 1994).

The Abstractedness (Factor M) scale measures individual’s ability to solve problems, and more specifically, which components they attend to in decision-making. On the high scoring end, these individuals tend to be more imaginative and abstract in their thinking. They also can be seen as impractical as a result of attending to internal processes and fantasies. Those who obtain lower scores on this scale tend to be seen as practical, as they use a solution-oriented approach to problems. However, these individuals may be less likely to generate alternative solutions, and utilize less creative approaches (Karson, Karson & O’Dell, 1997; Russell & Karol, 1994).

Individuals who score high on the Privateness factor (Factor N) tend to be seen as private, discrete, and non-disclosing. These individuals prefer to remain reserved, and are likely guarded in their willingness to disclose and discuss their lives. Contrary to high scorers, low scores tend to be forthright, and typically are seen as more genuine. They are more inclined to be open and share information about themselves with others (Karson, Karson & O'Dell, 1997; Russell & Karol, 1994).

For Factor O, the Apprehension factor, high scorers are generally self-doubting and often worried. These individuals may tend to feel chronically guilty and anxious, often experiencing an overall sense of dread, and being highly sensitive to criticism. On the other hand, low scorers appear self-assured and unworried. They also tend to be self-confident and self-sufficient. At the extreme end, these individuals may deny negative components of the self (Karson, Karson & O'Dell, 1997; Russell & Karol, 1994).

Factor Q1, the openness to change factor, addresses primarily one's attitude toward change. Those on the high scoring end are more likely to experiment, and be more willing to change. Low scorers tend to prefer, and be more comfortable with, more traditional and familiar activities, choices, and solutions. It has been suggested that high scorers on this scale may have some levels of dissatisfaction in their current life circumstances which motivates them to desire change. Conversely, low scorers may display more levels of satisfaction in their circumstances relating

to their preference for remaining consistent in their approach (Karson, Karson & O'Dell, 1997; Russell & Karol, 1994).

The Self-Reliance factor (Factor Q2), is a measure of an individual's preference regarding task completion, specifically relating to individual or group work. Individuals scoring on the high end tend to prefer individual work, utilizing more self-reliance in tasks, and tend to be solitary and individualistic. Those on the low end are individuals who prefer group work, and may be more affiliative with others. It is important to note this scale is in reference to work, as opposed to social activities or preference (Karson, Karson & O'Dell, 1997; Russell & Karol, 1994).

Factor Q3, also referred to as the Perfectionism factor, tends to measure one's degree of tolerance for disorder. Those scoring higher on this scale tend to be organized and self-disciplined, preferring order, and as a result, often have perfectionistic tendencies. Others may view high scorers as impatient. Contrarily, low scores are more likely to tolerate disorder, and may be more flexible and patient than their high scoring counterparts (Karson, Karson & O'Dell, 1997; Russell & Karol, 1994).

Lastly, Factor Q4, the Tension factor, measures traits which include feeling tense, high energy, impatient, and driven on the high scoring end. Lower scores tend to display feeling relaxed, placid and patient. Higher scores on this scale can suggest conducive to the development of anxiety, and higher levels of subjective distress (Karson, Karson & O'Dell, 1997; Russell & Karol, 1994).

Global Factors

The first global factor, Extraversion (EX) measures traits associated with one's level of extraversion or introversion. The primary factors which load high on the Extraversion factor include Warmth (A+), Liveliness (F+), Social Boldness (H+), Privateness (N-), and Self-Reliance (Q2-). These primary factors tend to share the element of extroversion, and as such, reflect individuals who tend to be less shy, more enthusiastic, and require less need for privacy, whereas the those scoring in the opposite directions are seen as much more introverted in nature. (Karson, Karson & O'Dell, 1997; Russell & Karol, 1994).

Anxiety (AX), the second global factor, generally measures an individual's degree of generalized anxiety. For example, low scorers on this global factor generally are more calm, unperturbed and placid, while high scorers are usually seen as more anxious, overwhelmed, and easily upset. Emotional Stability (Factor C-), Vigilance (Factor L+), Apprehension (Factor O+), and Tension (Factor Q4+), yield higher scores on this global factor. A higher score on AX typically depicts an individual who is overwhelmed, or acknowledging considerable problems (Karson, Karson & O'Dell, 1997; Russell & Karol, 1994).

The third global factor, Tough Mindedness (TM), measures on the lower end receptiveness, open-mindedness, and intuitiveness, and on the higher end tough mindedness, and resoluteness. Individuals on the higher end of this scale may be less empathetic compared to their lower scoring counterparts. The primary factors that reflect higher Tough Mindedness scores include Warmth (Factor A-),

Sensitivity (Factor I-), Abstractedness (Factor M-), and Openness to Change (Factor Q1-) (Karson, Karson & O'Dell, 1997; Russell & Karol, 1994).

For the fourth global factor, Independence (IN), low scoring individuals tend to be accommodating, agreeable, and selfless, when compared to the higher scoring individuals who appear independent, persuasive, and willful. The primary factors which load onto the Independence global factor include, Assertiveness (E+), Social Boldness (H+), Vigilance (L+), and Openness to change (Q1+). The independence global factor can suggest anger and/or aggression, however aggression levels should not be solely concluded from scores on this factor (Karson, Karson & O'Dell, 1997; Russell & Karol, 1994).

The fifth global factor, Self-Control (SC), describes individuals on the lower scoring end as unrestrained and who follow their urges, compared to higher scorers, who appear self-controlled and inhibited in relation to their urges. The factors loading on the higher end of this scale include Liveliness (Factor F-), Rule-Consciousness (Factor G+), Abstractedness (Factor M-), and Perfectionism (Factor Q3+) (Karson, Karson & O'Dell, 1997; Russell & Karol, 1994).

THE DEVELOPMENT OF THE SCALE OF ACCURATE PERSONALITY PREDICTION (SAPP)

Across the span of self-knowledge measurements, there appears to be a dearth of direct easily obtainable measure of self-knowledge. One that may

hopefully serve this function is the aforementioned SAPP, which combines one's obtained scores on the 16PF with self-predictions on the 21 primary and global scales (Miller, 2000). This measure thus utilizes two types of personality assessment; a self-report measure and a standardized objective test.

Miller (2000) addressed three main purposes for the SAPP, including 1) the development of a measure the accuracy of an individual's self-predictions of his/her personality traits, 2) the examination of which personality traits best contribute to accurate self-prediction, and 3) the identification of which personality traits are most common in those individuals who had good predictive ability, and those with poorer predictive ability.

Miller (2000)'s study involved a sample of 196 subjects. The subjects completed the 16PF, and following the administration were provided a scoring form of the 16PF (the 16PF Fifth Edition Record Form, see Appendix A), and instructed to rate themselves on each of the sixteen personality factors and five global factors. The participants' individual predicted self-report scores were then compared to their objectively derived scores on the 16PF. Miller (2000) was then able to obtain a total score of the degree of accuracy of self-prediction for each participant, by utilizing the following formula:

$$\begin{aligned} \text{SAPP} = & [\text{OSA} - \text{PSA}] + [\text{OSB} - \text{PSB}] + [\text{OSC} - \text{PSC}] + [\text{OSE} - \text{PSE}] + \\ & [\text{OSF} - \text{PSF}] + [\text{OSG} - \text{PSG}] + [\text{OSH} - \text{PSH}] + [\text{OSI} - \text{PSI}] + \\ & [\text{OSL} - \text{PSL}] + [\text{OSM} - \text{PSM}] + [\text{OSN} - \text{PSN}] + [\text{OSO} - \text{PSO}] + \\ & [\text{OSQ1} - \text{PSQ1}] + [\text{OSQ2} - \text{PSQ2}] + [\text{OSQQ3} - \text{PSQ3}] + \end{aligned}$$

$$[OSQ4 - PSQ4] + OSEX - PSEX + [OSAX - PSAX] + \\ [OSTM - PSTM] + [OSIN - PSIN] + [OSSC - PSSC].$$

In the above formula, the OS components stands for obtained scores across the respective 21 factors, and the PS scores refer to the predicted score for each factor. The formula continues by summing the absolute values of the differences for all 16 primary and five global factors. Absolute differences between the obtained and predicted sten scores are calculated, because the directionality of the differences is not relevant. The SAPP score then becomes the sum of the absolute differences of all the obtained and predicted sten scores across all 21 factors.

The above formula reveals that low scores are reflective of good self predictors, whereas higher scores reflect a decreased ability to self predict. The lowest obtainable score on the SAPP is 0, reflecting 100% accuracy in self-prediction, and the highest obtainable score on the SAPP is 189, indicating the weakest self-prediction.

Miller (2000) compared the obtained factor scores across two groups; Group 1 - those individuals whose SAPP scores were one standard deviation below the SAPP mean, and Group 2- those individuals whose scores were one standard deviation above the mean. Group 1 subjects were then those who did relatively well in predicting their scores, while those in Group 2 did relatively poorer in predicting their scores. Results indicated the Group 1 subjects showed the higher (+) and lower (-) scores across the following factors: Tough Mindedness (-), Openness to

Change (+), Sensitivity (+), Reasoning (+), Extraversion (+), Privateness (-), Vigilance (-), Warmth (+), and Liveliness (+).

To determine which 16PF personality traits were the best predictor of one's SAPP score, Miller (2000) performed a regression analysis, where the SAPP score became the criterion variable and the obtained scores for twenty-one scales of the 16PF the predictor variables. Results indicated Tough Mindedness (-) was the best predictor of individuals SAPP scores in this study. Reasoning (+) was the second best predictor, Independence (-) as the third best predictor, Tension (+) as the fourth best predictor, and Anxiety (-) as the fifth and final significant predictor. Individuals who scored higher on the SAPP (those with less predictive ability) were found to be introverted, reserved, restrained, concrete, unsentimental, private, tradition, wary of others, and unempathic. Those scoring lower on the SAPP, with higher predictive accuracy tended to be sensitive in nature, abstract, warm, lively, trusting of others, open to change, outgoing, and intuitive. Following Miller's (2000) study, research investigating the reliability and validity of the SAPP have been conducted.

Validation Efforts of the SAPP

Following Miller's (2000) research, Hood (2001) conducted a study to validate the SAPP. Hood (2001)'s study replicated Miller's findings and set out to validate the SAPP score by utilizing the Private Self-Consciousness score of the Self-Conscious Scale (SCS) (Osberg, 1975) to establish convergent validity,

believing there may well be some overlap between the two measures. Hood also tested the discriminant validity of the SAPP score by comparing it with a trait not expected to be related to the individuals' SAPP scores. To do so, Hood had her subjects also complete the Tennessee Self-Concept Scale (Brown, 1998). Hood hypothesized the SAPP would produce low, or insignificant correlation with the self-esteem measure. Sixty-two students participated in the study. Hood's results indicated the SAPP score did not correlate significantly with either of the measures. Hood concluded that the SAPP score does not measure the construct of self-reflection, stating "while an individual may participate in self-reflection, this study demonstrate that this does not necessarily relate to accurate prediction of their own personality traits."

Another study attempting to validate the SAPP was conducted by Anderson (2002). The study also aimed to measure the SAPP's convergent validity by correlating it with results from the Self-Monitoring Scale (Snyder, 1974). Anderson hypothesized low scorers on the SAPP, those with high degrees of accurate self-prediction would correlate highly with high self-monitors. She further elaborated high self monitors tend to be more self-aware, manipulate their behaviors to adapt to their environment, interpret other people's reactions, and compare themselves to those around them. The aforementioned characteristics appear to be linked to building a greater self-knowledge. Seventy-seven participants completed the study, which involved completing a packet including the SAPP and Self-Monitoring Scale. Results found no significant correlations between

the SAPP and the Self-Monitoring Scale, and thus concluded the SAPP is not considered likely a measure of an individual's level of self-monitoring.

Winter (2003), conducted a different study to provide construct validation for the SAPP. Winter's study had two groups of participants identified as being different in their ability to self-predict their scores. The two groups Winter used in her study were twenty-two psychology students, and ten engineering students. She hypothesized the psychology students would display better prediction rates of their personality compared to engineering students. Participants were provided a packet including the 16PF, and a blank 16PF scoring form to fill out where they predict their scores to fall. Winter utilized independent t-tests to determine significant group differences. However, results indicated there were no significant group differences, thus failing to offer construct validation for the SAPP measure.

Since Winter's study appeared to be limited by its small sample size, Grossenbacher (2006)'s study set out to reexamine the validation study proposed by Winter (2000). Grossenbacher (2006), replicated Winter's study, and expanded on the inclusion criteria to include individual's who obtained degrees, and were practicing in their field of study (psychology or engineering). Grossenbacher (2006) found significant differences between the mean SAPP scores for the two groups ($t = -4.247, p \leq .01$), with the psychology participants scoring lower on the SAPP, indicating better prediction accuracy when compared to the engineering participants, regarding their levels of self-knowledge.

Additional validation efforts have been also attempted. Glywasky (2003) was unable to find a correlation between the SAPP and Private self-consciousness score of the Private Self-Consciousness Scale, indicating the SAPP is not measuring the same construct as the Self-Consciousness Scale. Hickey (2005), stated there is an assumption that the more a person is in agreement with the way others seem them, the more accurate they are in predicting their own self-awareness. Therefore, Hickey (2005), had participant and their families predict the personality traits of the participant. A measure of concordance was developed to determine the amount of agreement between the raters. Hickey (2005) found no significant correlation between SAPP scores and a measure of concordance, however correlation between SAPP scores and the concordance measure approached significance in the predicted directed. Furthermore, no significant findings were established when comparisons were made between those who had high and low SAPP scores, however trends followed the predicted direction. Similar to previous validation efforts, Layton's (2005) study did not find significant results to support the validity of the SAPP. However, Wolf (2006), replicated Layton's study, in an attempt to increase sample size, and found strong and significant correlation between the SAPP and the derived concordance measure.

Reliability of the SAPP

Silva (2011), developed the first study to test reliability of the SAPP. She did this by having participants complete the 16PF and then predict their scores

during their initial testing, and repeated it again two weeks later. Sixty-two participants completed the testing. SAPP scores were calculated for all subjects across the two testing periods, and results indicated a significant correlation was found between two SAPP scores, although below what is typically considered acceptable test-retest correlation ($r^2=.397$, $p<.05$).

Hirsch (2012) replicated the Silva study, using 58 individuals who initially completed the 16PF, and two weeks later completed it again. A SAPP score Person correlation indicated a significant moderate correlation between the two SAPP scores ($r^2=.566$, $p<.05$).

Sverdlova (2012) studied the reliability of the SAPP across a four-week interval between testing sessions. Fifty-eight participants participated in the study. The results indicated a significant correlation ($r^2=.466$, $p<.05$), which was higher than the obtained reliability in Silva's study. Lastly, Elghossain (2012) studied the reliability of the SAPP across a six-week interval between testing sessions. Forty-seven participants participated in the study. The results indicated a significant correlation ($r^2=.722$, $p<.01$). In sum, the data to date collected suggest that the SAPP is a generally reliable measure of self-knowledge.

Adjustment of the SAPP Score

McElligot (2015) adjusted the overall SAPP scoring procedure from Miller's (2000) original scoring system. Following Miller's scoring system, the lower the SAPP score meant the higher the individuals' level of self-knowledge.

McElligot modified the scoring system to reflect a linear relationship between higher scores on the SAPP meaning stronger levels of self-knowledge. McElligot did this by subtracting the obtained SAPP score from 189 (the highest achievable SAPP score). This is the scoring system which was used in the present study.

STATEMENT OF PURPOSE & HYPOTHESIS

The purpose of this current study was to further explore Miller's (2000) measure of accurate self-prediction regarding self-knowledge. Specifically, the aim of this study was to re-evaluate the test-retest reliability of the SAPP across a four-week interval of test completion. Moreover, this study attempted to increase the overall sample size as a means to increase its correlational strength.

It was hypothesized that a subject's SAPP score will remain stable over time, and will yield a similar score between both testing trials. Additionally, it was hypothesized that the test-retest reliability of the SAPP in the present study will be equivalent or stronger than which Sverldlova (2012) obtained over a four-week period. It was hypothesized that the present study would yield equally, or slightly lower test-rest reliability scores compared to those which Silva (2011) and Hirsch (2012) obtained over a two-week retest period. Lastly, it was hypothesized that the present study would yield equivalent or stronger test-retest reliability than which Elghossain (2012) obtained over a six-week retest period. If the SAPP is proven to be a reliable measure of self-knowledge, it will help improve the ability of professionals in the field of mental health to tailor treatment goals and interventions to the specific individual based on their level of self-knowledge. Additionally, it will serve its purpose as being one of the only psychological tools currently developed to measure self-knowledge as a trait.

METHOD

SUBJECTS

The instructions for this study were disseminated via electronic mail, using a convenience sample. Subjects of this study were asked to participate on a voluntary basis, and had the option of discontinuing their participation at any given time. Participants were obtained through social media outlets, undergraduate classes for research credit at Florida Institute of Technology, and word of mouth. Testing materials were disseminated online through electronic email. A total of 72 letters of instruction were emailed for the first Trial. However, only 51 people completed the first trial. Those who completed the first trial were retained in the database and received the instruction email again four weeks later. Participants who completed both trials in their entirety, without any missing data, and met inclusion criteria were included in the findings of this research project, resulting in 29 number of participants.

INSTRUMENTS

The instruments utilized in this study included the 16PF Fifth edition (online administration), an equivalent form to the 16PF Individual Record Form (see Appendix D), demographic table (Table 1), a letter of instruction (see

Appendix A), a first trial email (see Appendix B), a second trial email (see Appendix C).

PROCEDURE

To establish test-retest reliability of the SAPP, participants were asked to participate in two separate testing sessions, an initial session, and a second one four-weeks later. Participants were provided a word document attachment with written instructions, and informed to return all completed test materials provided to them. To see an example of the instructions, trial 1 email, and trial 2 email that was sent to the participants, please refer to Appendices A, B, and C (Adapted from Silva 2011, and Hirsch 2012).

Each participant was emailed with instructions on how to complete the testing materials for the first trial of the online study. The email included instructions for participation, a unique log in and password for the IPAT website, and unique ID code for Trial 1 and Trial 2. They were provided a link to the IPAT website first to fill out a blank 16PF profile form (Appendix D). At the completion of the 16PF they were provided a link to a Qualtrics questionnaire to evaluate themselves on bipolar continuums for the five global and sixteen personality factors scales. Identical to Miller's (200) study, the SAPP score was obtained for each participant by adding together the total amount of absolute difference from all

twenty-one total scales of the 16PF. In order to obtain this score as mentioned earlier, the predicted score was subtracted from the obtained score, and then the absolute values of these sums were added together. Lastly the final score was subtracted from 189. A score of 189 is indicative of high levels of predictive ability, whereas a SAPP score of 0 indicates low self-predictive ability.

The participants who completed both parts of the first trial were sent a second email 28 days later to remind them to complete the second trial of the study within two days, which was four weeks from their first trial. The participants completed the blank 16PF questionnaire on IPAT's website first and upon completion were prompted to follow the link to complete the second trial of the SAPP, where they predicted their scores on the 21 factor scales, and then took the 16PF through the IPAT website. The second trial SAPP scores were then compared to the first trial SAPP through a correlational analysis.

RESULTS

A subject pool of 29 participants completed both testing trails with no missing data. A total of 72 letters with instructions were sent out to participants who volunteered to participate in the first trial. 52 participants completed the first trial in its entirety, and participants who did not complete the first trial in its entirety were removed from the database. The demographic variables that were recorded for each individual participant included age, highest level of education obtained, race, ethnicity, gender, geographic location, and marital status. Of the 29 participants who completed both trials, the mean age was 32.9, with ages ranging from 18 to 58. Three participants chose not to provide their age. Most of the participants were females (79.3%), with males representing 20.7% of the participants. The demographic analysis indicated that 96.6% of the participants identified as Caucasian (N=28), and 3.4% of the participants identified as other, and one wrote in Middle Eastern (N=1). Of the total participants, 3.4% (N=1) identified as Spanish, Hispanic, or Latino. In regard to level of education, 3.4% (N=1) of the subject pool reported being a high school graduate, 13.8% (N=4) reported attending college but not completing it, 6.9% (N=2) reported having an associate's degree, 17.2% (N=5) reported possessing a bachelor's degree, 37.9% (N=11) stated they hold a master's degree, and 20.7% (N=6) reported they have a doctoral degree.

In regard to geographic location, 11 participants reported they were from Florida, nine endorsed being a resident of New York, two participants identified as

being a resident of Mississippi, and one participant reported being a resident of California. Additionally, two participants identified being a resident of North Carolina, one participant indicated they reside in California, and one participant endorsed being a Louisiana resident. Lastly, one participant identified being a resident of Texas, one participant identified being a resident of Ontario, and one participant identified being a resident of Turkey. The majority of participants (55.2%, N=16) indicated their marital status as single, while 34.5% (N=10) reported their marital status as married, and 10.3% (N=3) identified their marital status as divorced. A comparison of the demographic data of the present study, with the normative sample of the 16PF and full database of research completed on the Scale of Accurate Personality Prediction can be found in Table 1. The sample of the current study is predominately comprised of individuals who identified as female, whereas both normative samples have more of an equal gender distribution. In regards to racial and ethnic diversity, the current sample has a greater percentage of Caucasian participants. A greater percentage of participants in the present study have master's or doctorate degree compared to both the normative sample and the SAPP database sample. Similar to the SAPP database sample, the majority of participants in the present study live in the southeast.

As a means to examine test-retest reliability, participant's SAPP scores were gathered from the initial trial, and from the second trial, which occurred four weeks later for each subject. SAPP scores of the sample during the initial testing trial (trial one) revealed a mean of 151.83 and a standard deviation of 11.28, with

scores ranging from 130 to 170. For the SAPP scores derived during the second testing session four weeks later (trial two), a mean of 152.10 was found, with a standard deviation of 10.84, with scores ranging from 131 to 169. A Pearson correlation revealed a significant strong correlation between the SAPP scores obtained during both trials ($r^2=.727$, $p<.01$). The four-week test-retest reliability values obtained for each of the 21 factors are comparable to the 16PF's normative sample (Table 2). Additionally, the means from the present study are comparable to the mean of SAPP scores from the full database of research on the measure (Table 3). The present study yielded standard deviations that are somewhat lower than the standard deviation of the entire database, which can be attributed to the smaller range of scores, and sample a smaller size.

DISCUSSION

The present study was conducted as a replication of Sverdlova's (2012) study to further examine the test-retest reliability of the SAPP over a four-week period. A Pearson correlation was conducted with the present study and revealed a significant strong correlation ($r^2=.727$, $p<.01$), thus suggesting the participant's SAPP scores from trial one and trial two remained stable across a four-week period, as hypothesized. Additionally, the present study ($r^2=.727$, $p<.01$) displayed stronger test-retest reliability compared to Sverdlova in 2012 ($r^2=.466$, $p<.05$), as hypothesized. It was further hypothesized that the present study would have lower test-retest reliability than Silva in 2011 ($r^2=.397$, $p<.05$), and Hirsch in 2012 ($r^2=.566$, $p<.05$), due to the lengthier time between trials of the present study, which typically leads to a decrease in reliability coefficients. However, results indicated the present study had a stronger correlation than both Silva (2011) and Hirsch (2012), which may be related to the present study's smaller sample size. Lastly, it was hypothesized that the present study's correlational strength ($r^2=.727$, $p<.01$), would yield an equal or stronger correlation than Elghossain's (2012) study ($r^2=.722$, $p<.01$), which was confirmed. As a result of the present study's strong correlation, it is reasonable to consider the SAPP a reliable measure of self-knowledge. The high reliability value can be explained by the correlations across the sixteen factors, which were all found to be in an acceptable range and are largely similar to those of the original 16PF standardization sample studies.

Nevertheless, it is important to caution the possibility of large effect size in the present study which may have increased the probability of a false-positive result.

LIMITATIONS

The present study's sample size ($N=29$) is a primary limitation to the current study. The researcher utilized social media to gather participation for the sample, and received 72 individual emails from individuals interested in participating in the study. Of the 72 individuals, 51 participants completed the first trial. Individuals received reminder emails to complete both the first and second trials. There was a large attrition rate ($N=22$) from trial one to trial two, resulting in 29 individuals completing both trials. Additionally, some participants ($N=3$) did not initiate responses to the SAPP questionnaire and could not be included in the final results. This is likely due to an oversight of the requirement that participants follow a hyperlink after completion of the 16PF to open the webpage for the SAPP. While results indicated a significant strong correlation, the small sample size may have enhanced the effect size, which could have increased the probability of a false-positive result.

The method of which the research was conducted proved to be another limitation of the study. One participant emailed the researcher and stated they were unable to open the instructions on their desktop, and requested a different version.

The participant was provided the instructions in the body of the email to remedy the situation. It is possible similar technical difficulties occurred which were not reported to the researcher. Additionally, the participants were provided two separate ID codes, and two sets of logins and passwords, which potentially added to the complexity of the instructions. The technical and instructional difficulties were avoided by pen and paper format conducted by Silva (2011).

The homogeneity of the current study, specifically regarding education and diversity is another limitation to the study. Majority of participants (58.6%) reported possessing a master's degree or doctorate degree. Additionally, 96.6% of the participants identified as Caucasian. The results indicated the participant pool is largely unrepresentative of the general population, and the results of this study cannot be applied or generalized to a broader population.

The final limitation that should be taken into consideration is the researcher's lack of control over testing settings. It was difficult to guarantee that the participants completed both trails in similar settings, without interruptions, despite instructions indicating to do so. However, the online administration provided date and time stamps which aided in the researcher's ability to confirm testing was completed in a timely manner, and exactly four-weeks later. Even considering the limitations of the current study, the correlations between obtained and predicted sten scores in this study were quite high, and comparable to the two-week and two-month test-retest reliability coefficients that have been obtained for the 16PF in the past. In sum, although the current study is not without limitations,

the present results are valuable. The current study, in conjunction with previous psychometric studies on the SAPP, suggests the SAPP is at least a reliable measure. Possessing a reliable and valid self-knowledge measure could aid in mental health professional's ability to formulate more effective treatment plans and intervention utilization based on an individual's measure of self-knowledge. Furthermore, an accurate measure of self-knowledge has the potential to shed light on an individual's level of insight, and further aid in measuring certain treatments' effectiveness.

Future studies should focus on implementation of strategies to decrease attrition rate, increase sample size, and improve the sample's representation of the population by utilizing a more diverse sample. It can be assumed that test length, instruction length, and the fact that more than one testing was required no doubt contributed to drop out rates and ultimately a smaller than expected sample size. It is recommended that future studies adapt a more concise version of instruction with less steps for participants to follow. Additionally, future studies should consider ways to present the 16PF and the SAPP in one web format, rather than providing a hyperlink to follow to complete the SAPP to increase completion of testing materials. In conclusion, the current study provided some supportive evidence that the SAPP is consistent across a four-week interval of time, and as such supports its overall level of reliability and psychometric strength. Thus, further examination of the psychometric properties of the SAPP is worth pursuing and examining.

Reference

- Abrams, L.D., Ellis, A., & Abrams, M. (2009). *Personality theories: Critical perspectives*. Sage Publications.
- Afandor, J. H. (2006). Expert validation of the scale of accurate personality prediction (SAPP). Unpublished master's thesis, Florida Institute of Technology.
- Allen, B. P. (2003). *Personality theories fourth edition*. Boston, MA: Pearson Education, Inc.
- Anderson, A. (2002). A validation of the Scale of Accurate Personality Prediction using the Self-Monitoring Scale. Unpublished master's thesis, Florida Institute of Technology.
- Bishop, P. (2009). *Analytical psychology and German classical aesthetics: Goethe, Schiller & Jung*. London: Routledge.
- Brown, R. (1998). Review of the Tennessee Self-Concept Scale. In Impara, J.C. & Plake, B.S. (eds). *The thirteenth mental measurements yearbook*.
- Carver, C. S., & Glass, D. C. (1975). The self-consciousness scale: a discriminant validity study. *Journal of Personality Assessment*, 40 (2).
- Elghossain, R. M. (2012). Examining test retest reliability of the scale of accurate personality prediction. Unpublished master's thesis, Florida Institute of Technology.
- Engler, B. (2005). *Personality Theories*. Belmont, CA: Cengage Learning, Inc.

- Glywasky, K. E. (2003). Re-evaluation of the scale of accurate personality prediction as a valid measure of self knowledge. Unpublished master's thesis, Florida Institute of Technology.
- Grossenbacher, A. S. (2006). A validation of the Scale of Accurate Personality Prediction using a priori comparison. Unpublished master's thesis, Florida Institute of Technology.
- Hickey, J. L. (2005). A validation of the scale of accurate personality prediction. Unpublished master's thesis, Florida Institute of Technology.
- Hirsch S. J. (2012). Test-retest reliability of the scall of accurate personality prediction. Unpublished master's thesis, Florida Institute of Technology.
- Hjelle, L. A., Ziegler, D. J., (1992). *Personality theories: Basic assumptions, research, and applications*. New York, NY: McGraww-Hill.
- Hood, C. A. (2001). A measure of the accuracy of self-predicted personality traits. Unpublished master's thesis, Florida Institute of Technology.
- Karson, M., Karson, S., & O'Dell, J. (1997). *16PF Interpretation in clinical practice: A guide to the fifth edition*. Champaign, IL: Institute for Personality and Ability Testing, Inc.
- Layton, E. E. (2005). A validation of the scale of accurate personality prediction using self versus other comparisons. Unpublished master's thesis, Florida Institute of Technology.
- Leary, M. R., Tangney, J.P. (2005). *Handbook of self and identity*. New York, NY: The Guilford Press.

- Miller, C.S. (2000). A measure of accuracy of self-predicted personality traits. Unpublished master's thesis, Florida Institute of Technology.
- Olson, E. T. (1998). There is no problem of the self. *Journal of Consciousness Studies*, 5(5-6), 645-657. Retrieved from <http://www.ingentaconnect.com.proxy.lib.umich.edu/content/imp/jcs/2000/00000007/F0020001/1142>.
- Osberg, T.M. (1985) The self consciousness scale. In Keyser, D.J. & Sweetland, R.C. (eds), test critiques, Volume V, Kansas City, MO: Test Corporation of America.
- Oyserman, D., Elmore, K. C., & Smith, G. C. (2012). *Self, self-concept, and identity. Handbook of Self and Identity*. https://doi.org/10.1300/J082v10n03_13.
- Robins, R. W., Norem, J. K., & Cheek, J. . (1999). Naturalizing the Self. In *Handbook of Personality: Theory and Research* (pp. 443–477). <https://doi.org/10.1016/B978-012134645-4/50022-6>.
- Russell, M. & Karol, D. (1994). 16PF Fifth edition, administrator's manual. Institute for personality and ability testing, Inc. Champaign, Illinois.
- Self Knowledge [Def. 1]. (n.d.). Merriam-Webster Online. In Merriam-Webster. Retrieved from <http://www.merriam-webster.com/dictionary/self-knowledge>.

- Silva, M.C. (2011). Test re-test reliability of the Scale of Accurate Personality Prediction. Unpublished master's thesis, Florida Institute of Technology.
- Snyder, M. (1974). The self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, 30, 526-537.
- Sverdlova, M. (2012). Test-Retest reliability of the scale of accurate personality prediction. Unpublished master's thesis, Florida Institute of Technology.
- Vazire, S. & Wilson, T. D., (2012). *Handbook of self-knowledge*. New York, NY: The Guilford Press.
- Wolf, A. T. (2006). A re-validation of the scale of accurate personality prediction using self versus peer comparisons. Unpublished master's thesis, Florida Institute of Technology.
- Vogt, D. S., & Colvin, C.R. (2005). Assessment of accurate self-knowledge. *Journal of Personality Assessment*.
- Winter, S. N. (2002). A measure of construct validity of the Scale of Accurate Personality Prediction. Unpunished master's thesis, Florida Institute of Technology.

Table 1. Demographic Percentages

Demographic Category		Current Sample Percent (N=29)	SAPP Database Percent ^a (N=645)	Normative Sample Percent ^b (N=2500)
Gender	Male	20.7%	42.0%	48.7%
	Female	79.3%	58.0%	51.3%
Race^c	African	0%	2.3%	12.1%
	American/Black	0%	9.3%	2.9%
	Asian	0%	71%	80.2%
	Caucasian	96.6%	.2%	1.0%
	Native American	0%	5.3%	3.8%
	Other	3.4%		
	Hispanic Origin	3.4%	11.9%	9.0%
Age Group	15 to 17	0%	1.2%	4.6%
	18 to 24	6.8%	51.3%	13.8%
	25 to 44	72.1%	34.4%	41.7%
	45 to 54	3.4%	5.3%	12.9%
	55 to 64	6.8%	5.9%	10.8%
	65 and older	0%	1.9%	16.2%
Education Level	≤12 years	3.4%	5.0%	61.5%
	12.5-16 years	37.9%	55.8%	22.7%
	16+ years	58.6%	39.2%	15.8%
Marital Status	Single	55.2%	72.9%	
	Married	34.5%	20.6%	
	Divorced	10.3%	4.8%	
	Separated	0%	1.1%	
	Widowed	0%	.6%	
Geographic Location	Southeast	54.9%	78.9%	
	Southwest	3.4%	3.8%	
	Northeast	27.5%	13.1%	
	Midwest	0%	4.0%	
	Canada	0%	.2%	

^a From the SAPP Database, which is an accumulation of data from multiple studies on the SAPP.

^b From "Characteristics of the Norm Sample" by S.R. Conn & M.L. Rieke, in press. In S.R. Conn & M. L. Rieke (Eds.). *The 16PF Fifth Edition Technical Manual*, Champaign, IL: Institute for Personality and Ability Testing, Inc. Please note that information regarding marital status and geographic location is not available.

^c Totals may exceed 100% since participants of the present study were allowed to choose more than one race. Additionally, in the present study and in the normative sample those who identified as Hispanic also endorsed at least one race category.

^d Totals may be less than 100% since participants of the present study were allowed to choose not to answer demographic information.

Table 2. Test-Retest Reliability Data of the 16PF, Fifth Edition

<i>Test-Retest Interval</i>				
	Primary Factor	Two-Week (N=29)	Two-Week^a (N=204)	Two Month^a (N=159)
A	Warmth	.84	.83	.77
B	Reasoning	.57	.69	.65
C	Emotional Stability	.80	.75	.67
E	Dominance	.80	.77	.69
F	Liveliness	.91	.82	.69
G	Rule-Consciousness	.77	.80	.76
H	Social Boldness	.83	.87	.79
I	Sensitivity	.92	.82	.76
L	Vigilance	.90	.76	.56
M	Abstractedness	.81	.84	.67
N	Privateness	.92	.77	.70
O	Apprehension	.89	.79	.64
Q1	Openness to Change	.83	.83	.70
Q2	Self-Reliance	.84	.86	.69
Q3	Perfectionism	.81	.80	.77
Q4	Tension	.83	.78	.68
	SAPP	.73		
Global Factor				
	Extraversion	.93	.91	.80
	Anxiety	.90	.84	.70
	Tough- Mindedness	.92	.87	.82
	Independence	.93	.84	.81
	Self-Control	.90	.87	.79

^aFrom "Comparison of the 16PF Fifth Edition and Form A (Fourth Edition)" by S.R. Conn, in press. In S.R. Conn & M.L. Rieke (Eds.), *The 16PF Fifth Edition Technical Manual*. Champaign, IL: Institute for Personality and Ability Test, Inc.

Table 3. SAPP Ranges, Means, and Standard Deviations

Sample	Range	Minimum	Maximum	Mean	Standard Deviation
Current Sample Trial One (N=29)	40.00	130.00	170.00	151.83	11.28
Current Sample Trial Two (N=29)	38.00	131.00	169.00	152.10	10.84
Current Sample Both Trials (N=29)	40.00	130.00	170.00	151.97	10.97
SAPP Database ^a (N=643)	69.30	101.30	170.60	147.04	13.05

^aFrom the SAPP Database, which is an accumulation of data from multiple studies on the SAPP.

APPENDIX A

INSTRUCTIONS FOR PARTICIPANTS

Dear Participant,

Thank you for agreeing to participate in this research study. The purpose of this study is to further explore the reliability of a new scale of self-knowledge for the Sixteen Personality Factor Questionnaire (16PF), which is the Scale of Accurate Personality Prediction (SAPP). More specifically, this study aims to establish test-retest reliability. This will require your participation on two distinct testing occasions. The second testing session will take place four weeks following the first testing session.

Please read all of the following steps before beginning the study. After reading them carefully, follow them in order:

1. To complete the 16pf assessment, go to: <https://www.netassess.ipat.com/> and enter the unique user name and password provided in the body of the email. Read the Terms of Service Provision and select Yes, I will. Select Continue. Please note: Exiting your web browser without agreeing to the Terms of Service, or responding No, will result in your passcode being locked and will require the code to be reset by the project team.
2. If this is your first trial, please enter the Trial One ID Code when prompted. You will use the Trial Two ID Code four weeks later. If this is your second trial, please enter the Trial Two ID Code.
3. You should use your ID Code for the answer blank regarding your name.
4. The 16PF should take approximately 30 minutes to complete. Once you have completed the 16PF, click on the link at the end. This link will redirect you to take to the second questionnaire, the SAPP.
5. Please reuse the same ID you used on the first questionnaire on the second questionnaire.
6. Answer each question on the questionnaire.
7. After a four-week delay you will complete the questionnaires again with a second username, password, and ID code. As a reminder, you will receive an email two days prior to when you are to complete the second trial. These instructions will be sent to you again. **Please remember to use the second username, password, and ID code that will be provided to you.** It is requested that you complete the second trial within a 24hr period of the date that is exactly four weeks from when you completed the first trial.
- 8.

Please be assured that the information you provide us is confidential. Your completion of the materials will serve as your consent to participate in this study. If you are interested in summary feedback concerning this study, please contact me via email, provided at the end of this page. Please note, to protect anonymity, individual feedback cannot be provided; only group summary results will be available. These results will be available upon completion of the research project.

Again, your assistance is appreciated. Please contact me if you have any further questions regarding the research.

Regards,
Taylor Anderson, M.S.
Doctoral Candidate
Florida Institute of Technology

APPENDIX B

FIRST TRIAL EMAIL TO PARTICIPANTS

Dear Participant,

Thank you for agreeing to participate in my doctoral research study. Attached is a Word Document with detailed instructions on how to participate. There are two questionnaires to complete for the study, and the instructions will explain how to do so.

This is your first trial of the study. Four weeks from the day you decide to complete the first trial, you will complete the second trial. I will send you a reminder email two days prior to when you are to complete the second trial. Please make sure you complete the first trial on a date you know you can complete the second trial exactly four weeks later. Reference the ID codes below when completing the questionnaires today and four weeks later. Additionally, you should use the ID code in the blank which asks for your name.

Trial One ID Code: XXXX (Use the first time you complete the questionnaires)

Trial Two ID Code: YYYY (Use the second time you complete the questionnaires)

You will be prompted to enter a username and password before completing the first part of the first trial. Please use the following username and password when completing the first trial:

Username: 22fc8faf0

Password: rimazuty

You are welcome to send me an email with any questions or concerns you may have regarding this project. I greatly appreciate your time!

Regards,

Taylor Anderson, M.S.
Doctoral Candidate
Florida Institute of Technology
andersont2014@my.fit.edu

APPENDIX C

SECOND TRIAL EMAIL TO PARTICIPANTS

Dear Participant,

Thank you for participating in the first trial of my study. Attached to this email is a Word Document with instructions on how to complete your participation in this study. Please read the instructions carefully. As with the first trial of the study, there are two questionnaires to complete. The instructions will explain how to do so.

Two days from now you will complete the **second** trial of my study. Reference the ID code below when completing the questionnaire two days from now.

Trial Two ID Code: YYYY

You will be prompted to enter a username and password before completing the first part of the second trial. Please use the following username and password when completing the second trial:

Username: 22fc8faf1

Password: zecimagi


You are welcome to send me an email with any questions or concerns you may have regarding this project. I greatly appreciate your time!

Regards,

Taylor Anderson, M.S.
Doctoral Candidate
Florida Institute of Technology
anderson2014@my.fit.edu

APPENDIX D

BLANK 16PF INDIVIDUAL RECORD FORM



16PF® Fifth Edition Individual Record Form
Profile Sheet

Name _____

Date _____

Instructions: Write the sten score for each factor in the second column. Starting with Factor A, place a mark over the spot representing the appropriate sten score. Repeat for each factor. Connect the marks with straight lines.

PRIMARY FACTORS

Factor	Sten	Left Meaning	Standard Ten Score (STEN)										Right Meaning
			1	2	3	4	5	6	7	8	9	10	
A: Warmth		Reserved, Impersonal, Distant	*	*	*	*	*	*	*	*	*	*	Warm, Outgoing, Attentive to Others
B: Reasoning		Concrete	*	*	*	*	*	*	*	*	*	*	Abstract
C: Emotional Stability		Reactive, Emotionally Changeable	*	*	*	*	*	*	*	*	*	*	Emotionally Stable, Adaptive, Mature
E: Dominance		Deferential, Cooperative, Avoids Conflict	*	*	*	*	*	*	*	*	*	*	Dominant, Forceful, Assertive
F: Liveliness		Serious, Restrained, Careful	*	*	*	*	*	*	*	*	*	*	Lively, Animated, Spontaneous
G: Rule-Consciousness		Expedient, Nonconforming	*	*	*	*	*	*	*	*	*	*	Rule-Conscious, Dutiful
H: Social Boldness		Shy, Threat-Sensitive, Timid	*	*	*	*	*	*	*	*	*	*	Socially Bold, Venturesome, Thick-Skinned
I: Sensitivity		Utilitarian, Objective, Unsentimental	*	*	*	*	*	*	*	*	*	*	Sensitive, Aesthetic, Sentimental
L: Vigilance		Trusting, Unsuspecting, Accepting	*	*	*	*	*	*	*	*	*	*	Vigilant, Suspicious, Skeptical, Wary
M: Abstractedness		Grounded, Practical, Solution-Oriented	*	*	*	*	*	*	*	*	*	*	Abstracted, Imaginative, Idea-Oriented
N: Privateness		Forthright, Genuine, Artless	*	*	*	*	*	*	*	*	*	*	Private, Discreet, Non-Disclosing
O: Apprehension		Self-Assured, Unworried, Complacent	*	*	*	*	*	*	*	*	*	*	Apprehensive, Self-Doubting, Worried
Q ₁ : Openness to Change		Traditional, Attached to Familiar	*	*	*	*	*	*	*	*	*	*	Open to Change, Experimenting
Q ₂ : Self-Reliance		Group-Oriented, Affiliative	*	*	*	*	*	*	*	*	*	*	Self-Reliant, Solitary, Individualistic
Q ₃ : Perfectionism		Tolerates Disorder, Unexact, Flexible	*	*	*	*	*	*	*	*	*	*	Perfectionistic, Organized, Self-Disciplined
Q ₄ : Tension		Relaxed, Placid, Patient	*	*	*	*	*	*	*	*	*	*	Tense, High Energy, Impatient, Driven

GLOBAL FACTORS

Factor	Sten	Left Meaning	Standard Ten Score (STEN)										Right Meaning
			1	2	3	4	5	6	7	8	9	10	
EX: Extraversion		Introverted, Socially Inhibited	*	*	*	*	*	*	*	*	*	*	Extraverted, Socially Participating
AX: Anxiety		Low Anxiety, Unperturbed	*	*	*	*	*	*	*	*	*	*	High Anxiety, Perturbable
TM: Tough-Mindedness		Receptive, Open-Minded, Intuitive	*	*	*	*	*	*	*	*	*	*	Tough-Minded, Resolute, Unempathic
IN: Independence		Accommodating, Agreeable, Selfless	*	*	*	*	*	*	*	*	*	*	Independent, Persuasive, Willful
SC: Self-Control		Unrestrained, Follows Urges	*	*	*	*	*	*	*	*	*	*	Self-Controlled, Inhibits Urges

Copyright ©1990 by the Institute for Personality and Ability Testing, P.O. Box 1168, Champaign, Illinois, U.S.A. 61824-1168. All rights reserved. May not be reproduced in whole or in part, stored in a retrieval system, or transmitted in any form or by any means, photocopying, mechanical, electronic, recording, or otherwise, without prior permission in writing from the publisher. Printed in the U.S.A.