CROSS-NATIONAL STUDY OF CHILDREN'S TEMPERAMENT: STRUCTURAL VALIDITY OF THE STUDENT STYLES QUESTIONNAIRE

By

CARMELO M. CALLUENG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE
UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2012

ACKNOWLEDGMENTS

I would like to express my sincerest appreciation to the significant persons who immensely contributed to the completion of my dissertation. I thank my advisor and committee chair, Dr. Thomas Oakland, for his competent advising and unwavering support for this project and throughout my graduate studies at the University of Florida. His expertise in temperament and international psychology has strongly motivated me in the pursuit of this project. His dedication to hard work, excellence, scholarly rigor, and research has inspired me to enhance my academic competence and professional commitment. On a personal note, I am blessed with Dr. Oakland's gift of person. I am insurmountably fortunate to experience a "family" with him during my stay in Gainesville.

I thank my co-chair, Dr. Diana Joyce, for her guidance and reliability. I am indebted to her for the countless assistance, positive regard, and encouragement in getting through this project and in my doctoral program. Her endearing qualities and values as a professor and as a person are worth emulating for. I also gratefully acknowledge my committee members, Drs. John Kranzler and Walter Leite, for their commitment and contributions to this project. Their feedback, guidance, and suggestions were valuable in the fruition of this project.

Lastly, I would like to express deepest gratitude to my family and acknowledge the value of their unconditional love and support. I am sincerely grateful to my mother, Francisca Callueng, my late father, Domingo Callueng, and my siblings (Teddy, Roderick, Joyce, Rudolph, and Clifford) and their families.

TABLE OF CONTENTS

LIST OF TABLES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

CROSS-NATIONAL STUDY OF CHILDREN'S TEMPERAMENT: STRUCTURAL VALIDITY OF THE STUDENT STYLES QUESTIONNAIRE

By

Carmelo M. Callueng

December 2012

Chair: Thomas Oakland
Cochair: Diana Joyce
Major: School Psychology

Temperament has a long history of scholarship dating back as early as 350 BC when Hippocrates (1984) associated body fluids or temperament with behavior. Temperament is broadly described as stylistic and relatively stable traits that subsume intrinsic tendencies to act and react in somewhat predictable ways to people, events, and other stimuli. Advances in temperament research include the development of scientific and practical measures that can be used to obtain information to better understand a person's social-emotional and behavioral functioning. The Student Styles Questionnaire (SSQ) is a measure of temperament based on the type theory of Jung, and Myers and Briggs. The SSQ was originally developed for use with children in the U.S. and is now adapted in more than 30 countries. Few studies have reported the psychometric characteristics of the SSQ adapted versions. To this end, this research examined the cross-national factor structure of temperament comprising of extroverted-introverted, practical-imaginative, thinking-feeling, and organized-flexible traits. Using the item level factor analytic framework with dichotomous items as indicators, results indicated that the four factor-model of temperament did not fit reasonably well to the data in each of the 21 countries. Most item indicators converged on the traits they intended to measure. However, factor loadings generally were low. In most countries, the bipolar traits are distinct on what they measure. However, multicollinearity was evident in data from China, Egypt, and Gaza. Findings are discussed in light of the shortcomings of factor analysis using dichotomous items when examining construct validity of the SSQ and sample size requirements of robust weighted least square estimator in CFA. Implications and limitations of the study also are discussed.

CHAPTER 1
REVIEW OF LITERATURE

Scholarship on temperament has had an important agenda that includes its conceptual meaning, measurement, and establishing a network of variables that are believed to be associated with temperament. Temperament currently is thought to characterize stylistic and relatively stable traits that subsume intrinsic tendencies to act and react in somewhat predictable ways to people, events, and other stimuli (Teglasi, 1998a; 1998b). A number of temperament theories present divergent views, resulting to a lack of consensus definition. Scholars differ on their views on what behaviors constitute temperament and the number of dimensions or traits that constitute temperament. Temperament can encompass behavioral styles, emotions, degrees of stability, and heritability (Goldsmith, et al., 1987).

Outweighing these areas of divergence are unifying views that serve as foundational ideas in understanding the concept of temperament (Goldsmith et. al., 1987). Scholarship generally supports the belief that temperament constitutes a group of related traits rather than a single trait, addresses the issue of individual differences rather than specifies-general characteristics, has biological origin, adheres to continuity in behavioral manifestations, is dynamic despite adherence to behavioral continuity, and refers to specific moods, attitudes, and dispositions that affect behaviors.

Temperament may help us understand children’s behavioral problems in light of goodness of fit model (Chess & Thomas, 1984) as well as the effects of children’s temperament in later life (Wach, 1994). Oakland, Glutting, and Horton (1996) expanded the temperament perspective in children by applying the Jungian typology and the Myers/Briggs type theory in the development of the Student Styles Questionnaire (SSQ) for children ages 8 through 17. The SSQ provides information on temperament preferences in four bipolar traits: extroverted-introverted, practical-imaginative, thinking-feeling, and organized-flexible.

An understanding of children's temperament preferences aids our understanding of the origins of behaviors and attitudes children display at home and school and its normal development. Temperament can assist in identifying talent, adjusting for possible weaknesses, enhancing personal and social development, promoting an understanding of others, assessing learning styles, promoting educational development, exploring career interest, and facilitating research and evaluation studies (Oakland, Glutting, & Horton, 1996). These broad uses of temperament are closely aligned to an ultimate goal of school psychology, namely to provide comprehensive and multilevel psychological services to children and youth that enable them to succeed academically, socially, behaviorally, and emotionally.

The quest to understand children's development, including their temperament, from cross-national perspectives has increased recently (Buss, 2011). This surge of interest in cross-national psychology over the past two decades with the goal to develop a broadly shared framework that recognizes the bidirectional roles of person and the environment in understanding behavior. Specifically, temperament research can be viewed from the lens of cultural and global diversity by engaging in studies that focus on theoretical and current practices in child and adolescent development.

The current study recognizes the value of valid measures to obtain information that can be used to make decisions relating to adjustment of children and youth. It primarily contributes in understanding the meaning of temperament cross-nationally and explores the emerging view that the development of temperament may be tempered by the environment and personal choice through the process of enculturation and accommodation, an idea closely related to psychological anthropology that views culture as constitutive of behavior (Miller, 1999). Few studies examined the cross-national structural validity of the Student Styles Questionnaire (SSQ) as a measure of temperament in children. This study addresses the psychological meaning of temperament by examining its construct equivalence using data from children in 21 countries.

Theoretical Perspective of Temperament

The concept of temperament has evolved from a simple idea related to body fluids to a more sophisticated and broad theoretical explanation that encompasses biological and environmental influences. The following discussion presents theoretical and historical perspective of temperament.

Classic Theory of Temperament

Hippocrates. Temperament was introduced in the Greek culture as early as the 4th century B.C. by Hippocrates. Temperament was described as behavior that clusters in four humors, namely, yellow and black bile, blood, and phlegm. They were considered to be neurochemical precursors of bipolar behaviors thought to oppose each other (i.e., warm and cool; dry and moist). Hippocrates' temperament theory was based on the premise that all mental, emotional, and behavioral disorders are caused by natural factors such as inherited susceptibility and an imbalance of the four humors (Hergenhahn, 2001).

Galen. Galen expanded on Hippocrates' four humors by translating them into physical and emotional characteristics of temperaments (Galen, trans, 1992;

Hergenhahn, 2001; Hippocrates, trans. 1939): choleric, phlegmatic, melancholic, and sanguine. Galen believed that the emotional and behavioral functioning was due to one pair of bodily states dominating the complementary pair (e.g., warm and moist dominating cool and dry). The choleric type has an over-supply of yellow bile that would incline a person to be both easily angered and easily calmed, to quickly change moods and likes, to be irascible, and, in its extreme, to be seen as a fool. The phlegmatic type has an over-supply of phlegm that would incline a person to be pale, slow, drowsy, apathetic, weak, to engage in fantasy, be mild mannered, and exhibit somatic complaints. The melancholic type has an over-supply of black bile that would incline a person to display extremes of happiness or malaise, sadness, and depression. The sanguine has an over-supply of blood that would incline a person to be pleasant, loving, affectionate, happy, optimistic, and hopeful (Galen, trans, 1992; Hergenhahn, 2001)

Kant. In 1798, Kant published *Anthropologie*, which included a chapter on temperament. He discussed how anthropology can become a viable alternative to psychology in studying a person's behavior (Eysenck & Eysenck, 1985). Specifically, Kant believed that anthropology can provide sufficient informational cues to predict and control a person's behavior. Similar to Galen, Kant believed that humors were the basic elements of the four distinct temperaments: choleric, phlegmatic, melancholic, and sanguine. Persons who display a choleric temperament type generally are easily irritated and become upset when others do not listen to what they say. Persons who display a phlegmatic temperament type generally are persistent and use reason rather than instinct to guide their actions. Persons who display a melancholic temperament type generally are fearful and sad. Persons who display a sanguine temperament type generally are carefree and sociable. In terms of behavioral characteristics, Kant further described a person's temperament as energetic (choleric vs. phlegmatic) and emotional (sanguine vs. melancholic). A person can have a distinct temperament based on his or her dominant behavioral characteristics. However, a combination of temperament types (choleric-phlegmatic) does not exist.

Modern Theory of Temperament

Wundt's Theory. Wundt, known as the father of experimental psychology, introduced a somewhat modern view of temperament in 1903 that closely resembled the four temperaments proposed by Galen and Kant (Eysenck & Eysenck, 1985). Contrary to Galen and Kant's unidimensional and categorical classification of temperament, Wundt believed that temperament can be measured and quantified and that a person can display one or a combination of positions on the energetic and emotional temperament dimensions. Wundt used the terms strong emotions in combination with weak emotions, and changeable in combination with unchangeable temperament to describe a person's behavioral characteristics.

Jung's Theory. Jung's theory of temperament was a product of his direct observation and interaction with clinic patients. He noted that patients with similar temperament qualities displayed similar adjustment problems. For example, he reported that patients with hysteria displayed extroverted attitudes regardless of their emotional instability. He described hysteric patients as being consistently aware of their surroundings and interacting actively with the therapist. On the other hand, he reported that patients with schizophrenia displayed introverted attitudes and preferred to be alone and withdrawn from others (Storr, 1991).

While acknowledging the ideas of Hippocrates and Galen about temperament, Jung proposed a differentiated temperament theory of psychological typology. Unlike Galen who believed that temperament is based on emotions or feelings, Jung explored the unconscious mind through the study of attitudes and values. Jung described attitude as “the psyche’s ability to act or react in a certain way” (Jung, 1921/1971, p. 414). The fundamental ideas in Jung’s psychological typology are the two attitudes referred to as introversion and extroversion. He believed that every person displays both introversion and extroversion tendencies and that a person is not entirely introverted or extroverted in attitude (Jung, 1921/1971). However, each person is thought to have a particular preference for introverted or extroverted attitudes and most often display behaviors consistent with their preferred attitude. Introversion and extroversion also can be understood as opposing attitudes along a continuum. Thus, persons may vary from strongly introverted to slightly introverted or strongly extroverted to slightly extroverted (Jung, 1921/1971). Jung also noted that extroversion and introversion attitudes bring both positive and negative qualities depending on a person’s general disposition (Wehr, 1971).

An introvert is drawn to their own thoughts and inner feelings. They prefer to be alone or in small groups, are introspective, tend to be cautious, and make decisions slowly. In contrast, extroverts are drawn to their outside environment. They are sociable, outgoing, and concerned with other’s expectations (Jung, 1921/1971).

In addition to the introversion and extroversion, Jung identified four basic psychological functions: thinking, feeling, sensation, and intuition (Jung, 1921/1971). Each of these functions can be displayed in tandem with introversion and introversion.

Jung (1920/1926) believed that persons tend to have well developed and dominant functions that are used more on a conscious level. There also are secondary functions that are not well developed and typically are used on a conscious level, yet can be potentially developed if desired by a person. The functions operate in pairs on a conscious level at any given time (Jung, 1921/1971).

Jung viewed thinking and feeling as rational functions (Jung, 1921/1971). Persons who display a thinking function make decisions through careful deliberation together with the use of logical and objective information. They highly value justice, truth, and facts. Persons who display a feeling function make decisions through a subjective process using one’s personal values (e.g., loyalty, sympathy, harmony with others). This value creates a sense of liking, disliking, or general mood that integrates past experiences and can lead to either accepting or rejecting a choice. “Feeling is a kind of judgment, differing from intellectual judgment in that its aim is not to establish conceptual relations but to set up a subjective criterion of acceptance or rejection” (Jung, 1921/1971, p. 434). According to Jung (1921/1971), feeling is a rational quality because the laws of reason are used in establishing value.

Jung believed intuition and sensation dimensions represent two different functions for acquiring and assessing information. Intuition and sensation operate in opposite directions and are considered irrational decision-making styles. Persons who display an intuition function assimilate quickly, thus integrating past experiences and unconscious perceptions. In contrast, persons who display a sensation function require direct physical experiences of the world, with facts or external stimuli acquired through the physical senses. Unlike intuition, sensation is a conscious perception and is governed more predominantly by concrete experience than by analysis.

The pairing of the combination of temperament functions resulted in eight temperament types. Jung believed that a person displays a dominant and frequently used type and an auxiliary or secondary type that is used less frequently (Jung, 1921/1971). He referred to four of these types as rational and the other four as irrational. The rational types are extroversion-thinking, introversion-thinking, extroversion-feeling, and introversion-feeling.

Jung identified famous persons who exemplify the qualities of each rational type. Charles Darwin, known for his penchant for scientific reasoning and facts, exemplifies an extroversion-thinking type. Immanuel Kant, known for the importance he placed on subjective reality and rational thinking, exemplifies an introversion-thinking type. Although both types are strongly influenced by ideas, the extroversion-thinking type is interested in objective data and will follow ideas externally. The introversion-thinking type is influenced by subjective ideas and will ponder those inwardly (Jung, 1921/1971). Jung believed that the extroversion-feeling and introversion-feeling types commonly are displayed in women. These types are guided by personal value systems comprised of subjective feelings and place strong value on harmony.

Jung’s four irrational types are extroversion-intuitive, introversion-intuitive, extroversion-sensing, and introversion-sensing. His example of an introversion-intuitive type is a person who is a dreamer or artist who enjoys contemplative moments. His example of an extroversion-intuitive type is a person who exhibits strong dependence on the external environment and is searching for new possibilities. Each of these types is strongly influenced by subjective factors and ideas. Compared to an extroversion-sensing type who seeks external facts, concrete objects, and ideas, an introversion-sensing type analyzes or thinks deeply on ideas and information (Jung, 1921/1971).

Jung discussed the phenomenon of falsification of type in relation to his theory of psychological type. Falsification of type occurs when a person uses his or her non-preferred functions more than his or her highly efficient and natural preferred functions. This phenomenon is caused when a person is placed in an environment that lessens or impedes opportunities to fully allow the expression of his or her natural preference. On the other hand, non-falsification of type occurs when a person is placed in an environment that allows the expression of his or her natural preference, resulting in a positive psychological well-being. A favorable adjustment is achieved when a person experiences a close match between his or her strengths and the demands of the environment (Jung, 1921/1971). This belief is consistent with Thomas and Chess' (1977) notion of goodness of fit.

Myers and Briggs Theory. Myers was fascinated with Jung's work and added a fourth dimension, judging and perceiving. Judging or perceiving are concepts that relate to how persons prefer to structure their lives in relation to the environment (Myers & Myers, 1980). A person with a judging orientation has a preference for planning, systems, order, routine, standards, self-regimentation, purposeful actions, decisiveness, and closure. A person with a perceiving orientation has a preference for spontaneity, understanding, tolerance, curiosity, zest for experience, and adaptability (Myers & Myers, 1980).

In 1942, Briggs and Myers started to develop test items that measure Jung's psychological types. The Myers-Briggs Type Indicator (MBTI) was published 20 years later (Myers & Myers, 1980). The MBTI combines Jung's three psychological dimensions [extroversion (E) - introversion (I), sensing (S) – intuition (N), and thinking (T) – feeling (F)] and Briggs' judging (J) – perceiving (P) dimension.

The extroversion-introversion dimension describes a person's energy orientation. A person who prefers an extroverted orientation draws energy from the outer world of people and events while a person who prefers an introverted orientation draws energy from the self, including thoughts and introspection (Joyce, 2010).

The sensing-intuition dimension describes a person's perception or learning processes. A person who prefers a sensing function acquire information from the physical senses, is realistic, and oriented to details. A person who prefers an intuitive function acquires information through logical and theoretical deductions, lesser attention to details, and more analytical in processing information (Joyce, 2010).

The thinking-feeling dimension describes a person's decision-making process. A person who prefers a thinking function make decisions based on facts, logic, and objective data. Fairness is achieved by applying the principles of justice and truth. A person who prefers a feeling function make decisions based on subjective values, such as empathy and well-being of others (Joyce, 2010).

The judging-perceiving dimension describes a person's environment or lifestyle orientation. A person who prefers a judging orientation values structure when interacting with the outside environment, is organized, and plans ahead. A person who prefers a perceiving orientation interacts with the outside environment is a flexible manner, is adaptive, and likes to keep options open (Joyce, 2010).

The four dimensions yield 16 type combinations: ISTJ, ISFJ, INFJ, INTJ, ISTP, ISFP, INFP, INTP, ESTP, ESFP, ENFP, ENTP, ESTJ, ESFJ, ENFJ, and ENTJ. Guidelines for administration, scoring, and interpretation of the MBTI are described in the *Essentials of Myers-Briggs Type Indicator Assessment, Second Edition* (Quenk, 2009).

As a measure of temperament or personality, the MBTI was conceptualized as a tool to obtain information that aids in understanding individual differences rather than a measure of pathology (Myers & Myers, 1980). Today, the MBTI is one of the most widely used personality test by psychologists, counselors, social workers, and other mental health professionals in the world. Industrial/Organizational psychologists and human resource specialists also utilize the MBTI for career assessment, employee training, and team building programs (Joyce, 2010).

Keirsey Theory. David Keirsey's *The Keisey Temperament Sorter* provides a brief, self-scoring, temperament measure that yields the MBTI 16 types (Keirsey & Bates, 1978). However, he introduced a modified interpretation of the 16 types into four clusters of interpretation based on the work of other theorists (i.e., Ernst Kretschmer, Eduard Spranger, Eric Adickes, and Eric Fromm). The four clusters included sensing-judging, sensing-perceiving, intuition-thinking, and intuition-feeling.

The sensing-judging temperament is characterized as responsible, conservative, stable, productive, organized, and a strong work ethic. Persons who prefer this temperament type are compelled to fulfill obligations, are industrious, have a high need for belongingness, and often are caretakers. The sensing-perceiving temperament is characterized as active, spontaneous, open-minded, and adaptable. Persons who display this temperament are cheerful and enjoy exploration and discovery. They have a high need for freedom and have a strong play ethic (Keirsey & Bates, 1978).

The intuition-thinking temperament is characterized as rational, analytical, systematic, and research-oriented. Persons who display this temperament generally exhibit high standards, are curious, display a high need for achievement, are perfectionistic, and may be compulsive. The core value of an intuitive-thinking person is to develop competencies and skills. They put work before play and can incorporate play into work by striving to develop recreation skills, such as golfing expertise. The intuitive-feeling temperament is characterized as friendly, imaginative, caring, and sensitive to the needs of others. Persons who display this temperament are non-competitive and passionate about social causes and the impact of their actions on humanity. The core value of intuitive-feeling person is personal integrity and self-actualization (Keirsey & Bates, 1978).

Temperament Construct in Children and Youth

Early theory and research on temperament were conceptualized on adult behavior patterns. However, theorists agree that temperament dispositions are early appearing and that temperament differences are present as early as infancy, thus suggesting temperament is innate or biologically rooted (Jung, 1928/1945; Rothbart, 1989; Teglasi, 1998). Differences in infants' activity levels, sociability, and emotionality characterize temperament (Buss, 2011). Jung (1928/1945) described infants' adaptation to the environment, especially how quick they interact with objects and other people, as an early sign of extroversion. Indicators of introversion in children include shyness, reflective thinking, and fearfulness of unknown objects.

Thomas and Chess Temperament Theory

Thomas and Chess (1977, 1984, 1986, & 1989) developed a framework for understanding children's temperaments based on behavioral characteristics. They explained the similarity between behavioral and temperament style. Behavioral style describes how a person behaves, rather than how well, what (i.e., abilities and content), or why (i.e., motivation) a person behaves. Moreover, behavioral style refers to behavioral characteristics that are present at birth through later life. In contrast, temperament refers to behavioral tendencies that appear in early infancy.

Thomas and Chess (1977, 1984, 1986, & 1989) considered development as a complex process of interplay between the child and the environment. They believed that temperament is best understood when it is explained in the context of the environment. They used the concept of goodness of fit to explain the temperament-environment interactive process. Goodness of fit is achieved when there is a match between the child's desired characteristics and the environment's desired characteristics. Poorness of fit is the consequence of a mismatch between the two sets of desired characteristics.

Thomas and Chess (1977, 1984, 1986, & 1989) described nine behavioral categories of temperaments based on data obtained though observation, parent questionnaires, and teacher interviews regarding the infancy periods of 22 children. Content and factor analysis of the data differentiated children's behaviors into three core temperament patterns: easy, difficult, and slow-to-warm-up. Easy temperament was displayed by approximately 40% of the infants. They were described as being able to establish regular routines, cheerful, and able to adapt easily to new situations. Parents further described these children as contented and easy going. A difficult temperament was displayed by approximately 10% of the infants. About 70% of these children were expected to encounter long-term adjustment problems at some point in the future. They were described as experiencing irregular routines, were slow to adapt to new circumstances, displayed problematic sleep cycles, and tended to react negatively. Slow-to-warm up temperament was displayed by approximately 15% of the infants. They were described as watchful with strangers, lethargic, tending to display a negative mood, and adjust rather slow to new situations. Approximately 35% of the infants displayed a combination of temperament patterns.

Psychological Type Approach to Temperament

The psychological type approach was adopted by the SSQ to understand children's temperament qualities. The SSQ provides information describing students' temperament in eight basic styles and grouped into four bipolar traits, namely extroverted or introverted, practical or imaginative, thinking or feeling, and organized or flexible. In turn, the basic styles can form 16 meaningful combinations with four styles (e.g., extroverted-practical-thinking-organized) per combination (Oakland, Glutting, & Horton, 1996).

Extroverted-introverted traits. Extroverted-introverted traits are associated with the source from which a student derives his or her energy. Approximately 65% of U.S. children prefer extroverted style and 35% prefer introverted style (Oakland, Glutting, & Horton, 1996). Children who display a preference for extroverted style derive energy from the external environment and are more oriented to people and events. On the other hand, children who prefer an introverted style derive energy from themselves or their immediate family and close friends and are more oriented to ideas and reflection.

Children who prefer an extroverted style enjoy meeting and interacting with people. They exhibit high energy to socialize with people and are open to discussions about different topics. Extroverted children are drawn to participate in group activities and would feel uncomfortable when required to spend too much time alone. They are quick to move from one space to another, getting to know people's names and faces. They develop their thoughts and ideas through oral discussion and thrive in a learning environment where there are ample opportunities to engage in hands-on activities. Positive behaviors of extroverted students are reinforced by encouragement and praises. Discussion is the primary means to develop their thoughts and ideas.

Extroverted children may not be comfortable to work on long assignments or seatwork. They also tend to act without thinking and planning. Their penchant for interaction may disrupt the classroom environment and when extreme, may be annoying to students who display strong preference for introverted style.

In contrast, children who prefer an introverted style prefer to be alone and may feel exhausted and bored with frequent and long interaction with others. They are slow in responding to their surrounding and have a wait-and-see tendency before joining an activity. Introverted children are cautious, have fewer friends, and prefer to listen than talk in a group. They can work for long periods on a project without interruption.

Introverted children tend to be selective of people they want to be close with. Their reserved and quiet manner sometimes is misinterpreted as displaying poor social skills. In school, they prefer to work with other children who match their behavioral qualities in small groups or in pairs. They like doing tasks that require methodical, reflective, and reasoning skills. They prefer to be called by the teacher to recite in class rather than volunteer to recite. At times, teachers and other adults in school would label introverted students as unfriendly, uncooperative, and withdrawn because of their preference to be alone and quiet.

Practical-imaginative traits. Practical-imaginative traits are associated with what students attend to. Approximately 65% of U.S. children prefer a practical style and 35% prefer imaginative style (Oakland, Glutting, & Horton, 1996). Children who prefer a practical style are more realistic and pragmatic in their view of the world. They tend to pay more attention to details and facts when acquiring information. Practical children are keen observers and use their physical senses to know more about their environment. They enjoy family traditions, recreation, and leisure time. They find the company of their extended family members satisfying.

Practical children enjoy learning through a step-by-step approach. They are organized, have good rote memory, and persist in working toward their goals. For them, the present is more important and they prefer to do things that have practical value in their lives. Practical children are more concrete learners and thus, become disinterested in learning complex material and abstract concepts.

In contrast, children who prefer an imaginative style enjoy activities that harness their creative skills and qualities. They also enjoy acquiring new skills and do things in non-traditional ways. They love ideas and concepts rather than factual and sensory knowledge. Imaginative students are more engaged in tasks that require analysis, comprehension, and formulating conclusions. They prefer to learn the theory first and proceed to know its usefulness and application. Imaginative children are challenged by difficult situations and may confidently take risks.

Imaginative children value quickness of understanding, especially in learning novel tasks. They enjoy testing hypotheses and generating insights about their experiences. Within the family, imaginative children generally have close attachments to their parents, siblings, and the extended family members. In school, they prefer to work on projects and assignments that require imagination and analysis. They are fascinated to learn new theories, even trivial ones that lack empirical evidence. At times, they come up with very complex and sophisticated plans that can be difficult to implement.

Thinking-feeling traits. Thinking-feeling traits are associated with how children make decisions. These traits are significantly influenced by gender. Approximately 65% of U.S. males and 35% of U.S. females prefer a thinking style. On the other hand, approximately 65% of U.S. females and 35% of U.S. males prefer a feeling style (Oakland, Glutting, & Horton, 1996).

Children who prefer a thinking style are likely to be fair and objective in making decisions. They attach more importance to logic than emotions. Thinking children are more engaged in activities that require analysis of facts and ideas. They are more likely to exhibit a skeptical stance on what other says and typically ask questions to challenge other's ideas. They are consistent with their opinions and beliefs.

In social situations, children who prefer thinking style are likely to enjoy the company of peers who share common interests. When with friends, they prefer to do tasks that stimulate their logical and analytical skills rather than engage in social conversation and other more socially oriented activities. They are open with their opinions that may not be supported by evidence. Their tendency to express criticisms may be offending to others. In the family, students who prefer thinking style tend to be close to their parents and other family members. However, they are less likely to express their love and affection, both verbally and non-verbally. They cooperate and follow family rules that seem fair and reasonable.

As students, they are more likely to sustain attention during lectures and presentation that are organized logically. They are more likely to enjoy subjects (e.g., math and science) that require the use of logical and analytical skills. Feedback to them is beneficial when teachers point out their errors and how to correct them.

In contrast, children who prefer a feeling style rely on their emotions and interpersonal relationship when making decisions. They are more engaged with people, hence are friendly, charming, sympathetic, and generous with appreciation. They support or participate in projects that promote human welfare, peace, and harmony.

In social relationships, children who prefer feeling style tend to avoid conflict or disagreement with others that may strain relationships. They give more importance to friendships and are careful not to hurt feelings of others. In the family, children who prefer feeling style tend to be very expressive with their love and affection to parents and other family member. They are likely to please their parents by doing chores and following rules at home. They are caring, empathetic, and supportive. Conflicts or disagreements in the family cause them distress.

As students, they tend to enjoy subjects or lessons that deal with people. Praises and other forms of affirmation are effective strategies to motivate them to perform well. In group activities, they enjoy working with friends and may even accept more responsibilities to fill in for others. They shy away from activities that promote competition rather than cooperation.

Organized-flexible traits. Organized-flexible traits are associated with a tendency to either make decisions rather quickly or to postpone them. Approximately equal proportions of U.S. children prefer organized and flexible styles (Oakland, Glutting, & Horton, 1996).

Children who prefer an organized style are likely to have structure and plan their activities. They have an admirable work ethic; are persistent, dependable, and exhibit self-discipline. They generally complete tasks in a systematic way. Organized students generally show respect to authority.

In social relationships, children who prefer organized style generally want their friends to conform to their expectations. Loyalty and support to friends are defining qualities. They tend to devote much time to prepare for social activities and may be upset when changes are made on the plans, especially at the last minute. In the family, they enjoy a well-planned home life and expect parents and other family members to fulfill their promises. They generally are obedient to rules and do chores in an orderly way.

As students, they generally have good study habits and are in control of their school activities. They are organized in their work and devote ample time to prepare for their activities. Teachers and other adults in school perceive students who display proclivity for organized style as responsible, obedient, and dependable. They expect teachers to provide explicit instructions and guidelines related to grading, examinations, assignments, and classroom behavior. They generally put forth their best effort to attain goals and usually are motivated to maintain good work and behavior through consistent praise and rewards.

In contrast, children who prefer a flexible style tend to adjust easily to changes in their lives and environment. They enjoy surprises and may tend to avoid following rules. They are open in their commitments and simply want to take as many opportunities as possible to enjoy life. They are likely to be accepting of diversity and tolerant of opinions from others.

In social relationships, children who are inclined to flexible style may be considered the clown of the group because they enjoy making people laugh and have fun. They seek fun most of the time and would have difficulty coping with stress and problems. In the family, children who prefer a flexible style enjoy fun activities with family members. They have a high need for autonomy and may resist rules at home. They may not keep their room or personal space organized or neat.

As students, they are likely to enjoy lessons and activities if presented in a game-like manner. They like to participate and perform in school activities, especially when there are incentives for winning. They thrive in classroom environments in which students are given options, have flexible deadlines in assignment and projects, can be permitted to move around the room, and are expected to follow only a few necessary rules.

Developmental Perspective of Temperament

The New York Longitudinal Study reported that temperament differences between males and females appear shortly after infancy and increase with age on the following qualities: adaptability, approach/withdrawal, activity, and sensory threshold (Chess & Thomas, 1991). During the period from four months to four years, males are more adaptable and approaching than females. Between ages 8 to 12, males display higher levels of activity and sensitivity (Maziade et al., 1986).

Studies of temperament style of U.S. children, ages 8-17, confirm the presence of gender differences (Bassett & Oakland, 2009; Oakland, Glutting, & Horton, 1996). As noted previously, more males than females prefer thinking and flexible styles while more females than males prefer organize and feeling styles. Gender differences on thinking-feeling appear early, at least by age 8, are sustained through adulthood, and may be universal (Hammer & Mitchell, 1996; Myers, McCaulley, Quenk, & Hammer, 1998; Myers & McCaulley, 1985).

Temperament Preferences of Children in Special Populations

Temperament traits are important characteristics that can provide a non-pathological framework for developing children's strengths and identifying characteristics that require intervention (Joyce & Oakland, 2005). An understanding of temperament traits of children with special needs can contribute to building a more positive interactions and relationships between these children and their significant adults (e.g., their parents and teachers). Likewise, psychologists and other mental health professionals can use temperament information when designing interventions and determining possible reinforcement strategies in behavioral modification plans for children. Few studies used the four bipolar traits to describe temperament of children with special needs.

Children with Anxiety Disorder and Depressive Disorder

The temperament traits of 70 middle school children who met the criteria for eligibility as emotionally disturbed, with a primary diagnosis of either depression or anxiety, were examined (Jennings, 2005). No specific temperament style was associated with conditions of depression and anxiety. In general, children diagnosed with either depression or anxiety indicated stronger preference for extroversion than introversion, practical than imaginative, thinking than feeling, and organized than flexible styles.

In light of the goodness-of-fit hypothesis (Thomas & Chess, 1984), preferences for organized-flexible styles in depressed and anxious children were examined in relation to children's preferences for classroom setting. Compared to children with depression, children with anxiety were expected to prefer an organized style and a classroom setting that placed greater emphasis on structure and conformity to rules. Furthermore, children who displayed both depression and anxiety and who prefer a flexible style were expected to have a stronger preference for a classroom setting that placed lesser emphasis on structure and conformity to rules. Both hypotheses were disconfirmed. Children who were anxious did not differ from those who were depressed in their preferences for organized style and classroom setting that placed greater emphasis on structure and conformity to rules. Moreover, preference for a flexible style was not associated with preference for a classroom setting that place lesser emphasis on structure and conformity to rules. Children who either were anxious and depressed indicated a preference for a teacher who displayed with high emotional responsiveness regardless of classroom conformity (Jennings, 2005).

Children with Oppositional Defiant Disorder (ODD) and Conduct Disorder (CD)

Temperament traits of 80 children and youth ages 8 through 17, and with a current diagnosis of either Oppositional Defiant Disorder (ODD) or Conduct Disorder (CD), were compared (Joyce & Oakland, 2005). Children and youth with CD and ODD differed only on practical-imaginative styles; those with ODD displayed a stronger preference for practical style. Furthermore, those with ODD also displayed a greater preference for practical and thinking styles than do their general population peers.

Knowledge that children and youth with ODD prefer practical and thinking styles can be useful in developing treatment plans. Children with practical styles generally prefer simplicity, details, facts, and the concrete and sequential processing of ideas. A preference for a practical style among children and youth with ODD can be associated with their need for explicit behavioral expectations along with specific consequences (Joyce & Oakland, 2005). In addition, concrete and sequential communication of expectations by parenting programs is beneficial in modifying children’s defiant behaviors (Barkley, 1997; Patterson, 1982).

On the other hand, the stronger preference of children and youth with ODD for a thinking style can be considered a weakness since this can be associated with a negative attribution bias (Dodge & Newman, 1981). Consequently, children with ODD may tend to express hostility in the way they relate with people. Moreover, their inability to regulate the expression of their blunt and critical opinions may increase their defiance and aggressive behaviors (Joyce & Oakland, 2005).

Children with Attention Deficit-Hyperactivity Disorder (ADHD)

Temperament traits of 83 children, ages 8 through 12, with a current diagnosis of Attention Deficit-Hyperactivity Disorder (ADHD) were compared with a matched group comprising of 84 non-ADHD children (Harrier, 2005). Children with ADHD generally prefer extroverted over introverted, practical over imaginative, thinking over feeling, and flexible over organized styles. Although more non-ADHD children prefer imaginative over practical style and organized over flexible style, the proportion of children with and without ADHD who displayed preferences for each of the four bipolar traits was comparable. Furthermore, children with and without ADHD displayed moderate levels of preferences for all four bipolar traits.

Regardless of the diagnosis, gender differences in the preferences for thinking-feeling styles were significant. More males displayed a preference for thinking style and more females displayed a preference for a feeling style. In addition, more females displayed a preference for organized styles and more males displayed a preference for flexible styles. Among males, those with ADHD displayed a higher preference for a flexible style than those who did not display ADHD.

Thus, while ADHD is not strongly associated with temperament traits in children, the preference among males with ADHD for a flexible style may warrant attention to academic and behavioral accommodations in school and home settings (Harrier, 2005)

Children with Autism Spectrum Disorder (ASD)

The temperament traits of children with high functioning autism were examined, in part, to determine whether a mismatch occurs between temperament styles of ASD children and their parents. Additionally, the study examined whether a mismatch in temperament styles increases parenting stress experienced by parents of children with ASD (Darby, 2009). In general, children with high functioning autism generally prefer introverted over extroverted, practical over imaginative, feeling over thinking, and organized over flexible styles. Temperament preferences of children with ASD and their parents were more often matched on the extroverted-introverted (62%) and organized-flexible (53%) styles. On the other hand, the match between temperament preferences of children with ASD children and their parents was less likely on practical-imaginative (68%) and thinking-feeling (44%) styles. Parents report higher levels of stress when they and their children with ASD display similar preferences for thinking-feeling styles and lower levels of stress when they and their children display different preferences for thinking-feeling styles along with their children with ASD (Darby, 2009).

Children with Learning Disabilities

Temperament traits of 84 Canadian children, ages 6 through 14, with diagnosed learning difficulties in reading, spelling, language arts, and arithmetic were compared with a matched group of children without any reported learning difficulties (Danielsen, 1991). Children with and without learning difficulties preferred extroverted over introverted, imaginative over practical, feeling over thinking, and flexible over organized styles. However, more children with learning difficulties preferred extroverted style while approximately equal proportions of children without learning difficulties prefer extroverted and introverted styles

Children Identified as Gifted

A number of studies examined the temperament styles of gifted children across gender and ages. In general, gifted children displayed preferences for imaginative over practical (Bireley, 1991; Cross, Speirs Neumeister, & Cassidy, 2007; Oakland, Joyce, Horton, & Glutting, 2000) and flexible over organized (Cross, Speirs Neumeister, & Cassidy, 2007; Oakland, Joyce, Horton, & Glutting, 2000) styles. Their preferences for extroverted-introverted (Beckner, 1990; Dempsey, 1975; Gallagner, 1990; Mills, 1993; Parker & Robinsor, 1989; Oakland, Joyce, Horton, & Glutting, 2000) and thinking-feelings (Beckner, 1990; Gallagner, 1990; Oakland, Joyce, Horton, & Glutting, 2000) styles were somewhat balanced. Gender differences among gifted children generally were apparent in practical-imaginative, thinking-feeling, and organized-flexible styles. More females than males prefer imaginative and feeling styles and more males than females prefer thinking and flexible style. Compared to non-gifted children, gifted children are more likely to prefer an imaginative style, especially among gifted females (Cross, Speirs Neumeister, & Cassidy, 2007; Oakland, Joyce, Horton, & Glutting, 2000).

Children with Visual Impairments

Temperament traits of visually impaired children ages 10 through 17, were examined across gender and age. In general, children who are visually impaired displayed preferences for practical over imaginative and organized over flexible styles (Oakland, Banner, & Livingston, 2000). Differences in preferences for extroverted-introverted styles (Oakland, Banner, & Livingston, 2000; Pinner & Forlano, 1943; Zahran, 1965) or thinking-feeling styles (Oakland, Banner, & Livingston, 2000) were not apparent. Gender differences were apparent. Males were more likely to prefer extroverted, thinking, and flexible styles while females more likely to prefer introverted, feeling, and organized styles. Age differences were apparent in thinking-feeling and organized-flexible styles. Younger children were more likely to prefer thinking and organized styles while older children more likely to prefer feeling and flexible styles (Oakland, Banner, & Livingston, 2000).

Correlation between Temperament and Cognitive Variables

Temperament traits are known to impact children's behaviors and performance. Thus, several studies explored the association of children's temperament traits with cognitive variables.

Several studies reported mix results on the relationship between temperament preferences and cognitive variables. One study found consistent positive relationships between introverted style and students' achievement and intelligence test scores (Myers & McCaulley, 1985). However, another study of elementary and high school students who displayed a preference for extroverted style displayed higher achievement than students who displayed a preference for introverted style (Tobacyk, Hearn, & Wells, 1990). Children who preferred imaginative and organized styles were found to do better academically than children who preferred practical and flexible styles (Fourgurean, Meisgeier, & Swank, 1988; Kaufman, McLean, & Lincoln, 1996; Myers & McCaulley, 1985; Tobacyk, Hearn, & Wells, 1990). In addition, students who displayed a preference for thinking styles displayed higher math achievement (Davis, 2007). However, students who displayed a preference for feeling style had higher grades than students who display a preference for thinking style.

Contrary to these summarized results, the SSQ manual reports students' temperament preferences were not related to achievement as measured by California Achievement Test (CTB/MacGraw-Hill, 1985) and to their intelligence as measured by the Wechsler Intelligence Scale for Children-Revised (WISC-R: Wechsler, 1994). Obtained correlations between extroverted-introverted, practical-imaginative, thinking-feeling, and organized-flexible styles and achievement and IQ generally were low and not significant. Thus, temperament traits may be independent of achievement and intelligence (Oakland, Glutting, & Horton, 1996).

In conclusion, some evidences support the relationship between imaginative style, intelligence, and achievement. Studies that reported this significant relationship obtained data from samples with a broad age range (e.g., ages 16-94). Thus, the non-significant results obtained by Oakland, Glutting, and Horton (1996) can be explained by having younger participants and limited age range (i.e., ages 8-17).

Cross-national Differences in Temperament in Children

International interest on children's temperament has led to considerable scholarship, including cross-national studies on children's temperament traits.

Temperament preferences of children from Australia (Oakland, Faulkner, & Bassett, 2005), Costa Rica (Oakland & Mata, 2007), Gaza (Oakland, Alghorani, & Lee, 2006), Greece (Oakland & Hatzichristou, 2010), Hungary (Katona & Oakland, 2000), India (Oakland, Singh, Callueng, & Goen, 2011), Japan (Callueng, de Carvalho, Isobe, & Oakland, 2012), Nigeria (Oakland, Mogaji, & Dempsey, 2006), Pakistan (Oakland, Rizwan, Aftab, & Callueng, 2011), People's Republic of China (Oakland & Lu, 2006), Romania (Oakland, Illiescu, Dinca, & Dempsey, 2009), South Africa (Oakland & Pretorisus, 2009), Samoa (Callueng, Lee Hang, Gonzales, Ling-So'o, & Oakland, 2011), South Korea (Oakland & Lee, 2010), United States (Bassett & Oakland, 2009), Venezuela (Leon, Oakland, Wei, & Berrios, 2009), and Zimbabwe (Oakland, Mpofu, & Sulkowski, 2007) were examined in light of the four bipolar traits measured by the SSQ.

A synthesis of data from these 17 countries suggests that children display several prevailing temperament qualities. For example, children from 13 of the 17 countries generally show a preference for an extroversion style: Australia, Costa Rica, Greece, India, Japan, Pakistan, People's Republic of China, Romania, Samoa, South Korea, United States, Venezuela, and Zimbabwe. In contrast, children from Hungary and Nigeria generally show a preference for an introversion style. Children from Gaza and South Africa show a somewhat balanced preference for extroversion and introversion styles. U.S. children's preference for extroversion increases from 8 to 13 (Oakland et al., 1996; Bassett, 2005).

Gender differences in the thinking and feeling styles are apparent in children from Australia, Costa Rica, Gaza, Greece, Hungary, Japan, Nigeria, Pakistan People's Republic of China, South Africa, Venezuela, and United States. Males generally prefer a thinking style and females generally prefer a feeling style. Both male and female children from Egypt, India, South Africa, South Korea, and Zimbabwe show a general preference for a feeling style. Samoan male and female children show a preference for a thinking style.

Children from 12 countries also show a general preference for a practical style: Gaza, Greece, Hungary, India, Japan, Nigeria, Pakistan, People’s Republic of China, Romania, Samoa, South Africa, Venezuela, and Zimbabwe. In contrast, children from Australia, Costa Rica, South Korea, and United States generally show a preference for an imaginative style. Children from Greece have a somewhat balanced preference for practical and imaginative styles. Gender differences are evident among children from China, Pakistan, and United States. In these countries, more males prefer an imaginative style and more females prefer a practical style. Age differences also are evident. Increased preference for a practical style is seen in older children from South Korea while increased preference for imaginative style is seen in older children from Egypt and the United States.

Children from 16 of the 17 countries generally show a preference for an organized style: Australia, Costa Rica, Gaza, Greece, Hungary, India, Japan, Nigeria, Pakistan, People’s Republic of China, Romania, Samoa, South Africa, United States., Venezuela, and, Zimbabwe. In contrast, only children from South Korea show a general preference for a flexible style. Children from Costa Rica, Gaza, Greece, Japan, Nigeria, Pakistan, and the United States show gender related differences in organized-flexible styles. Males from these countries are more likely to prefer an organized style and females are more likely to prefer a flexible style. Similarly, age related differences are evident in children from Costa Rica, Greece, Japan, Nigeria, Pakistan, People’s Republic of China, Romania, South Africa, South Korea, Venezuela, and Zimbabwe.

Preference of children from these countries for flexible style increases with age. However, a greater proportion of older children from Greece and the United States display a preference for organized style.

Test Validity

Validity is an essential property of a test that refers to the "*degree to which evidence and theory support the interpretation of test scores entailed by proposed uses of test* (AERA, APA, NCME, 1999, p. 9). Validity pertains to a judgment about test data obtained for a specific purpose and setting. Test validation is an on-going activity of collecting information to increase understanding of test results. The validation process is a shared task of the test developer and the test user. The test developer is responsible in formulating the conceptual framework and the rationale for a test. On the other hand, a qualified test user conducts research using test data to support the purposes of a test or even expand the uses of an existing test (Urbina, 2004).

The following are the sources of validity evidence.

Evidence based on test content.. Test content includes tasks or questions on a test, including item format and wording. It also covers guidelines and procedures for test administration and scoring. Validity evidence based on test content entails establishing a match or congruence between the content and the construct the test is purported to measure. This approach requires test developers to ensure that test content is adequate and representative of the content domain as well as the relevance of content domain to test score interpretation. Developing a test specification of the content domain is a common requirement to guarantee adequacy and representative of content domain to what the test intends to measure. Another approach to establish validity evidence based on content is through expert judgment. This approach requires experts to evaluate the extent to which test parts and construct are related (AERA, APA, NCME, 1999).

Evidence based on response processes. Response processes of examinees can give information about the congruence of the construct and the response or actual performance of examinees. Evidence based on response processes requires the analysis of responses such as inquiring examinees about their response styles or actual responses to items. Information obtained from this procedure can clarify and enhance the definition of the construct. Analyzing patterns of responses in various test parts and how these are related to other individual and situational factors also contributes to understanding the construct (AERA, APA, NCME, 1999)

Group variations in meaning and interpretation of test scores can be clarified by obtaining information response processes. In addition, assessment that requires the recording or evaluating of performance or products of examinees by observers or judges can impact test validity. Hence, it is essential to examine how observers and judges record and evaluate data, and how these data are relevant to the test interpretation (AERA, APA, NCME, 1999).

Evidence based on internal structure. Validity evidence that is intended to examine internal structure contributes to the general understanding of the conceptual framework of the test. Evidence of internal structure pertains to whether the correlations among test items or test domains fit to the construct on which test interpretation is anchored. The type of analysis used to determine internal structure depends on the purpose of the test. Empirical evidence is provided to support interpretation of test scores for a multidimensional or a unidimensional internal structure. In some cases, evidence of internal structure is demonstrated through differential item functioning, That is, items are analyzed to examine if different groups of examinees that have similar characteristics have different responses on the same item. Differences in item functioning can suggest possible multidimensionality of the internal structure of a test (AERA, APA, NCME, 1999).

Evidence based on relations to other variables. Correlations of test scores with other variables constitute another source of validity evidence. Evidence based on relations to other variables can be in the form of convergent and discriminant evidence, test criterion relationships, and validity generalization. Convergent validity refers to the relationships between test scores with other measures that assess similar constructs. Discriminant validity pertains to the relationships of test scores with other measures that assess different constructs. Convergent and discriminant validity evidences can be examined using experimental and correlational research designs (AERA, APA, NCME, 1999).

Evidence on test-criterion relationships has the goal to determine how accurate test scores predict criterion behavior or performance. A criterion variable is a measure of a characteristic or quality that is determined by test users. When conducting a test-criterion validity study, findings would be interpreted in light of the relevance and reliability of criterion measure to the intended use of a test. Test-criterion relationships are established through concurrent or predictive design. In concurrent validity study, the test scores (predictor) and the criterion variable are obtained at the same time. Predictive validity study assesses the extent to which a test can predict a criterion scores obtained at a later time **(**AERA, APA, NCME, 1999).

Validity generalization refers to how accurately findings on test-criterion relationships can be applied to a new situation without a systematic investigation on validity in that new situation. Validity generalization employs meta-analytic design that requires collecting validation studies conducted in similar situations and make statistical summaries of these studies that may be useful in calculating test-criterion relationships in a new situation. Meta-analysis for the purpose of validity generalization can vary in terms of situational factors: 1) differences in predictor variable measurement, 2) job or curriculum involved, 3) criterion measures used, 4) test taker characteristics, and 5) time when the study was conducted. These situational factors are included in meta-analysis to determine if variations in these factors have significant effect on test-criterion relationships **(**AERA, APA, NCME, 1999).

Research Objectives and Hypotheses

This study examines cross-nationally the four-factor structure of temperament as measured by the Student Styles Questionnaire. Independent factor solution by country is conducted within the item factor analysis framework for categorical or dichotomous variables. Subsequently, measurement invariance is examined in countries that exhibit a satisfactory fit to the four-factor model of temperament.

In light of the research objectives, it is hypothesized that the data from each country have a good fit to the four-factor model of temperament. It is further hypothesized that the four-factor model of temperament is invariant across countries

CHAPTER 2
METHODS

Participants

Data were collected on 17,867 children from 21 countries as part of an international research program of Dr. Thomas Oakland in collaboration with approximately 50 on-site researchers. I also acquired some of the data independently of these efforts. The sample consists of children ages 8 through 17 from diverse geographical regions: Africa (Egypt, Nigeria, and Zimbabwe), Central America (Costa Rica), North America (United States), South America (Brazil and Venezuela), Europe (Hungary, Poland, and Romania), East Asia (China, Japan, and Mongolia), Southeast Asia (Philippines and Singapore), South Asia (Pakistan), Middle East (Gaza, Iran, and Israel), and Oceania (Australia and Samoa). As seen in Table 2-1, sample sizes of individual countries range from 253 (Israel) to 7,902 (United States). Gender distribution was comparable in almost all countries and for the combined group (49.53% males).

The children attended public and private schools and lived mostly in urban settings. Samples from Romania and the U.S. were drawn from their respective standardization samples and stratified according to their national census statistics. Samples from other countries were obtained through research in order to describe the preferred temperament styles of children and to examine possible age and gender differences. Children came from families of diverse socio-economic status, which was indirectly measured based on type of school (i.e., public and private) and place of residence (i.e., rural and urban). Sample characteristics by country are described below.

Australia. The Australia sample included 308 children enrolled in public and private primary and secondary schools in the provincial city of Bendigo located in the state of Victoria. Children's ages ranged from 9 through 15, with approximately equal number of males and females. The sample included children from families who display diverse socio-economic status.

Brazil. The Brazil sample included 461 children enrolled in public and private schools. Children's ages ranged from 9 through 15, with approximately 55% females. The sample included children from families who display diverse socio-economic status.

China. The China sample included 400 children enrolled in public schools in the city of Taiyuan, the provincial capital of Shanxi Province. Children's ages ranged from 9 through 15, with equal number of males and females. The sample included children from middle to upper lower socio-economic families.

Costa Rica. The Costa Rica sample included 432 children enrolled in public and private urban schools. Children's ages ranged from 9 through 15, with approximately 50% males in each age. The sample also included children who generally come from middle class families.

Egypt. The Egypt sample included 800 children enrolled in public schools in the city of Asyut in the southern region of Upper Egypt. Children's ages ranged from 9 through 15, with approximately 50% males. Children came from middle to upper middle socioeconomic families.

Gaza. The Palestine sample included 400 children enrolled in public schools in the city of Gaza. Children's ages ranged from 9 through 17, with equal number of males and females. The sample included children from families who display diverse socio-economic status.

Hungary. The Hungary sample included 400 children enrolled in public schools. Children's ages ranged from 9 through 16, with equal number of males and females. The sample included children from families who display diverse socio-economic status.

Iran. The Iran sample included 511 children enrolled in public schools. Children's ages ranged from 8 through 15, with approximately 55% females. The sample included children from families who display diverse socio-economic status.

Israel. The Israel sample included 253 children enrolled in public schools. Children's ages ranged from 10 to 12, with approximately 50% males. The sample included children from families who display diverse socio-economic status.

Japan. The Japan sample included 493 children enrolled in public schools in Miyakonojo, a city of 170,000 inhabitants and one of the main metropolitan areas of the southern Kyushu region in Japan. Children's ages ranged from 9 through 16, with approximately 50% males. All children came from lower to upper middle class families.

Mongolia. The Mongolia sample included 1015 children enrolled in public schools in the cities of Ulaanbaatar and Khovd. Children's ages ranged from 9 through 16 years, with approximately equal number of males and females. The sample included children from lower to upper middle class families.

Nigeria. The Nigeria sample included 400 children enrolled in public schools located in Lagos. Children's ages ranged from 9 through 17, with 50% males across ages. The sample represents children from the country's three major ethic groups: Yoruba from the south-west region, Ibo from the southeast region, and Hausa/Fulani from the north region. The sample included children from families who display diverse socio-economic status.

Pakistan. The Pakistan sample included 463 children enrolled in private schools in the cities of Karachi, Islamabad, and Quetta. Children's ages ranged from 9 through 15, with approximately 50% males for each age group. The sample included children from middle to upper class families.

Philippines. The Philippine sample included 400 children enrolled in public schools and private schools in Manila. Children's ages ranged from 9 through 15, with approximately 50% males. The sample included children from low, middle, and upper class families.

Poland. The Poland sample included 440 children enrolled in public schools. Children's ages ranged from 9 through 17, with approximately 50% males for each age group. The sample included children from lower to middle class families.

Romania. The Romania sample included 900 children enrolled in various public schools in Romania. Children's ages ranged from 9 through 17, with equal number of males and females. Children came from a cross-section of socio-economic backgrounds. Approximately 86% of the children are of Romanian nationality while the remaining 14% represented other nationalities, including Hungarian and German.

Samoa. The Samoa sample included 400 children enrolled in a public school in the city of Apia, the national capital of Samoa. Children's ages ranged from 9 through 16, with equal number of males and females. The sample included children from families who display diverse socio-economic status.

Singapore. The Singapore sample included 483 children enrolled in public schools Children's ages ranged from 9 through 15, with approximately 53% females. The sample included children from middle class families.

United States. The United States sample included 7902 children drawn from the standardization sample of the SSQ and representative of the 1990 U.S. Bureau of the Census Statistics. Children attending public and private schools located in 29 U.S. states and Puerto Rico, ages ranged from 8 through 17, with equal number of males and females. The sample was also representative of major racial/ethnic groups: Whites (n=5,547), African Americans (1,194), and Hispanics (868). The remaining 293 children were from other racial/ethnic groups (e.g., Asian).

Venezuela. The Venezuela sample included 411 children enrolled in public and private schools located in four major cities: Caracas, Anzoategui, Bolivar, and Zulia. Children's ages ranged from 9 through 16, with approximately equal number of males and females. Children came from middle to upper class families.

Zimbabwe. The Zimbabwe sample included 600 children enrolled in public schools. Children's ages ranged from 8 through 15, with equal number of males and females. Children came from middle to upper class families.

Measure

The SSQ is a self-report paper and pencil group administered measure of temperament traits for students ages 8 through 17. Each of its 69 forced-choice items has two alternatives that provide for an assessment of preferred behaviors associated with one of four bipolar traits: extroversion-introversion (EI), practical-imaginative (PM), thinking-feeling (TF), and organized-flexible (OL). These items were developed and empirically verified to support the Jungian and Myers-Briggs temperament theory. The EI scale has 23 items, the PM scale has 16 items, the TF scale has 10 items, and the OL scale has 23 items. Additionally, six items provide information simultaneously on two scales.

The SSQ can be completed in 20 to 30 minutes. Each item requires a student to choose between two options related to the same bipolar dimension (e.g., a thinking option and a feeling option). SSQ scores provide information corresponding to three levels of interpretation. The basic level provides information that focuses on the eight styles and four bipolar traits. The second level provides information pertaining to the combination of two styles. The third level provides information on 16 possible combinations of style preferences, and is patterned after the Myers-Briggs model that combines four styles.

The reliability of SSQ scores of U.S. children was established through test-retest procedure over a period of seven months using prevalence-based standard scores of 137 students, with ages ranging between 8 through 17. The sample comprised of 75 males and 62 females, with gender distribution approximately equal across age levels. Test-retest reliability coefficients were .80 for EI, .67 for PM, .78 for TF, and .74 for OL; with a mean reliability coefficient of .74. Change scores between completions ranged from .50 (EI) through 2.20 (PM) (Oakland, Glutting, & Horton, 1996).

The validity of the SSQ was assessed using several methods. Internal validity of the SSQ was established following the substantive-construct model of instrument development. Items describing behavioral indicators of the four bipolar traits were developed and subjected to item review by nationally known scholars in temperament theory and learning styles. A total of 100 items were selected after the initial factor analysis and were subjected to further analysis during SSQ’s standardization phase. Sequential exploratory factor analysis (EFA) was conducted to address the problem of restricted variance of dichotomous items. First, an EFA was employed directly on the dichotomous scores with forced rotation of the hypothesized four factors. Results indicated that 69 out of the 100 items met the two criteria set to evaluate factor solution: 1) ≥ .20 item loading on the hypothesized factor, and 2) 10% or fewer items cross loading on two or more factors. Overall, the direct item EFA supported SSQ's hypothesized four-factor structure. Factor analysis of item parcels was repeated 17 times, with one of the parcels "opened". This procedure of using open and closed parcels was intended to control the problem of inadequately conditioned correlation matrices as well as confirming if individual items in the opened parcel would load on their hypothesized factor. The four-factor structure of the SSQ was stable when separate factor analysis using item parcels were conducted across three age groups of students who comprised the standardization sample (Oakland, Glutting, & Horton, 1996).

Independence of the SSQ's four scales was shown by the generally low intercorrelations, ranging from -.03 (for EI and PM) through .24 (for PM and OL). Age, gender, and racial/ethnic differences were examined in several studies to provide support to SSQ's construct validity. Results from these studies consistently indicated high agreement on SSQ factor structure and latent means of the four temperament traits across age, gender, and racial/ethnic groups (Strafford & Oakland, 1996).

Subsequent evaluation of interval validity was conducted using item parceling to overcome low total variance and item loadings when using dichotomous item scores. Factor loadings in the previous EFA became the basis to form 17 parcels, with four to five items per parcel, distributed across the four SSQ factors: five for EI, four for PM, two for TF, and six for OL. Six items were assigned to two parcels because of their significant loadings (≥ .20) on two hypothesized factors. Similar to the direct item EFA results, the hypothesized four-factor model was supported when using item parcels, with much larger total variance (53%) explained and higher factor loadings (≥ .44) of item parcels (Oakland, Glutting, & Horton, 1996).

Equivalence at the item level also was assessed through differential item functioning. Response patterns of Hispanic students were similar to those of African American and White students. However, items function somewhat differently on the OL trait between African American and White students (Stafford, 1994).

External validation of the SSQ was obtained through contrasted groups, convergent and divergent validation procedures. Using contrasted groups, students' temperament preferences were differentiated as a function of their career preferences, classes they liked and disliked, activities they most and least enjoyed, and involvement in special programs in school. Findings indicated that personal (e.g., interest) and contextual (e.g., school programs) qualities significantly influence students temperament style preferences (Oakland, Stafford, Horton, & Glutting, 2001).

Evidence of SSQ's convergent validity was examined by correlating students' temperament preferences and values. Correlations between temperament preferences and the values of helpfulness and loyalty were significant. For example, students who display a high preference for thinking style are more likely to value being helpful and students who display a high preference for a feeling style are more likely to value loyalty. Convergent validity also was examined by correlating scores from the SSQ and the Myers-Briggs Type Indicator (MBI). Expected results were evident. High correlations were found on the scales that measure similar qualities while low correlations were found on scales that measure different qualities (Oakland, Glutting, & Horton, 1996).

Divergent validity is achieved when a construct measured by a test is independent or not related to some hypothesized variables. Evidence of divergent validity of the SSQ is supported by showing that students' temperament preferences of students were not related to their achievement as measured by California Achievement Test (CAT: CTB/MacGraw-Hill, 1985) and to their and intelligence as measured by the Wechsler Intelligence Scale for Children-Revised (WISC-R: Wechsler, 1994). Obtained correlations between scores in the four SSQ traits and achievement and IQ scores generally were low and not significant. Thus, temperament constructs measured by the SSQ are independent of achievement and intelligence (Oakland, Glutting, & Horton, 1996).

The use of the SSQ in various countries has been increasing. SSQ adaptations have been successfully standardized in Romania, South Korea, and Taiwan. Content validity of the SSQ adapted versions was systematically guided by following the guidelines for test adaptations promulgated by the International Testing Commission (ITC: Hambleton, 2002).

Factorial structures of the SSQ scores of children from Australia, China, Costa Rica, Philippines, and Zimbabwe achieved excellent fit while factor structure of SSQ scores from Gaza and Nigeria achieved a modest fit (Benson, Oakland, & Shermis, 2009).

Procedure

Secondary data are used in this study and thus are exempt from formal review by the University Florida Institutional Review Board (UFIRB). Data gathering in each country was preceded by evaluating the suitability of the SSQ following the ITC guidelines for test adaptations. Researchers from all countries reviewed the SSQ's 69 items and direction and determined whether the items were suitable for use in their respective countries. Researchers from Australia, Nigeria, Pakistan, Singapore, and Zimbabwe deemed the SSQ English version appropriate for use because target participants in these countries are fluent in the English language and have at least a third-grade English reading level.

Researchers in countries where English was not the primary language formed advisory groups consisting of bilingual psychologists and educators who are fluent in both their native language and English to translate the SSQ following the back translation sequential method. The procedure essentially involves one or two bilingual speakers to translate the SSQ to the native language of the children. Then, different bilingual speakers back translated the SSQ into English. Discrepancies between the original English version and translated version of the SSQ were discussed and a consensus agreement made on the most developmentally and culturally appropriate translation. Countries in which the SSQ was translated into native languages include Brazil, China, Costa Rica, Egypt, Gaza, Hungary, Iran, Israel, Japan, Mongolia, Philippines, Poland, Romania, Samoa, and Venezuela. Table 2-2 reports the translation languages used and on-site researchers from 21 countries included in the study.

Participation of children was voluntary. Informed consent was obtained from the parents or guardians of the participating children by the researchers. Data gathering and control for extraneous testing factors were facilitated by request from the researchers to school heads to allow them to administer the SSQ to children in their respective classrooms. Children indicated their answers either using a separate response form or directly on the item booklet. Test administration was conducted either by the researchers or trained research assistants who followed administration procedures stated in the SSQ manual.

Data Analyses

Different statistical procedures were used to address the goals of the study. These procedures utilized either item scores or trait scores of the SSQ. The six SSQ items assigned to two bipolar traits were not included in the analysis to control the problem of cross-item variance. Thus, data from 63 items with unique factor solutions were included in the analysis.

Data Screening

Missing data in each country data set was handled in the following manner. First, children who did not answer six or more items from the SSQ were dropped from further analysis. Second, when calculating trait scores, children who missed to answer more than one item in any temperament trait will be treated as missing data in that specific trait (Schimitt, Allik, McCrae, & Benet-Martinez, 2007). Children with missing data will not be included in the analysis.

Internal Consistency of the SSQ

Preliminary analyses included examining internal consistency using Cronbach alpha and tetrachoric correlation of SSQ scores from each country. Cronbach alphas were calculated for each of the four temperament traits using dichotomous item scores and following the method used by Fan & Thomson (2001). A 95% confidence interval was established for each alpha coefficient to account for the band of values that accurately estimates internal consistency of dichotomous item scores.

Structural Validity

Structural equivalence of the hypothesized four factor structure was examined in U.S. children across age, gender, and racial/ethnic groups using EFA (Stafford, 1994; Stafford & Oakland, 1998) and across countries using CFA (Benson, Oakland, & Shermis, 2009). Both studies used item parcels instead of individual items as observed or manifest variables to remedy the problem of restricted variance of dichotomous items. Each parcel contained four to five items that measure the same trait. In forming the parcels, items with high and low *p* values were grouped together to counterbalance the difficulty levels within each parcel (Oakland, Glutting, & Horton, 1996).

Item parceling has been recommended as a method for lengthy tests with categorical or ordinal items when subjected to structural equation modeling (Yang, Nay, & Hoyle, 2010). Use of item parcels in factor analysis also has empirical advantages that include increasing reliability, meeting normality assumptions, resolving concerns on small sample sizes, simplifying interpretations, and achieving better model fit (Bandalos & Finney, 2001).

Item parceling is not without criticisms. Empirical arguments for the use of item parcels in psychometric analysis are associated with the problem of dimensionality of the construct, the misspecification of factor model, and the inaccurate meaning of parameter estimates (Little, Cunningham, Shahar, & Widaman, 2002). Parceling is problematic because items in any parcels may represent different meanings other than the hypothesized construct. Because of the multidimensionality of the item parcels, measurement models may have biased loading estimates that may have serious implications in explaining the variance of the latent construct (Badalos & Finney, 2001). Misspecification of factor model occurs when there is no information of the manifest variable (i.e., item) that underlie a latent construct. Hence, the measurement model is misspecified on account of the variance explained by individual items that directly describe the construct being measured.

The factor structure of the SSQ was highly similar across age, gender, and racial/ethnic groups (Stafford, 1994; Stafford & Oakland, 1998). Likewise, independent CFA of the SSQ data from seven countries replicated the four-factor structure of the SSQ. However, partial invariance was achieved when the data were examined for multigroup CFA (Benson, Oakland, & Shermis, 2009).

In the current study, the categorical nature of the SSQ item options was considered when determining the factor structure of temperament traits cross-nationally. Using Mplus 6.1 (Muthén & Muthén, 2006), a tetrachoric correlation matrix of item data for each country was fitted for CFA using a mean and variance-adjusted weighted least-squares estimator (WLSMV) with theta parameterization. The WLSMV is an advanced estimator for non-normal data when multivariate normality is not assumed. In addition, the WLSMV is a highly efficient estimator for non-normal data compared to other asymptotically distribution-free (ADF) estimators (Muthén & Muthén, 2006).

When using WLSMV, the root mean square error approximation (RMSEA) is used as the primary index of model fit, with ≤ .06 indicating good fit (Loehlin, 1998; Yu & Muthen, 2002). Comparative fit index (CFI) and Tucker-Lewis Index (TLI), commonly used model fit indexes, do not perform well in WLSMV and thus, were considered secondary criteria (Ivanova et al., 2007), in this study. Hu and Bentler (1999) suggested CFI and TLI values of ≥ .95 as criteria for a good model fit. However, alternative criteria of ≥ .90 for a good fit and .80 to .89 for an acceptable fit were adopted for use in this study (Brown & Cudeck, 1993; Marsh & Hau, 1996).

Sequential multigroup CFA was implemented starting with a single-group analysis in each country followed by a multiple-group analysis. A CFA for each country was conducted to examine the factor structure of temperament as measured by the SSQ. Only the countries with good model fit (i.e., RMSEA ≤ .06; CFI and TLI ≥ .90) were included in the multigroup CFA. Multigroup analysis was conducted to assess factorial invariance between a target country (e.g., Australia) and a reference country (U.S.). The U.S. is chosen as the reference country because the SSQ was developed in the U.S. and a stable SSQ four-factor structure has been established using data from more than 7,000 U.S. children.

The test of factorial invariance is a two-group CFA (i.e., a target country and the U.S. as the reference country) that examines a four factor structure of temperament in terms of the configural, metric, and scalar invariance (Dimitrov, 2010). A configural invariance is tested by fitting a baseline model with the same patterns of zero and non-zero loadings for a target country and the U.S. A metric invariance is tested by constraining the factor loadings to be equal between a target country and the U.S. A scalar invariance is tested by constraining both the factor loadings and the item intercepts to be equal for a target country and the U.S. (Meredith, 1993; Muthén & Muthén, 2006). Metric and scalar invariance models are tested at the same step because when using categorical data, factors loadings and item intercepts must be constrained in tandem (Muthén & Muthén, 2006).

The multigroup CFA was conducted following important assumptions. First, each item would be associated with only the factor it was intended to measure. Second, the covariation of the four factors of temperament would be allowed. Third, the post hoc model fitting would be set to a minimum and correlated error terms would be allowed when supported by strong empirical evidence. Fourth, the group-specific error covariance would be left unconstrained for a target country and the U.S. through the factorial invariance testing procedure (Byrne, 1994; Byrne, Shavelson, & Muthen, 1989).

Factorial invariance is determined through the statistical significance of the difference in the chi-square ($\Delta\chi^2$) values and CFI between the two nested models. A partial invariance is conducted if the difference test result is significant based on the modification indices (MI > 10) information (Cheung & Rensvold, 2002; Vandenberg & Lance, 2000).

Table 2-1. Sample sizes of countries by total, gender, and age group

Country	Total	Gender		Age Group			
		Male	Female	8-10	11-12	13-14	15-17
Australia	308	157	151	78	78	74	78
Brazil	461	212	249	124	139	114	84
China	400	200	200	100	100	100	100
Costa Rica	432	212	220	108	108	108	108
Egypt	800	399	401	178	197	213	212
Gaza	400	200	200	100	100	100	100
Hungary	401	201	200	99	100	102	100
Iran	511	237	274	133	161	140	77
Israel	253	114	139	87	166	—	—
Japan	493	223	270	149	141	93	112
Mongolia	1009	502	507	251	245	252	261
Nigeria	400	200	200	100	100	100	100
Pakistan	463	232	231	126	118	111	108
Philippines	400	200	200	100	100	100	100
Poland	440	202	238	44	155	73	168
Romania	900	450	450	225	225	225	225
Samoa	400	200	200	100	100	100	100
Singapore	483	255	228	193	139	90	61
United States	7,902	3,950	3,952	2,111	2,293	2,083	1,415
Venezuela	411	203	208	101	100	102	108
Zimbabwe	600	300	300	150	150	150	150
Total	17,867	8,849	9,018	4,657	5,015	4,430	3,767

Table 2-2. SSQ translation languages, names and institutional affiliations of on-site researchers

Country	Translation language	Lead on-site researcher	Institutional Affiliation
Australia	-	Michael Faulkner, Ph.D.	La Trobe University
Brazil	Portuguese	Raquel Guzzo, Ph.D.	Pontifucal Catholic University of Campinas
China	Mandarin	Li Lu, Ph.D	Shanxi Medical University
Costa Rica	Spanish	Ana Mata, M.A.	
Egypt	Arabic	Mahmoud Emam, Ph.D.	Assiut University
Gaza	Arabic	Mohammed Adnan Alghorani, Ph.D.	United Arab Emirates University
Hungary	Hungarian	Nora Katona	Eotvos Lorand University
Iran	Arabic	Mohammed Adnan Alghorani, Ph.D.	United Arab Emirates University
Israel	Hebrew	Sharone Maital, Ph.D.	University of Haifa
Japan	Nihonggo	Moisés Kirk de Carvalho Filho, Ph.D.	Kyoto University
Mongolia	Mongolian	Seded Bathkuyag, Ph.D.	National University of Mongolia
Nigeria	-	Andrew Mogaji, Ph.D.	University of Lagos
Pakistan	-	Muhammad Rizwan, Ph.D.	University of Karachi
Philippines	Tagalog	Carmelo Callueng, M.S.	De La Salle University
Poland	Polish	Tomasz Rowinski, Ph.D.	Cardinal Stefan Wyszynski University
Romania	Romanian	Dragos Iliescu, Ph.D.	SNSPA/D & D Research
Samoa	Samoan	Desmond Lee Hang, Ph.D.	National University of Samoa
Singapore	-	Yoke Fong Lau, M.Ed	University of Florida
Venezuela	Spanish	Carmen Leon, Ph.D.	Universidad Catolica Andres Bello
Zimbabwe	English	Elias Mpofu, Ph.D.	Pennsylvania State University

Note: *Countries with dash (-) used English language version of the SSQ*

CHAPTER 3
RESULTS

Data Screening

Before subjecting the data to CFA, data screening was conducted for each country data set to identify cases with at least 10% (6 items) of missing responses. Of the 21 countries included in the study, nine (43%) did not have a single case with substantial missing responses. The remaining 12 (57%) countries have missing data that range from one to 35 cases. In general, the occurrence of missing data (i.e., ≥ 6 items) was very minimal. Table 3-1 reports the number of cases per country with missing data.

Internal Consistency

As reported in Table 3-2, Cronbach alpha was calculated as a measure of internal consistency of item responses for each temperament trait per country. Using Cicchetti and Sparrow's (1990) guidelines for evaluating Cronbach alpha coefficient, EI data in 16 countries display acceptable (i.e., $\alpha \geq .60$) estimates of internal consistency. PM data in all the 21 countries display unacceptable (i.e., $\alpha \leq .60$) estimates of internal consistency. TF data in 2 countries display acceptable level of internal consistency. OL data in 16 countries display acceptable levels of internal consistency. Data from Egypt, Mongolia, and Samoa display unacceptable estimates of internal consistency across the four temperament factors.

Structural Validity

The fit of the four-factor model of temperament as measured by the SSQ was examined through the use of confirmatory factor analysis (CFA) for each country. Tetrachoric correlation matrix was fitted to the model, and the parameters were estimated using the mean and variance-adjusted weighted least-squares (WLSMV) estimator with theta parameterization. An initial CFA solution used the default paths of Mplus 6.1 with the first factor loading of each set of parameters automatically constrained to 1.0. However, using this default version resulted to non-convergence solution for data sets from Brazil, China, Costa Rica, Philippines, and Samoa. As an alternative, the variance of the *latent* factors was fixed to 1.0 for all country data sets. As mentioned in the previous chapter, the overall goodness-of-fit was evaluated using multiple criteria: 1) a non-significant WLSMV χ^2, 2) RMSEA $\leq .06$ as the primary model fit index (Loehlin, 1998; Yu & Muthen, 2002), and 3) CFI $\geq .90$ and TLI $\geq .90$ as secondary model fit indices (Brown & Cudeck, 1993; Marsh & Hau, 1996).

As reported in Table 3-3, CFA results were consistent for all the countries. That is, χ^2 values were significant, RMSEA values were less than .06, and both the CFI and TLI values were less than .90. RMSEA values indicate that each country data set has a good fit to the four-factor model of temperament. However, both the CFI and TLI indicate that the model did not fit the data for each of the 21 countries. The statistically significant χ^2 values are due to large sample size. Most item indicators converged on the traits they intended to measure. However, factor loadings generally were low. In most countries, the bipolar traits are distinct on what they measure. Taken together, the results of the model fit indices indicate that the four-factor model of temperament did not fit the data reasonably well for each of the 21 countries. Hence, the hypothesis that the four-factor model of temperament is not confirmed. The lack of model fit precludes tests of cross-national invariance of the four-factor model of temperament.

CFA for China, Egypt, and Gaza reported a non-positive definite matrix that was likely the result of an observed linear dependency of factors. As a result, a series of exploratory factor analysis (EFA) was performed for these three countries to determine an alternative factor solution.

Some items have significant loadings on factors other than what they intend to measure. These misspecified items were more frequent in EI (e.g., items 40 and 59) and OL (e.g., items 1 and 62). To address item misspecifications and consequently improve model fit, a modified four-factor solution was performed for each of the 18 countries with positive definite matrix. This was done by allowing some items to cross load on other factors. A modified solution was not done with the Samoa data because the misspecified items were not logically and substantively related to other factors that they seem to load. The modified solution did not significantly improve model fit in all the 18 countries. Similar to the initial CFA solution, the WLSMVχ^2 values were statistically significant, RMSEA values were within acceptable range indicating good model fit, and the CFI and TLI were consistently below .90 indicating non-fit to the model. A report on the structural validity of the SSQ is presented below for each country. Included in the report is a summary of factor loadings, standard error, and threshold of item indicators.

Australia (N = 368). Australian children prefer extroverted more than introverted, imaginative more than practical, thinking more than feeling and flexible more than organized styles.

The initial CFA solution yielded a significant $\chi^2_{(1884)} = 2606.560$, $p < .001$; RMSEA = .032; CFI = .733; and TLI =.723. As a whole, these fit statistics suggest that the data from Australian children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the squared standardized factor loadings of item indicators revealed that 53 (17 EI, 9 PM, 7 TF, and 20 OL items) out of 63 item indicators have appreciable loadings (i.e., ≥ .30) to their associated factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 10 items (2 EI, 5 PM, 1 TF, and 2 OL items) did not have substantial factor loadings (i.e., < .30) to their associated factors and conversely have high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated factor.

Factor correlations range from low (r = -.074 for EI and TF) to moderate (r = -.508 for PM and OL), with most correlations low yet significant. Factor correlations for the Australian data indicate lack of multicollinearity (i.e., r ≥ .85). Although most factors are related (e.g., PM and OL), they are not overlapping and thus are distinct in terms of the construct they measure.

A review of the initial CFA solution indicated significant misspecifications of factorial structure for nine items: 1, 11, 21, 28, 40, 51, 52, 55, and 62. Allowing some of these items to cross load on other factors would likely improve the model fit statistics. After evaluating the content of these items, three items appeared to have substantive and logical meaning to factors other than what these items intend to measure. Consequently, items 28 and 40 were allowed to cross load on EI factor and item 1 was allowed to cross load on OL factor. The resulting model fit statistics of the modified CFA solution for the Australian data yielded a significant $\chi^2_{(1881)} = 2548.690$, $p < .001$; RMSEA = .031; CFI = .753; and TLI =.744. These fit statistics suggest that the data from Australian children did not exhibit a good fit to the four-factor structure of temperament, albeit with modifications. Table 3-4 reports the parameter estimates of the four-factor solution in the Australia data.

Brazil (N = 461). Brazilian children prefer extroverted more than introverted, imaginative more than practical, and organized more than flexible styles. Approximately equal proportion of Brazilian children prefers thinking and feeling styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2990.948$, $p < .001$; RMSEA = .033; CFI = .639; and TLI =.626. As a whole, these fit statistics suggest that the data from Brazilian children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the squared standardized factor loadings of item indicators revealed that 50 items (18 EI, 10 PM, 8 TF, and 14 OL items) have appreciable loadings (i.e., ≥ .30) to their associated factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measures. The remaining 13 items (1 EI, 4 PM, and 8 OL items) did not have substantial factor loadings (i.e., < .30) to their associated factors and conversely have very high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated latent factor.

Factor correlations generally are low, ranging from r = .059 for TF and OL to r = -.252 for PM and OL. This suggests that factor correlations for the Brazil data indicate an absence of multicollinearity (i.e., r ≥ .85). Thus, the factors are distinct in reference to the constructs they measure.

A review of the initial CFA solution indicated significant misspecifications of factorial structure for items 1, 15, 16, 23, 33, 39, 47, 48, 50, 57, 59, 62, 64, 65, and 66. Allowing some of these items to cross load on other factors would likely improve the model fit statistics. After evaluating the content of these items, three items appeared to have substantive and logical meaning to factors other than what these items intend to measure. Consequently, item 59 was allowed to cross load on EI factor and items 1 and 21 were allowed to cross load on OL factor. The resulting model fit statistics of the modified CFA solution yielded a significant $\chi^2(1881) = 2893.080$, $p < .001$; RMSEA = .033; CFI = .67; and TLI =.658. These fit statistics suggest that the data from Brazilian children did not exhibit a good fit to the four-factor structure of temperament, albeit with modifications. Table 3-5 reports the parameter estimates of the four-factor solution for the Brazil data.

China (N = 399). Chinese children prefer practical more than imaginative, thinking more than feeling, and organized more than flexible styles. Approximately equal proportion of Chinese children prefers extroverted and introverted styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2754.924$, $p < .001$; RMSEA = .034; CFI = .741; and TLI =.731. As a whole, these fit statistics suggest that the data from Chinese children did not exhibit a good fit to the four-factor structure of temperament. Moreover, the initial solution indicates a non-positive definite correlation matrix that may be due to very high correlation coefficients that indicate substantial overlap of some factors. Specifically, correlation coefficients are 1.060 for PM and TF and -1.050 for TF and OL. All other inter-factor correlation coefficients range from r = -.233 (EI and OL) to r = -.789 (PM and OL). Although sample Pearson product moment correlation coefficients cannot be more extreme than ±1.0 in CFA sample inter-factor correlation coefficients can be more extreme than ±1.0 Such estimates are called inadmissible or improper and indicate a potential problem with the model even when goodness of fit indices are adequate. Because the CFA resulted to an improper solution, an EFA was conducted to determine an alternative model that may be salient and meaningful to the data in Chinese children.

Among the EFA solutions for two to five factors, the four-factor solution seemed to provide a more logical and meaningful convergence of items despite a poor fit to the model: $\chi^2(1767) = 2288.786$, $p < .001$; RMSEA = .027; CFI = .845; and TLI =.828. Factor 1 has 23 items with substantial factor loadings (i.e., ≥ .30), with 18 items (78%) measuring OL trait and the remaining 5 items (28%) measuring EI, PM, or TF trait. Factor 2 has 17 items with substantial factor loadings (i.e., ≥ .30), with 10 items (59%) measuring PM trait and the remaining 7 items (41%) measuring EI, TF, or OL trait. Factor 3 has 18 items with substantial factor loadings (i.e., ≥ .30), with 16 items (89%) measuring EI trait and the remaining 2 items (11%) measuring TF or OL trait. Seven items (i.e., 17, 23, 25, 32, 62, 65, & 66) loaded on more than one factor. Inter-factor correlation are low ranging from r = .072 to r = -.187. In summary, the three-factor solution represents a stable structure for items that measure OL, PM, and EI traits. Items that measure TF trait did not cluster together to form a stable and independent factor. Table 3-6 reports the EFA model fit indices and number of items with significant loadings for the different factor solutions in the China data.

Costa Rica (N = 431). Costa Rica children prefer extroverted more than introverted, practical more than imaginative, feeling more than thinking, and organized more than flexible styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 3164.511$, $p < .001$; RMSEA = .026; CFI = .590; and TLI =.575. As a whole, these fit statistics suggest that the data from Costa Rica children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the squared standardized factor loadings of item indicators revealed that 39 items (14 EI, 6 PM, 6 TF, and 13 OL items) have appreciable loadings (i.e., ≥ .30) with their associated factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 24 items (8 EI, 8 PM, 2 TF and 9 OL items) did not have substantial factor loadings (i.e., < .30) to their intended factors and conversely displayed with very high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated factor.

Most factor correlations are significant and range from very low (r = -.077 for PM and TF) to moderate (r = -.439 for PM and OL). This suggests that factor correlations for the Costa Rica data indicate an absence of multicollinearity (i.e., r ≥ .85) and these factors are distinct on the constructs they measure.

A review of the initial CFA solution indicated significant misspecifications of factorial structure for several items. Allowing some of these items to cross load on other latent factors would likely improve the model fit statistics. After evaluating the content of these items, four items seemed to have substantive and logical meaning to other later factors. Consequently, items 18 and 23 were allowed to cross load on EI factor and items 36 and 62 were allowed to cross load on OL factor. The resulting model fit statistics of this modified CFA solution yielded a significant $\chi^2(1880) = 3088.953$, $p < .001$; RMSEA = .025; CFI = .613; and TLI =.598. These fit statistics suggest that the data from Costa Rica children did not exhibit a good fit to the four-factor structure of temperament, albeit with modifications. Table 3-7 reports the parameter estimates of the four-factor solution in the Costa Rica data.

Egypt (N = 920). Egyptian children prefer extroverted more than introverted, practical more than imaginative, feeling more than thinking, and organized more than flexible styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 3146.025$, $p < .001$; RMSEA = .027; CFI = .513; and TLI =.495. As a whole, these fit statistics suggest that the data from Egyptian children did not exhibit a good fit to the four-factor structure of temperament. Moreover, the initial solution indicates a non-positive definite correlation matrix that may be due to a very high correlation coefficient for TF and OL (r = -1.051). Also, PM highly correlates with TF (r = -.859) and OL (r = .795). These high correlation coefficients indicate substantial overlap on what PM, TF, and OL intended to measure. All other inter-factor correlation coefficients are moderate and range from r = .393 (EI and TF) to r = -.556 (EI and OL). Although sample Pearson product moment correlation coefficients cannot be more extreme than ±1.0 in CFA sample inter-factor correlation coefficients can be more extreme than ±1.0 Such estimates are called inadmissible or improper and indicate a potential problem with the model even when goodness of fit indices are adequate. Because the CFA resulted in an improper solution, an EFA was conducted to determine an alternative model that is salient and meaningful to the data.

Among the EFA solutions for two to five factors, no factor solution showed a discernible pattern of item clustering, and model fit indices for all factor solutions generally are not adequate. In the two-factor solution ($\chi^2(1828) = 2795.259$, $p < .001$; RMSEA = .024; CFI = .627; and TLI =.601), 16 items have substantial loadings (i.e., ≤

.30) on factor 1 with 10 or 63% of the items measuring OL trait and four or 25% of the items measuring PM trait. The remaining two items measure EI and TF traits. For factor 2, 10 items have substantial loadings, with half of the items measuring OL trait and the other items measuring EI, PM, and TF traits. Item 15 cross-loaded on factors 1 and 2.

Using a three-factor solution ($\chi^2(1767) = 2496.157$, $p < .001$; RMSEA = .021; CFI = .718; and TLI =.689), 18 items have substantial loadings (i.e., ≥ .30) on factor 1, with 11 or 61% of the items measuring OL, five or 28% of the items measuring PM, one item each measuring EI and TF. For factor 2, 10 items have substantial loadings, with half of the items measuring OL and the other items measuring EI, PM, or TF traits. For factor 3, all nine items that have substantial factor loadings measure EI trait. Two items (i.e., item # 42 and 57) cross-load on factors 1 and 3.

Using a four-factor solution ($\chi^2(1707) = 2235.302$, $p < .001$; RMSEA = .018; CFI = .796; and TLI =.767), 10 items have substantial loadings (i.e., ≥ .30) on factor 1, with more than half (60%) of the items measuring OL and the other items measuring PM or TF. For factor 2, 11 items have substantial loadings mostly measuring PM and OL traits. For factor 3, nine items have substantial factor loadings that measure EI, PM, TF, or OL traits. For factor 4, nine of the 10 items that have substantial factor loadings measure EI trait. Four items (i.e., item # 15, 32, 47 and 66) cross-load on more than one factor.

Using a five-factor solution ($\chi^2(1648) = 2062.220$, $p < .001$; RMSEA = .017; CFI = .840; and TLI =.810), 16 items have substantial loadings (i.e., ≥ .30) on factor 1, with half of the items measuring OL and the other items measuring EI, PM, or TF. For factor 2, three items have substantial loadings that measure PM or TF. For factor 3, 11 items have substantial factor loadings, with most items measuring OL and few items measuring EI, PM, or TF. For factor 4, six of the eight items with substantial factor loadings measure OL and the remaining two items measure PM. For factor 5, all eight items with substantial factor loadings measure EI. Nine items (i.e., item # 3, 15, 20, 31, 32, 42, 47, 57, and 66) cross-load on more than one factor. Table 3-8 reports the EFA model fit indices and number of items with significant loadings for the different factor solutions in the Egypt data.

Gaza (N = 400). Gaza children prefer extroverted more than introverted, practical more than imaginative, thinking more than feeling, and organized more than flexible styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2461.500$, $p < .001$; RMSEA = .029; CFI = .755; and TLI =.746. As a whole, these fit statistics suggest that the data from Gaza children did not exhibit a good fit to the four-factor structure of temperament. Moreover, the initial solution indicates a non-positive definite correlation matrix that may be due to very correlation coefficients and indicate substantial overlap of some factors. Most inter-factor correlation coefficients are evident of multicollinearity (i.e., $r \geq .85$). For example, TF highly correlates with EI ($r = -.904$), PM ($r = .890$), and OL ($r = -.938$). OL also highly correlates with EI ($r = -.852$) and PM ($r = -.806$). Correlation between EI and PM ($r = -.605$) is moderate. A very high correlation of factors in CFA indicates a potential problem with the model even when goodness of fit indices are adequate. Because the CFA resulted to an improper solution, an EFA was conducted to determine an alternative model that is salient and meaningful to the data in Gaza children.

EFA solutions for two to five factors were run and among these solutions, the two-factor solution seemed to provide a more logical and meaningful convergence of items despite a poor fit to the model: $\chi^2(1828) = 2152$, $p < .001$; RMSEA = .022; CFI = .862; and TLI =.853. Factor 1 has 33 items with substantial factor loadings (i.e., ≥ .30) that mostly measure OL (19 items or 58%) and PM (8 items or 24%). The remaining six items are equally divided to measure EI and TF. In factor 2, 13 out of 14 items have substantial factor loadings (i.e., ≥ .30) that measure EI and one item measuring PM. The correlation between factors 1 and 2 is low ($r = -.090$). In summary, the two-factor solution represents a stable structure for items that measure EI on one factor and a combination of items that measure PM and OL on the other factor. Very few TF items have significant loadings and did not cluster as an independent factor. Table 3-9 reports the EFA model fit indices and number of items with significant loadings for the different factor solutions in the Gaza data.

Hungary (N = 494). Hungarian children prefer extroverted more than introverted and organized more than flexible styles. Approximately equal proportions of Hungarian children prefer practical and imaginative as well as thinking and feeling styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2617.849$, $p < .001$; RMSEA = .031; CFI = .475; and TLI =.456. As a whole, these fit statistics suggest that the data from Hungarian children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the squared standardized factor loadings of item indicators revealed that 37 items (15 EI, 5 PM, 4 TF, and 12 OL items) have appreciable loadings (i.e., ≥ .30) to their associated factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 26 items (4 EI, 9 PM, 4 TF and 10 OL items) did not have substantial factor loadings (i.e., < .30) to their intended latent factors and conversely have high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated latent factor.

Factor correlations for EI and PM (r = .049), EI and TF (r = -.086), and EI and OL (r = -.157) are in the low range. Moreover, significant and moderate correlations are reported for PM and TF (r = .686) and TF and OL (r = .571). Multicollinearity (i.e., r ≥ .85) is not evident on these correlations. EI and TF do not overlap with PM and OL and are distinct in terms of the constructs they measure. On the other hand, the correlation between PM and OL (r = -.867) provides evidence of multicollinearity and suggests an overlap on what PM and OL measure.

A review of the initial CFA solution indicated significant misspecifications of factorial structure for items 22, 38, 62, and 67. Allowing some of these items to cross load in other latent factors may likely improve the model fit statistics. After evaluating the content of these items, three items seemed to have substantive and logical meaning to other later factors. Consequently, items 62 and 67 were allowed to cross load on OL and TF factors respectively. The resulting model fit statistics of this modified CFA solution for the Hungary data yielded a significant $\chi^2(1882) = 2592.397$, $p < .001$; RMSEA = .031; CFI = .492; and TLI =.472. These fit statistics suggest that the data from Hungarian children did not exhibit a good fit to the four-factor structure of temperament, albeit with modifications. Table 3-10 reports the parameter estimates of the four-factor solution in the Hungary data.

Iran (N = 510). Iranian children prefer practical more than imaginative and organized more than flexible styles. Approximately equal proportions of Iranian children prefer extroverted and introverted as well as thinking and feeling styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2878.989$, $p < .001$; RMSEA = .032; CFI = .722; and TLI =.712. As a whole, these fit statistics suggest that the data from Iranian children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the squared standardized factor loadings of item indicators revealed that 36 items (8 EI, 7 PM, 2 TF, and 19 OL items) have appreciable loadings (i.e., ≥ .30) to their intended factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 27 items (11 EI, 7 PM, 6 TF, and 3 OL items) did not have substantial factor loadings (i.e., < .30) to their intended factors and conversely have very high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated latent factor.

Factor correlations range from moderate (r = .508 for PM and TF) to high (r = -.879 for EI and PM). The high correlation between EI and PM indicates multicollinearity (i.e., r ≥ .85) and overlap on the construct they measure. All other correlations generally are in the moderate range and the correlations are distinct in the construct they measure.

A review of the initial CFA solution indicated a significant misspecification of factorial structure for item 27. When item 27 was allowed to cross load on PM factor, the resulting model fit statistics of the modified CFA solution yielded a significant $\chi^2(1883) = 2866.009$, $p < .001$; RMSEA = .032; CFI = .725; and TLI =.715. These fit statistics suggest that the data from Iranian children did not exhibit a good fit to the four-factor structure of temperament, albeit with modification. Table 3-11 reports the parameter estimates of the four-factor solution in the Iran data.

Israel (N = 218). Israeli children prefer extroverted more than introverted, practical more than imaginative, feeling more than thinking, and organized more than flexible styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2175.017$, $p < .001$; RMSEA = .027; CFI = .749; and TLI =.739. As a whole, these fit statistics suggest that the data from Israeli children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the squared standardized factor loadings of item indicators revealed that 44 items (15 EI, 8 PM, 7 TF, and 14 OL items) have appreciable loadings (i.e., ≥ .30) on their intended factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 19 items (4 EI, 6 PM, 1 TF, and 8 OL) did not have significant factor loadings (i.e., < .30) on their intended factors and conversely have high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated factor.

Moderate and significant correlations are reported for PM and TF ($r = -.386$) and for PM and OL ($r = -.432$). All other factor correlations are low and not significant. Multicollinearity (i.e., $r \geq .85$) is not evident in all correlations, suggesting that the factors are distinct and do not overlap in the construct they measure.

A review of the initial CFA solution indicated significant misspecifications of factorial structure for items 1, 50, and 68. Allowing these items to cross load on other latent factors would improve the model fit statistics. However, only items 1 and 50 seemed to have substantive and logical meaning associated with OL and EI, respectively. When item 1 was allowed to cross load on OL trait and item 50 was allowed to cross load on EI trait the resulting model fit statistics of the modified CFA solution for the Israel data yielded a significant $\chi^2(1882) = 2144.574$, $p < .001$; RMSEA = .025; CFI = .773; and TLI =.763. These fit statistics suggest that the data from Israeli children did not exhibit a good fit to the four-factor structure of temperament, albeit with modifications. Table 3-12 reports the parameter estimates of the four-factor solution in the Israel data.

Japan (N = 494). Japanese children prefer practical more than imaginative and organized more than flexible styles. Approximately equal proportions of Japanese children prefer extroverted and introverted as well as feeling and thinking styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2904.025$, $p < .001$; RMSEA = .033; CFI = .732; and TLI =.722. As a whole, these fit statistics suggest that the data from Japanese children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the squared standardized factor loadings of item indicators revealed that 50 items (18 EI, 8 PM, 7 TF, and 17 OL items) have appreciable loadings (i.e., ≥ .30) to their intended factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 13 items (1 EI, 6 PM, 1 TF, and 5 OL items) did not have significant factor loadings (i.e., < .30) to their intended factors and conversely have high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated factor.

Factor correlations are generally low and range from r = -.028 (EI and OL) to r = -.199 (TF and OL), with most correlations not significant. This implies that factor correlations for the Japanese data lack multicollinearity (i.e., $r \geq .85$) and the factors are distinct in terms of the construct they measure.

A review of the initial CFA solution indicated significant misspecifications of factorial structure for items 9, 11, 18, 22, 23, 28, 31, 34, 35, 40, 43, 45, 50, 55, 61, 62, 63, and 68. Allowing some of these items to cross load on other latent factors may improve the model fit statistics. After evaluating the content of these items, items 31, 50, and 63 appeared to have substantive and logical meaning to other factors other than the factors they are associated to measure. Consequently, item 31, 50, and 63 were allowed to cross load on OL, EI, and PM, respectively. The resulting model fit statistics of the modified CFA solution for the Japan data yielded a significant $\chi^2(1881)$ = 2861.529, $p < .001$; RMSEA = .033; CFI = .743; and TLI =.733. These fit statistics suggest that the data from Japanese children did not exhibit a good fit to the four-factor structure of temperament, albeit with modifications. Table 3-13 reports the parameter estimates of the four-factor solution in the Japan data

Mongolia (N = 1,003). Mongolian children prefer practical more than imaginative, thinking more than feeling, and organized more than flexible styles. Approximately equal proportion of Mongolian children extroverted and introverted styles.

The initial CFA solution yielded a significant $\chi^2(1884)$ = 3166.944, $p < .001$; RMSEA = .026; CFI = .589; and TLI =.574. As a whole, these fit statistics suggest that the data from Mongolian children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the standardized factor loadings of item indicators revealed that 39 (14 EI, 6 PM, 6 TF, and 13 OL) items have appreciable loadings (i.e., ≥ .30) on their associated factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 24 (5 EI, 8 PM, 2 TF and 9 OL) items did not have substantial factor loadings (i.e., < .30) on their intended factors and conversely display high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated factor.

Most factor correlations are significant and range from low (r = -.077 for PM and TF) to moderate (r = -.439 for PM and OL). This suggests that factor correlations for the Mongolia data indicate an absence of multicollinearity (i.e., r ≥ .85) and the factors are distinct on the constructs they measure.

A review of the initial CFA solution indicated significant misspecifications of factorial structure for items 1, 7, 11, 16, 18, 23, 31, 33, 35, 39, 40, 49, 50, 52, and 59. Allowing some of these items to cross load on other latent factors may improve the model fit statistics. After evaluating the content of these items, items 1, 31, 33, and 59 appeared to have substantive and logical meaning to other factors other than the factors they are associated to measure. Consequently, the aforementioned items were allowed to cross load: item 59 on EI, item 33 on TF, and items 1 and 31 on OL. The resulting model fit statistics of the modified CFA solution for the Mongolia data yielded a significant $\chi^2(1880) = 3078.819$, $p < .001$; RMSEA = .025; CFI = .616; and TLI =.601. These fit statistics suggest that the data from Mongolian children did not exhibit a good fit to the four-factor structure of temperament, albeit with modifications. Table 3-14 reports the parameter estimates of the four-factor solution in the Mongolia data.

Nigeria (N = 392). Response proportion of Nigerian children who endorsed specific temperament styles indicated their preferences for response options that measure practical more than imaginative and organized more than flexible styles. Approximately equal proportions of Nigeria children prefer extroverted and introverted as well as thinking and feeling styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2282.416$, $p < .001$; RMSEA = .023; CFI = .666; and TLI =.654. As a whole, these fit statistics suggest that the data from Nigeria children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the standardized factor loadings of item indicators revealed that 50 items (13 EI, 7 PM, 5 TF, and 15 OL items) have appreciable loadings (i.e., ≥ .30) to their associated factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 23 items (6 EI, 7 PM, 3 TF, and 7 OL items) did not have significant factor loadings (i.e., < .30) to their intended factors and conversely have high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated factor.

Most factor correlations are significant and range from low ($r = .027$ for EI and OL) to moderate ($r = -.635$ for PM and OL), with no evidence of multicollinearity (i.e., $r \geq .85$). Although most factors are related, they do not overlap and appear to be distinct in terms of the construct they measure.

A review of the initial CFA solution indicated a significant misspecification of factorial structure for items 16, 52, 62, and 67. Allowing these items to cross load on other factors may improve the model fit statistics. However, only items 52 and 62 appear to have substantive and logical meaning associated with OL and TF, respectively. When item 52 and 62 were allowed to cross load on OL and TF, the resulting model fit statistics of the modified CFA solution yielded a significant $\chi^2(1882)$ = 2249.523, $p < .001$; RMSEA = .022; CFI = .692; and TLI =.681. These fit statistics suggest that the data from Nigerian children did not exhibit a good fit to the four-factor structure of temperament, albeit with modifications. Table 3-15 reports the parameter estimates of the four-factor solution in the Nigeria data.

Pakistan (N = 458). Pakistani children prefer practical more than imaginative, feeling more than thinking, and organized more than flexible styles. Approximately equal proportion of Pakistan children prefers extroverted and introverted styles.

The initial CFA solution yielded a significant $\chi^2(1884)$ = 2644.864, $p < .001$; RMSEA = .030; CFI = .504; and TLI =.486. As a whole, these fit statistics suggest that the data from Pakistani children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the standardized factor loadings of item indicators revealed that 36 items (10 EI, 8 PM, 5 TF, and 13 OL items) have appreciable loadings (i.e., ≥ .30) to their associated factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 27 items (9 EI, 6 PM, 3 TF, and 9 OL items) did not have significant factor loadings (i.e., < .30) on their associated factors and conversely have high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated factor.

Except for a moderate and significant correlation between PM and OL ($r = -.681$), all other factor correlations are low and not significant. Multicollinearity (i.e., $r \geq .85$) was not evident, thus suggesting that the factors are distinct in the constructs they measure.

A review of the initial CFA solution indicated a significant misspecification of factorial structure for items 7, 11, 13, 18, 19, 27, 40, 49, 55, 57, 62, and 67. Allowing these items to cross load on other factors may improve the model fit statistics. However, only items 27 and 40 seemed to have substantive and logical meaning associated with PM and EI, respectively. When items 27 and 40 were allowed to cross load on PM and EI the model fit statistics of the modified CFA solution yielded a significant $\chi^2(1882) = 2619.891$, $p < .001$; RMSEA = .029; CFI = .519; and TLI =.501. These fit statistics suggest that the data from Pakistani children did not exhibit a good fit to the four-factor structure of temperament as measured by the SSQ, albeit with modifications. Table 3-16 reports the parameter estimates of the four-factor solution in the Pakistan data.

Philippines (N = 399). Filipino children prefer practical more than imaginative and organized more than flexible styles. Approximately equal proportions of Filipino children prefer extroverted and introverted as well as thinking and feeling styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2357.626$, $p < .001$; RMSEA = .025; CFI = .678; and TLI =.666. As a whole, these fit statistics suggest that the data from Filipino children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the standardized factor loadings of item indicators revealed that 36 items (12 EI, 6 PM, 3 TF, and 15 OL items) have appreciable loadings (i.e., $\geq .30$) on their associated factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 27 items (7 EI, 8 PM, 5 TF, and 7 OL items) did not have significant factor loadings (i.e., < .30) on their associated factors and conversely have high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated factor.

All factor correlations are significant and generally in the low range (i.e., r ≤ .300), except for a moderate correlation between PM and OL (r = -.500). Multicollinearity (i.e., r ≥ .85) was not evident, thus suggesting that the factors are distinct on the constructs they measure.

A review of the initial CFA solution indicated a significant misspecification of factorial structure for items 1, 4, 40, 48, and 67. Allowing these items to cross load on other latent factors may improve the model fit statistics. However, only items 1 and 40 seemed to have substantive and logical meaning associated with OL and EI, respectively. When items 1 and 40 were allowed to cross load on the OL and EI factors, respectively, the resulting model fit statistics of the modified CFA solution yielded a significant $\chi^2(1882) = 2300.994$, $p < .001$; RMSEA = .024; CFI = .715; and TLI =.704. These fit statistics suggest that the data from Filipino children did not exhibit a good fit to the four-factor structure of temperament, albeit with modifications. Table 3-17 reports the parameter estimates of the four-factor model in the Philippine data.

Poland (N = 434). Polish children prefer introverted more than extroverted, practical more than imaginative, and organized more than flexible styles. Approximately equal proportions of Polish children prefer thinking and feeling styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2792.991$, $p < .001$; RMSEA = .033; CFI = .620; and TLI =.606. As a whole, these fit statistics suggest that the data from Polish children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the standardized factor loadings of item indicators revealed that 44 items (13 EI, 9 PM, 6 TF, and 16 OL items) have appreciable loadings (i.e., ≥ .30) on their associated factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 19 items (6 EI, 5 PM, 2 TF, and 6 OL items) did not have significant factor loadings (i.e., < .30) on their associated factors and conversely have high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated factor.

Most factor correlations are significant and low (i.e., $r \leq .300$), except for a moderate correlation between PM and OL ($r = -.487$). Multicollinearity (i.e., $r \geq .85$) is not evident. Although the factors generally are related with one another, they reflect distinct constructs.

A review of the initial CFA solution indicated a significant misspecification of factorial structure for items 7, 8, 22, 28, 33, 40, 45, 50, 59, 62, 64, and 68. Allowing these items to cross load on other latent factors would likely improve the model fit statistics. However, only items 8, 28, 40, 50, and 59 seem to have substantive and logical meaning associated with EI. When those five items were allowed to cross load on EI, the resulting model fit statistics of the modified CFA solution yielded a significant $\chi^2(1879) = 2687.857$, $p < .001$; RMSEA = .031; CFI = .662; and TLI =.648. These fit statistics suggest that the data from Polish children did not exhibit a good fit to the four-factor structure of temperament as measured by the SSQ, albeit with modifications.

Table 3-18 reports the parameter estimates of the four-factor solution in the Poland data.

Romania (N = 391). Romanian children prefer extroverted more than introverted, practical more than imaginative, thinking more than feeling and organized more than flexible styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2650.835$, $p < .001$; RMSEA = .032; CFI = .654; and TLI =.641. As a whole, these fit statistics suggest that the data from Filipino children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the standardized factor loadings of item indicators revealed that 45 items (12 EI, 9 PM, 6 TF, and 18 OL items) have appreciable loadings (i.e., ≥ .30) on their associated factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 18 items (7 EI, 5 PM, 2 TF, and 4 OL items) did not have significant factor loadings (i.e., < .30) on their associated factors and conversely have high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated factor.

Most factor correlations are significant and low (i.e., r ≤ .300), except for a moderate correlation between PM and OL (r = -.634). Multicollinearity (i.e., r ≥ .85) is not evident. Although the factors generally are related with one another, they are distinct in the constructs they measure.

A review of the initial CFA solution indicated a significant misspecification of factorial structure for items 1, 16, 33, 40, 43, 52, 62, and 67. Allowing these items to cross load on other latent factors may improve the model fit statistics. However, only items 1 and 40 seemed to have substantive and logical meaning associated with OL and EI factors, respectively. When these items were allowed to cross load, the resulting model fit statistics of the modified CFA solution yielded a significant $\chi^2(1882) = 2595.766$, $p < .001$; RMSEA = .031; CFI = .678; and TLI =.666. These fit statistics suggest that the data from Romanian children did not exhibit a good fit to the four-factor structure of temperament, albeit with modifications. Table 3-19 reports the parameter estimates of the four-factor solution in the Romania data.

Samoa (N = 400). Samoan children prefer extroverted more than introverted, practical more than imaginative, thinking more than feeling and organized more than flexible styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2501.372$, $p < .001$; RMSEA = .029; CFI = .620; and TLI =.606. As a whole, these fit statistics suggest that the data from Samoa children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the standardized factor loadings of item indicators revealed that 39 items (13 EI, 8 PM, 4 TF, and 14 OL items) have appreciable loadings (i.e., ≥ .30) on their associated factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 24 items (6 EI, 6 PM, 4 TF, and 8 OL items) did not have significant factor loadings (i.e., < .30) on their associated factors and conversely have high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated factor.

All factor correlations are high, ranging from r = .713 (EI and TF) to r = -.959 (TF and OL). Multicollinearity (i.e., r ≥ .85) is evident for EI and OL (r = -.863), PM and OL (r

= -.857), and TF and OL (r = -.959). These correlations imply that the factors overlap and may share similar meanings on what they measure.

A review of the initial CFA solution indicated a significant misspecification of factorial structure for item 52; that is, allowing item 52 to cross load on TF and OL traits may improve the model fit statistics. However, the content meaning of item 52 is not substantively and logically related to TF and OL. Hence, there was no compelling reason to pursue a modification of the CFA solution. Table 3-20 reports the parameter estimates of the four-factor solution in the Samoa data.

Singapore (N = 483). Singaporean children prefer practical more than imaginative, feeling more than thinking and organized more than flexible styles. Approximately equal proportions of Singaporean children prefer extroverted and introverted styles.

Initial CFA solution yielded a significant $\chi^2(1884) = 2463.218$, $p < .001$; RMSEA = .025; CFI = .732; and TLI =.725. As a whole, these fit statistics suggest that the data from Singaporean children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the standardized factor loadings of item indicators revealed that 49 (16 EI, 8 PM, 6 TF, and 19 OL) items have appreciable loadings (i.e., ≥ .30) on their associated factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 14 (3 EI, 6 PM, 2 TF, and 3 OL) items did not have significant factor loadings (i.e., < .30) on their associated factors and conversely have high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated factor.

Most factor correlations are not significant and in the low range (i.e., ≤ .200). As multicollinearity (i.e., r ≥ .85) was not evident, the factors seemingly measure distinct constructs.

A review of the initial CFA solution indicated a significant misspecification of factorial structure for items 17, 25, 43, 52, 59, 68, and 69. Allowing these items to cross load on other factors may improve the model fit statistics. Among the misspecified items, only item 59 seemed to have substantive and logical meaning associated with EI. When item 59 was allowed to cross load on EI, the resulting model fit statistics of the modified CFA solution yielded a significant $\chi^2(1883) = 2432.524$, $p < .001$; RMSEA = .025; CFI = .748; and TLI =.739. These fit statistics suggest that the data from Singapore children did not exhibit a good fit to the four-factor structure of temperament, albeit with modifications. Table 3-21 reports the parameter estimates of the four-factor solution in the Singapore data.

U.S. (N = 7,902). Children from the U.S prefer imaginative more than practical and organized more than and flexible styles. Approximately equal proportions of U.S. children prefer extroverted and introverted as well as thinking and feeling styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2488.575$, $p < .001$; RMSEA = .025; CFI = .759; and TLI =.750. As a whole, these fit statistics suggest that the data from U.S. children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the standardized factor loadings of item indicators revealed that 54 items (14 EI, 11 PM, 7 TF, and 22 OL items), including all OL have appreciable loadings (i.e., ≥ .30) on their associated factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 8 items (5 EI, 3 PM, and 1 TF items) did not have significant factor loadings (i.e., < .30) on their associated factors and conversely have high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated factor.

Factor correlations generally are low, ranging from r = .048 for EI and PM to r = -.396 for PM and OL. As multicollinearity (i.e., r ≥ .85) was not evident, the factors seemingly measure distinct.

A review of the initial CFA solution indicated a significant misspecification of factorial structure for items 1, 16, 26, 42, 55, 62, and 65. Allowing these items to cross load on other latent factors may improve the model fit statistics. After evaluating the content of these items, items 1, 16, 26, 52, and 62 appeared to have substantive and logical meaning to factors other than those they are intended to measure. Consequently, the aforementioned items were allowed to cross load: items 16, 26, and 52 on TF, and items 1 and 62 on OL. The resulting model fit statistics of the modified CFA solution yielded a significant $\chi^2(1880) = 2407.145$, $p < .001$; RMSEA = .024; CFI = .789; and TLI =.781. These fit statistics suggest that the data from U.S. children did not exhibit a good fit to the four-factor structure of temperament, albeit with modifications. Table 3-22 reports the parameter estimates of the four-factor solution in the U.S. data.

Venezuela (N = 411). Venezuelan children prefer extroverted more than introverted, practical more than imaginative, and thinking more than feeling styles. Approximately equal proportion of Venezuelan children prefer organized and flexible styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2586.444$, $p < .001$; RMSEA = .030 ; CFI = .648; and TLI =.635. These fit statistics suggest that the data from Venezuelan children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the standardized factor loadings of item indicators revealed that 40 items (15 EI, 4 PM, 5 TF, and 16 OL items) have appreciable loadings (i.e., ≥ .30) on their associated factors. Each of these items contributed a significant amount of variance that can be explained by the factor it is intended to measure. The remaining 23 items (4 EI, 10 PM, 3 TF, and 6 OL items) did not have significant factor loadings (i.e., < .30) on their associated factors and conversely have high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated factor.

Most factor correlations are in the low range (i.e., $r \le .262$), except for a high correlation between PM and OL ($r = -.757$). As multicollinearity (i.e., $r \ge .85$) was not evident, the factors seemingly measure distinct constructs.

A review of the initial CFA solution indicated a significant misspecification of factorial structure for items 1, 7, 18, 39, 59, 62, 67, and 68. Allowing these items to cross load on other latent factors improve the model fit statistics. Among the misspecified items, only items 1 and 59 seemed to have substantive and logical meaning associated with OL and EI, respectively. When items 1 and 59 were allowed to cross load on OL and EI factors, the resulting model fit statistics of the modified CFA solution yielded a significant $\chi^2(1882) = 2517.479$, $p < .001$; RMSEA = .029; CFI = .682; and TLI =.670. These fit statistics suggest that the data from Venezuelan children did not exhibit a good fit to the four-factor structure of temperament, albeit with modification.

Table 3-23 reports the parameter estimates of the four-factor solution in the Venezuela data.

Zimbabwe (N = 490). Zimbabwe children prefer extroverted more than introverted, practical more than imaginative, feeling more than thinking, and organized more than flexible styles.

The initial CFA solution yielded a significant $\chi^2(1884) = 2300.126$, $p < .001$, RMSEA = .021 , CFI = .664, and TLI =.652. As a whole, these fit statistics suggest that the data from Zimbabwe children did not exhibit a good fit to the four-factor structure of temperament. An inspection of the standardized factor loadings of item indicators revealed that only 30 (4 EI, 7 PM, 3 TF, and 17 OL) out of 63 items have appreciable loadings (i.e., ≥ .30) on their associated factors. Each of these 31 items contributed a significant amount of variance that can be explained by the factor it measures. The remaining 33 items (15 EI, 8 PM, 5 TF, and 5 OL items) did not have significant factor loadings (i.e., < .30) on their associated factors and conversely have high residual variances. Each of these items accounted for a significant proportion of variance that is not explained by its associated latent factor.

Most factor correlations are significant and range from low (r = .183 for EI and TF) to high (r = -.763 for PM and OL). As multicollinearity (i.e., r ≥ .85) is not evident, factors seemingly measure distinct constructs.

A review of the initial CFA solution indicated a significant misspecification of factorial structure for items 7, 12, 26, 46, and 52. Allowing these items to cross load on other factors improve the model fit statistics. Among the misspecified items, only item 52 seemed to have substantive and logical meaning associated with OL trait. When item 52 was allowed to cross load on OL trait, the resulting model fit statistics of the modified CFA solution yielded a significant $\chi^2(1883) = 2281.020$, $p < .001$; RMSEA = .021; CFI = .679; and TLI =.667. These fit statistics suggest that the data from Zimbabwe children did not exhibit a good fit to the four-factor structure of temperament, albeit with modification. Table 3-24 reports the parameter estimates of the four-factor solution in the Zimbabwe data.

Table 3-1. Participants per country with at least 6 item responses missing

Country	Total Sample	Number of Participants with ≥ 6 Missing Item Responses	Actual Sample Size
Australia	369	1	368
Brazil	461	0	461
China	400	1	399
Costa Rica	432	1	431
Egypt	954	34	920
Gaza	400	0	400
Hungary	401	0	400
Iran	511	1	510
Israel	253	35	218
Japan	493	0	493
Mongolia	1009	6	1003
Nigeria	400	8	392
Pakistan	463	5	458
Philippines	400	1	399
Poland	440	6	434
Romania	391	0	391
Samoa	400	0	400
Singapore	483	0	483
United States	7,902	0	7,902
Venezuela	411	0	411
Zimbabwe	492	2	490

Table 3-2. Cronbach alpha coefficients of the temperament traits in 21 countries

Country	EI (19 items)		PM (14 items)		TF (8 items)		OL (22 items)	
	α	95% CI	α	95% CI	α	95 %CI	α	95% CI
Australia	.70	.66-.75	.56	.49-.63	.64	.58-.70	.80	.77-.83
Brazil	.72	.68-.76	.52	.45-.58	.55	.48-.61	.46	.39-.53
China	.73	.69-.77	.55	.48-.61	.39	.30-.48	.73	.69-.76
Costa Rica	.71	.66-.75	.36	.26-.44	.41	.32-.49	.75	.71-.78
Egypt	.49	.44-.54	.33	.26-.39	.29	.22-.36	.34	.28-.40
Gaza	.61	.55-.66	.41	.32-.50	.38	.28-.47	.72	.69-.76
Hungary	.65	.60-.70	.48	.40-.55	.30	.20-.40	.55	.48-.61
Iran	.60	.55-.65	.31	.22-.39	.12	.01-.24	.76	.73-.79
Israel	.75	.68-.80	.56	.46-66	.62	.53-.69	.65	.58-.72
Japan	.78	.75-.81	.51	.44-.57	.57	.51-.62	.73	.69-.76
Mongolia	.54	.50-.59	.23	.16-.30	.41	.36-.47	.43	.38-.49
Nigeria	.62	.56-.67	.47	.39-.55	.39	.30-.48	.66	.60-.70
Pakistan	.56	.49-.62	.48	.41-.55	.39	.30-.47	.60	.54-.65
Philippines	.62	.57-.67	.42	.33-.50	.34	.24-.44	.61	.57-.67
Poland	.62	.57-.67	.54	.47-.60	.58	.52-.64	.71	.67-.75
Romania	.67	.62-.72	.49	.40-.55	.41	.32-.50	.76	.73-.80
Samoa	.56	.50-.62	.32	.22-.41	.21	.09-.32	.35	.25-.43
Singapore	.73	.69-.76	.49	.42-.55	.49	.42-.56	.69	.65-.73
United States	.69	.68-.70	.55	.54-.57	.35	.32-.37	.68	.67-.69
Venezuela	.62	.57-.68	.31	.21-.41	.33	.22-.42	.66	.61-.71
Zimbabwe	.51	.45-.57	.25	.15-.35	.20	.09-.31	.60	.54-.65
Mean α	.64		.44		.40		.63	

Table 3-3. Overall goodness-of-fit indices for individual country CFA

Country	WLSMVχ^2	*df*	RMSEA (90% C.I.)	CFI	TLI
Australia					
Initial solution	2606.560	1884	.032 (.029 - .035)	.733	.723
Modified solution	2548.690	1881	.031 (.028 - .034)	.753	.734
Brazil					
Initial solution	2990.948	1884	.033 (.031 - .035)	.639	.626
Modified solution	2893.080	1881	.033 (.031 - .035)	.670	.658
China	2754.924	1884	.034 (.031 - .037)	.741	.731
Initial solution					
Costa Rica					
Initial solution	3164.511	1884	.026 (.023 - .029)	.590	.575
Modified solution	3088.953	1880	.025 (.022 - .028)	.613	.598
Egypt					
Initial solution	3146.025	1884	.027 (.025 -.029)	.513	.495
Gaza					
Initial solution	2461.500	1884	.029 (.026 - .032)	.755	.746
Hungary					
Initial solution	2617.849	1884	.031 (.028 - .034)	.475	.456
Modified solution	2592.397	1882	.031 (.028 - .034)	.492	.472
Iran					
Initial solution	2878.989	1884	.032 (.030 - .035)	.722	.712
Modified solution	2866.009	1883	.032 (.030 - .034)	.725	.715
Israel					
Initial solution	2175.017	1884	.027 (.025 - .029)	.749	.739
Modified solution	2144.574	1882	.025 (.022 - .028)	.773	.765
Japan					
Initial solution	2904.025	1884	.033 (.031 - .035)	.732	.722
Modified solution	2861.529	1881	.033 (.031 - .035)	.743	.733
Mongolia					
Initial solution	3166.944	1884	.026 (.024 - .028)	.589	.574
Modified solution	3078.819	1880	.025 (.022 - .028)	.616	.601
Nigeria					
Initial solution	2282.416	1884	.023 (.020 - .027)	.666	.654
Modified solution	2249.523	1882	.022 (.018 - .026)	.692	.681
Pakistan					
Initial solution	2644.864	1884	.030 (.027 - .032)	.504	.486
Modified solution	2619.891	1882	.029 (.026 - .032)	.519	.501
Philippines					
Initial solution	2357.626	1884	.025 (.022 - .028)	.678	.666
Modified solution	2300.994	1882	.024 (.021 – .027)	.715	.704
Poland					
Initial solution	2792.991	1884	.033 (.031 - .036)	.620	.606
Modified solution	2687.857	1879	.031 (.028 - .034)	.662	.648

Table 3-3. Continued.

Country	WLSMVχ^2	*df*	RMSEA (90% C.I.)	CFI	TLI
Romania					
Initial solution	2650.835	1884	.032 (.029 - .035)	.654	.641
Modified solution	2595.766	1882	.031 (.028 - .034)	.678	.666
Samoa					
Initial solution	2501.372	1884	.029 (.026 - .032)	.633	.619
Singapore					
Initial solution	2463.218	1884	.025 (.022 - .028)	.732	.725
Modified solution	2432.524	1883	.025 (.022 - .028)	.748	.739
United States					
Initial solution	2488.575	1884	.025 (.023 - .028)	.759	.750
Modified solution	2407.145	1880	.024 (.021 - .027)	.789	.781
Venezuela					
Initial solution	2586.444	1884	.030 (.027 - .033)	.648	.635
Modified solution	2517.479	1882	.029 (.026 - .032)	.682	.670
Zimbabwe					
Initial solution	2300.126	1884	.021 (.018 - .024)	.664	.652
Modified solution	2281.020	1883	.021 (.018 - .024)	.679	.667

Table 3-4. Parameter estimates of the four-factor model in Australia data

Trait/Item #	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
4	0.501	7.300	0.749	0.213
7	0.494	7.011	0.756	-0.705
13	0.308	3.972	0.905	-0.525
16	0.389	5.704	0.849	-0.225
19	0.513	7.799	0.737	0.287
22	0.376	5.436	0.859	-0.164
26	0.376	5.268	0.859	0.092
33	0.623	10.605	0.612	0.371
36	0.442	6.211	0.805	-0.235
39	0.202	2.630	0.959	-0.296
40	0.541	8.577	0.707	0.214
42	0.633	11.481	0.599	-0.248
46	0.575	9.071	0.669	0.371
49	0.498	7.614	0.752	-0.075
52	0.426	6.165	0.819	-0.075
57	0.584	8.891	0.658	-0.707
62	0.206	2.754	0.957	-0.114
65	0.301	3.832	0.910	0.438
67	0.393	5.747	0.846	-0.038
PM				
3	0.451	5.680	0.797	0.221
6	0.527	6.450	0.722	-0.359
9	0.228	2.759	0.948	-0.137
11	0.037	0.398	0.999	0.531
15	0.401	5.014	0.839	-0.068
21	0.565	7.586	0.681	0.151
25	0.443	5.745	0.804	0.221
31	0.425	5.344	0.820	0.352
34	0.384	4.713	0.852	0.206
45	0.202	2.351	0.959	-0.052
48	0.581	7.401	0.662	0.486
51	0.220	2.621	0.952	-0.249
64	0.548	7.226	0.700	-0.294
68	0.225	2.614	0.949	0.231

Table 3-4. Continued

Trait/Item #	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.582	8.836	0.661	0.021
18	0.730	11.752	0.467	-0.096
28	0.742	11.894	0.450	0.130
37	0.702	11.016	0.507	0.565
43	0.213	2.493	0.955	-0.338
50	0.383	4.747	0.853	0.330
55	0.438	5.441	0.808	0.939
61	0.572	8.299	0.673	0.289
OL				
2	0.391	6.039	0.847	-0.178
5	0.599	11.535	0.641	-0.034
8	0.220	2.633	0.952	0.918
14	0.278	3.549	0.923	0.794
17	0.718	16.285	0.484	0.144
20	0.674	13.758	0.545	-0.481
23	0.506	8.505	0.744	0.106
27	0.453	7.335	0.794	-0.106
32	0.709	14.993	0.497	0.438
35	0.396	5.782	0.843	0.698
38	0.655	13.409	0.570	-0.116
40	0.340	4.855	0.884	0.393
44	0.293	4.130	0.914	-0.403
47	0.484	8.139	0.766	-0.456
53	0.396	6.245	0.843	-0.106
56	0.375	5.306	0.859	0.541
58	0.650	13.190	0.577	-0.182
59	0.509	8.720	0.741	-0.113
60	0.788	20.781	0.379	0.158
63	0.313	4.391	0.902	-0.446
66	0.651	13.869	0.576	-0.082
69	0.596	10.881	0.644	0.248

Table 3-5. Parameter estimates of the four-factor model in Brazil data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	0.453	8.422	0.795	-0.003
4	0.594	12.054	0.647	0.382
7	0.376	6.440	0.858	0.216
13	0.422	7.964	0.822	-0.194
16	0.619	13.556	0.617	-0.205
19	0.568	12.158	0.678	-0.112
22	0.558	10.959	0.688	-0.399
26	0.640	14.474	0.590	0.128
33	0.365	5.930	0.867	0.520
36	0.466	7.908	0.782	0.558
39	0.510	10.064	0.740	-0.014
42	0.501	7.892	0.749	-0.704
46	0.708	18.018	0.499	-0.101
49	0.532	10.481	0.717	-0.222
52	0.420	7.727	0.824	0.079
57	0.630	12.049	0.603	-0.649
62	0.429	7.469	0.816	-0.358
65	0.261	3.885	0.932	-0.732
67	0.428	7.618	0.817	0.284
PM				
3	0.311	4.842	0.903	0.596
6	0.341	5.123	0.884	-0.364
9	0.809	17.478	0.346	0.683
11	0.142	2.081	0.980	0.603
15	0.527	9.364	0.722	-0.183
21	0.231	3.388	0.947	-0.222
25	0.653	14.457	0.574	-0.041
31	0.276	4.274	0.924	0.656
34	0.335	-5.516	0.888	0.483
45	0.303	4.596	0.908	0.189
48	0.328	5.293	0.892	0.352
51	0.793	19.776	0.372	0.447
64	-0.242	-3.584	0.941	-0.228
68	0.566	10.316	0.679	0.520

Table 3-5. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.623	10.480	0.612	-0.483
18	0.746	12.396	0.443	0.435
28	0.453	6.868	0.795	-0.014
37	0.426	6.393	0.819	-0.284
43	0.425	6.262	0.820	0.465
50	-0.395	-6.027	0.844	-0.019
55	0.560	8.190	0.687	-0.718
61	0.505	8.234	0.745	0.189
OL				
2	0.322	4.831	0.896	0.435
5	0.360	5.750	0.870	-0.057
8	0.511	8.327	0.739	0.539
14	0.428	6.086	0.817	1.047
17	0.756	16.714	0.429	-0.539
20	0.397	6.024	0.842	0.489
23	0.380	5.668	0.856	0.636
27	0.143	2.039	0.980	0.429
32	0.609	11.135	0.629	-0.453
35	0.476	7.369	0.773	0.797
38	0.521	9.137	0.729	-0.211
40	0.263	4.036	0.931	0.347
44	0.105	1.595	0.989	0.312
47	0.258	3.974	0.934	0.150
53	-0.017	0.234	1.000	-0.558
56	0.147	1.952	0.979	-0.813
58	0.745	15.332	0.445	-0.739
59	0.070	-1.042	0.995	0.095
60	0.683	14.050	0.533	-0.387
63	0.202	3.070	0.959	0.256
66	0.721	16.024	0.480	0.417
69	0.497	8.570	0.753	-0.228

Table 3-6. EFA model fit indices and number of items with ≥ .30 factor loadings by factor solution for China data

Trait	No. of items	2-factor model		3-factor model			4-factor model				5-factor model				
		1	2	1	2	3	1	2	3	4	1	2	3	4	5
EI	19	1	16	1	1	16	-	17	1	3	2	1	18	3	-
PM	14	3	1	3	10	-	3	-	13	-	5	12	-	-	-
TF	8	1	1	1	3	-	1	-	-	4	1	-	-	4	-
OL	22	18	1	17	3	1	18	-	1	5	20	1	-	1	5
		Model Fit Indices													
χ^2		2524.415		2288.786			2123.139				1987.430				
df		1828		1767			1707				1640				
p		.000		.000			.000				.000				
CFI		.793		.846			.876				.899				
TLI		.779		.880			.858				.880				
RMSEA		.031		.027			.025				.023				

Table 3-7. Parameter estimates of the four-factor model in Costa Rica data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	0.216	4.711	0.911	-0.084
4	0.555	12.216	0.784	-0.957
7	0.054	0.980	0.931	0.450
13	0.439	9.966	0.747	-0.748
16	0.540	12.425	0.637	0.170
19	0.474	11.447	0.707	-0.372
22	0.333	7.200	0.845	0.146
26	0.557	14.515	0.560	0.123
33	0.310	6.970	0.958	0.088
36	0.377	8.618	0.950	-0.281
39	0.450	10.840	0.673	-0.579
42	0.490	11.154	0.781	-0.586
46	0.430	10.022	0.645	0.304
49	0.404	8.844	0.846	-0.348
52	0.256	5.634	0.724	-0.203
57	0.623	14.846	0.531	-0.935
62	0.449	10.553	0.776	-0.518
65	0.208	4.384	0.955	0.586
67	0.264	5.740	0.833	-0.372
PM				
3	0.075	1.052	0.910	0.513
6	0.010	0.150	0.973	-0.480
9	0.577	-7.442	0.988	-0.717
11	-0.186	-2.949	0.985	1.116
15	-0.150	-2.327	0.989	-0.012
21	0.391	5.801	0.982	-0.023
25	0.300	4.672	0.992	-0.052
31	0.034	0.516	0.890	0.801
34	0.336	5.045	0.999	0.296
45	0.191	3.017	0.875	0.059
48	0.376	5.863	0.594	-0.622
51	0.446	6.921	1.000	-0.600
64	0.026	0.412	0.922	-0.056
68	0.147	2.243	0.910	0.809

Table 3-7. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.527	9.225	0.723	-0.241
18	0.373	6.053	0.739	0.259
28	0.424	7.448	0.969	-0.131
37	0.478	8.100	0.876	1.427
43	-0.082	-1.409	0.995	-0.353
50	-0.132	-2.252	0.958	-0.247
55	0.408	6.809	0.975	0.867
61	0.621	10.788	0.974	-0.143
OL				
2	0.341	7.042	0.925	-0.310
5	0.412	8.725	0.706	0.420
8	0.121	2.331	0.939	0.926
14	0.367	6.550	0.783	1.729
17	0.339	6.413	0.281	0.982
20	0.482	10.431	0.614	0.433
23	0.330	6.642	0.813	1.217
27	0.248	4.996	0.897	-0.143
32	0.568	10.009	0.445	0.944
35	0.011	0.224	0.847	1.814
38	0.503	10.730	0.606	0.669
40	0.074	1.443	0.968	0.372
44	0.270	5.327	0.723	0.268
47	0.004	0.076	0.769	-0.152
53	0.377	6.896	0.762	0.363
56	0.195	3.882	0.682	0.190
58	0.365	7.401	0.652	0.755
59	0.488	10.604	1.000	0.061
60	0.415	8.561	0.364	0.676
63	0.269	5.396	0.969	0.642
66	0.487	9.150	0.455	0.738
69	0.178	3.355	0.681	0.410

Table 3-8. EFA model fit indices and number of items with ≥ .30 factor loadings by factor solution for Egypt data

Trait	No. of items	2-factor model		3-factor model			4-factor model				5-factor model				
		1	2	1	2	3	1	2	3	4	1	2	3	4	5
EI	19	1	2	2	2	9	-	-	2	9	2	-	2	-	8
PM	14	4	1	5	1	-	3	5	1	1	5	2	2	2	-
TF	8	1	2	1	2	-	1	1	2	-	1	1	2	-	-
OL	22	10	5	11	6	-	6	5	4	-	8	-	5	6	-
		Model Fit Indices													
χ^2		2795.259		2496.157			2235.302				2062.220				
df		1828		1767			1707				1648				
p		.000		.000			.000				.000				
CFI		.627		.718			.796				.840				
TLI		.601		.689			.767				.810				
RMSEA		.024		.021			.018				.017				

Table 3-9. EFA model fit indices and number of items with ≥ .30 factor loadings by factor solution for Gaza data

Trait	No. of items	2-factor model		3-factor model			4-factor model				5-factor model				
		1	2	1	2	3	1	2	3	4	1	2	3	4	5
EI	19	3	10	4	9	5	3	9	4	-	3	5	5	1	7
PM	14	8	4	11	4	6	8	4	5	-	8	5	5	-	1
TF	8	3	-	3	-	2	4	-	2	4	3	-	2	4	1
OL	22	19	1	18	-	1	19	1	1	1	16	2	1	2	1
		Model Fit Indices													
χ^2		2152.565		2037.012			1932.922				1834.679				
df		1828		1767			1707				1648				
p		.000		.000			.000				.000				
CFI		.862		.885			.904				.921				
TLI		.853		.873			.890				.906				
RMSEA		.022		.020			.019				.017				

Table 3-10. Parameter estimates of the four-factor model in Hungary data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	0.463	7.288	0.785	-0.091
4	0.574	8.379	0.671	-0.708
7	0.172	2.167	0.970	0.741
13	0.468	7.714	0.781	-0.128
16	0.321	4.630	0.897	0.016
19	0.360	5.163	0.870	-0.224
22	0.300	3.968	0.911	0.505
26	0.333	4.862	0.889	-0.022
33	0.197	2.569	0.961	-0.361
36	0.395	5.842	0.844	-0.282
39	0.591	10.089	0.651	-0.041
42	0.458	6.603	0.790	-0.519
46	0.342	5.085	0.883	0.047
49	0.485	6.903	0.765	-0.374
52	0.480	7.786	0.770	-0.078
57	0.460	6.796	0.788	-0.622
62	0.460	6.661	0.788	-0.368
65	0.221	2.925	0.951	-0.428
67	0.245	3.429	0.940	-0.173
PM				
3	0.321	4.390	0.897	0.749
6	0.531	8.071	0.718	-0.122
9	-0.010	-0.121	1.000	-0.835
11	0.340	4.809	0.884	0.415
15	0.243	3.357	0.941	-0.128
21	0.222	3.089	0.951	0.097
25	0.220	3.083	0.951	-0.110
31	0.390	5.512	0.848	0.505
34	0.269	3.896	0.928	0.198
45	0.463	6.673	0.786	-0.262
48	0.288	4.219	0.917	0.022
51	0.252	3.382	0.936	-0.505
64	0.192	2.580	0.963	-0.147
68	0.281	3.999	0.921	0.548

Table 3-10. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.436	5.244	0.809	-0.166
18	0.109	1.210	0.988	-0.512
28	0.172	1.921	0.970	-0.009
37	0.327	3.299	0.893	0.630
43	0.588	6.338	0.654	-0.230
50	-0.175	-1.958	0.969	0.154
55	0.105	1.004	0.989	0.880
61	0.427	5.107	0.818	0.205
OL				
2	0.255	3.836	0.935	0.072
5	0.655	12.391	0.571	0.456
8	0.403	5.536	0.838	0.645
14	0.255	3.826	0.935	0.097
17	0.615	9.898	0.622	0.630
20	-0.013	-0.186	1.000	0.003
23	0.398	5.760	0.842	0.630
27	0.053	0.703	0.997	-0.676
32	0.466	7.086	0.783	0.463
35	-0.083	-1.039	0.993	0.638
38	0.230	3.228	0.947	0.394
40	0.313	3.962	0.902	0.700
44	0.305	4.392	0.907	-0.321
47	0.304	4.656	0.908	-0.224
53	0.220	3.084	0.951	0.308
56	0.287	3.519	0.918	0.791
58	0.533	8.828	0.716	0.078
59	0.096	1.223	0.991	-0.669
60	0.700	12.135	0.510	0.615
63	0.387	5.926	0.850	0.003
66	0.456	6.628	0.792	0.519
69	0.203	2.838	0.959	0.301

Table 3-11. Parameter estimates of the four-factor model in Iran data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	0.417	6.830	0.826	0.485
4	-0.115	-1.739	0.987	-0.425
7	0.324	5.189	0.895	0.422
13	0.117	1.804	0.986	0.064
16	0.278	4.366	0.923	0.238
19	0.031	0.484	0.999	-0.054
22	0.419	6.838	0.825	0.377
26	-0.045	-0.695	0.998	-0.148
33	0.026	0.388	0.999	-0.146
36	-0.190	-2.970	0.964	-0.325
39	0.291	4.633	0.915	-0.111
42	0.520	8.950	0.729	-0.377
46	0.286	4.476	0.918	0.258
49	0.715	15.729	0.488	-0.785
52	0.380	6.214	0.856	0.304
57	0.352	5.889	0.876	-0.279
62	0.218	3.373	0.952	-0.356
65	0.416	6.935	0.827	0.325
67	0.233	3.623	0.946	0.027
PM				
3	-0.069	-1.034	0.995	0.436
6	0.215	3.522	0.954	-0.511
9	0.519	9.840	0.730	-0.665
11	-0.034	-0.512	0.999	0.228
15	0.274	4.449	0.925	-0.138
21	0.149	2.359	0.978	-0.144
25	0.512	9.519	0.738	-0.310
31	0.489	8.266	0.761	0.582
34	0.081	1.264	0.994	-0.012
45	0.495	9.214	0.755	-0.469
48	0.628	13.430	0.606	-0.678
51	0.565	11.120	0.681	-0.519
64	0.254	4.140	0.936	-0.139
68	0.399	6.683	0.840	0.404

Table 3-11. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.007	0.083	1.000	0.000
18	0.224	2.830	0.950	-0.335
28	0.179	2.215	0.968	-0.074
37	0.322	4.382	0.897	-0.108
43	0.135	1.668	0.982	-0.153
50	0.981	8.548	0.038	0.458
55	0.087	1.062	0.992	0.133
61	0.027	0.321	0.999	0.206
OL				
2	0.358	6.296	0.872	0.424
5	0.409	7.189	0.833	0.372
8	0.319	5.451	0.898	0.218
14	-0.080	-1.298	0.994	-0.208
17	0.658	15.221	0.567	0.653
20	0.459	8.705	0.789	0.218
23	0.571	12.244	0.674	0.623
27	-0.058	-0.942	0.997	-0.123
32	0.690	16.680	0.524	0.547
35	0.340	5.885	0.885	0.069
38	0.491	9.553	0.759	0.469
40	0.518	10.167	0.732	0.297
44	0.458	8.735	0.790	0.208
47	0.579	12.042	0.665	0.223
53	0.535	10.560	0.714	0.346
56	0.502	9.588	0.748	0.541
58	0.593	12.995	0.648	0.627
59	0.627	14.113	0.607	0.617
60	0.691	17.363	0.523	0.617
63	0.197	3.301	0.961	0.143
66	0.645	14.994	0.583	0.528
69	0.467	8.270	0.782	0.528

Table 3-12. Parameter estimates of the four-factor model in Israel data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	0.279	3.087	0.922	-0.037
4	0.763	12.682	0.418	-0.552
7	0.628	8.483	0.606	0.522
13	0.534	6.907	0.715	-0.491
16	0.434	4.728	0.811	0.576
19	0.530	7.001	0.719	0.047
22	0.459	4.340	0.789	1.248
26	0.551	7.463	0.696	0.355
33	0.203	2.276	0.959	-0.128
36	0.448	4.740	0.799	-0.678
39	0.540	6.945	0.709	-0.422
42	0.658	8.113	0.567	-0.734
46	0.708	10.458	0.498	0.660
49	0.470	5.252	0.780	-0.667
52	0.478	5.873	0.772	-0.486
57	0.693	9.892	0.520	-0.674
62	0.297	3.390	0.912	-0.122
65	0.285	2.827	0.919	0.594
67	0.466	5.808	0.783	-0.094
PM				
3	0.565	7.029	0.681	0.194
6	0.193	1.907	0.963	-0.152
9	0.798	13.734	0.363	-0.554
11	0.196	1.986	0.961	0.088
15	0.546	6.302	0.702	-0.286
21	0.082	0.802	0.993	0.256
25	0.505	6.143	0.745	-0.095
31	0.647	9.042	0.581	-0.232
34	0.355	3.634	0.874	0.719
45	0.237	2.502	0.944	-0.190
48	0.382	3.947	0.854	-0.426
51	0.822	14.179	0.324	-0.572
64	0.014	0.141	1.000	-0.311
68	0.142	1.240	0.980	0.715

Table 3-12. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.551	6.716	0.697	-0.282
18	0.752	10.859	0.434	-0.220
28	0.710	10.020	0.496	-0.046
37	0.536	5.648	0.713	0.435
43	0.068	0.631	0.995	-0.435
50	0.479	4.951	0.771	-0.253
55	0.589	6.370	0.653	0.530
61	0.632	8.349	0.600	-0.006
OL				
2	0.400	4.555	0.840	0.146
5	0.518	6.942	0.732	0.285
8	0.553	6.519	0.694	0.796
14	0.117	1.151	0.986	0.538
17	0.777	13.503	0.396	0.625
20	0.617	8.627	0.620	-0.439
23	0.448	5.172	0.799	0.401
27	-0.092	-0.825	0.992	-0.617
32	0.627	8.888	0.607	0.761
35	0.216	2.102	0.953	0.542
38	0.748	12.095	0.440	0.197
40	0.340	3.898	0.884	0.205
44	0.317	3.542	0.899	0.270
47	-0.031	0.280	0.999	-0.812
53	0.237	2.592	0.944	-0.133
56	0.246	2.239	0.940	0.902
58	0.510	6.420	0.740	0.328
59	0.019	-0.196	1.000	-0.190
60	0.732	12.322	0.464	0.469
63	-0.176	-1.407	0.969	-1.248
66	0.752	13.101	0.434	0.552
69	0.584	7.925	0.658	0.490

Table 3-13. Parameter estimates of the four-factor model in Japan data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	0.646	14.318	0.583	0.388
4	0.664	13.900	0.559	-0.651
7	0.290	5.022	0.916	0.421
13	0.518	10.121	0.732	-0.449
16	0.661	15.877	0.563	0.334
19	0.466	7.734	0.783	-0.832
22	0.577	12.108	0.667	0.302
26	0.606	13.856	0.632	0.074
33	0.437	7.564	0.809	-0.489
36	0.370	6.046	0.863	-0.524
39	0.504	10.225	0.746	0.059
42	0.550	8.462	0.698	-1.002
46	0.638	15.217	0.593	0.223
49	0.588	12.651	0.655	-0.140
52	0.534	10.754	0.715	-0.223
57	0.686	15.673	0.529	-0.589
62	0.387	6.692	0.851	-0.472
65	0.380	6.901	0.856	0.156
67	0.570	12.347	0.675	-0.275
PM				
3	0.178	2.530	0.968	0.003
6	-0.057	-0.795	0.997	-0.028
9	0.753	13.457	0.433	-0.768
11	0.383	5.638	0.854	-0.218
15	0.377	5.438	0.858	-0.318
21	0.200	2.709	0.960	-0.421
25	0.573	9.969	0.672	-0.244
31	-0.096	-1.336	0.991	0.115
34	0.129	1.797	0.983	0.583
45	0.342	4.859	0.883	-0.444
48	0.329	4.718	0.892	-0.461
51	0.871	16.369	0.241	-0.607
64	0.343	4.992	0.882	-0.270
68	0.251	3.594	0.937	-0.207

Table 3-13. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.523	8.889	0.727	0.140
18	0.747	11.086	0.442	-0.728
28	0.622	10.571	0.613	-0.048
37	0.586	9.284	0.656	-0.161
43	0.241	3.200	0.942	-0.313
50	0.311	4.492	0.903	0.048
55	0.482	7.529	0.768	0.339
61	0.520	8.668	0.729	0.595
OL				
2	0.471	8.854	0.778	-0.286
5	0.781	23.119	0.390	0.161
8	0.023	0.330	0.999	0.644
14	-0.065	-1.039	0.996	-0.104
17	0.608	12.124	0.630	0.559
20	0.611	13.755	0.626	-0.286
23	0.222	3.556	0.951	0.478
27	0.404	7.126	0.837	0.334
32	0.514	9.349	0.736	0.722
35	0.079	1.275	0.994	-0.094
38	0.736	19.670	0.459	0.018
40	0.348	5.994	0.879	0.361
44	0.397	7.447	0.843	-0.249
47	0.363	6.273	0.868	-0.291
53	0.087	1.403	0.992	-0.202
56	0.389	6.421	0.849	0.728
58	0.640	13.214	0.590	0.536
59	0.305	5.023	0.907	0.518
60	0.825	26.907	0.319	0.110
63	0.343	5.507	0.882	-0.489
66	0.609	12.932	0.629	-0.110
69	0.635	13.607	0.596	0.356

Table 3-14. Parameter estimates of the four-factor model in Mongolia data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	0.216	4.696	0.890	0.045
4	0.555	12.224	0.690	-0.913
7	0.035	0.641	0.999	0.982
13	0.439	9.97	0.806	-0.386
16	0.540	12.405	0.711	0.460
19	0.474	11.447	0.777	0.014
22	0.332	7.176	0.889	0.390
26	0.557	14.509	0.695	-0.151
33	0.309	6.945	0.796	-0.060
36	0.377	8.628	0.864	-0.265
39	0.450	10.83	0.799	0.306
42	0.490	11.169	0.758	-0.718
46	0.429	9.999	0.819	0.342
49	0.404	8.843	0.837	-0.611
52	0.255	5.615	0.616	-0.806
57	0.624	14.857	0.788	0.099
62	0.449	10.555	0.797	-0.085
65	0.207	4.363	0.959	0.248
67	0.264	5.735	0.932	-0.061
PM				
3	0.074	1.043	1.000	0.623
6	0.010	0.147	1.000	-0.369
9	0.577	7.437	0.706	-0.299
11	-0.187	-2.957	0.942	-0.013
15	-0.151	-2.330	0.977	-0.627
21	0.391	5.802	0.852	0.744
25	0.300	4.672	0.922	-0.583
31	0.033	0.507	0.861	0.335
34	0.336	5.049	0.889	-0.482
45	0.191	3.012	0.976	-0.441
48	0.376	5.865	0.862	-0.643
51	0.446	6.921	0.840	-0.885
64	0.026	0.410	1.000	-0.024
68	0.147	2.247	0.974	0.231

Table 3-14. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.527	9.222	0.717	-0.253
18	0.373	6.056	0.879	-0.712
28	0.424	7.447	0.826	-0.275
37	0.478	8.103	0.789	0.633
43	-0.082	-1.412	0.996	0.128
50	-0.132	-2.254	0.968	0.412
55	0.409	6.817	0.817	0.732
61	0.620	10.78	0.656	0.049
OL				
2	0.341	7.038	0.877	-0.120
5	0.412	8.724	0.824	0.130
8	-0.121	-2.331	0.988	-0.314
14	0.367	6.546	0.859	0.892
17	0.339	6.416	0.884	0.590
20	0.481	10.423	0.760	-0.111
23	0.330	6.647	0.903	0.850
27	0.248	4.991	0.931	0.005
32	0.568	10.01	0.689	0.390
35	-0.012	-0.228	1.000	-0.097
38	0.503	10.723	0.728	0.596
40	0.074	1.439	0.993	0.263
44	0.270	5.323	0.928	-0.380
47	0.004	0.075	1.000	-0.406
53	0.377	6.899	0.852	0.843
56	0.195	3.881	0.956	0.213
58	0.365	7.402	0.862	0.789
59	0.489	10.61	0.788	0.099
60	0.415	8.554	0.817	0.453
63	0.269	5.394	0.926	0.314
66	0.488	9.154	0.754	0.930
69	0.179	3.361	0.971	0.664

Table 3-15. Parameter estimates of the four-factor model in Nigeria data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	0.293	3.616	0.914	0.631
4	0.394	4.896	0.844	-0.726
7	0.197	2.554	0.961	0.697
13	0.436	6.314	0.810	0.352
16	0.428	6.152	0.817	-0.045
19	0.436	5.932	0.810	-0.280
22	0.377	4.995	0.858	0.375
26	0.542	7.476	0.706	-0.213
33	0.346	4.621	0.881	-0.259
36	0.272	3.653	0.926	-0.055
39	0.412	5.899	0.830	0.003
42	0.205	2.689	0.958	-0.223
46	0.463	6.587	0.786	0.492
49	0.342	4.642	0.883	-0.245
52	0.374	4.311	0.86	0.827
57	0.497	6.956	0.753	-0.417
62	0.272	3.585	0.926	-0.115
65	0.248	3.324	0.939	0.534
67	0.336	4.671	0.887	0.670
PM				
3	0.156	1.835	0.976	0.273
6	0.212	2.159	0.955	-0.809
9	0.688	8.127	0.527	-0.931
11	-0.091	-0.895	0.992	0.649
15	0.455	5.558	0.793	-0.343
21	0.448	4.787	0.799	-0.807
25	0.391	4.855	0.847	-0.325
31	-0.111	-1.242	0.988	0.329
34	0.172	2.006	0.970	-0.010
45	0.023	0.263	0.999	-0.543
48	0.517	5.813	0.732	-0.888
51	0.385	4.505	0.852	-0.707
64	0.380	4.451	0.855	-0.401
68	0.098	1.121	0.990	0.507

Table 3-15. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.467	4.865	0.781	-0.096
18	0.613	5.976	0.625	-0.681
28	0.309	3.260	0.904	-0.090
37	-0.067	-0.678	0.995	0.391
43	0.182	1.684	0.967	-0.584
50	0.453	4.347	0.794	-0.505
55	-0.149	-1.464	0.978	0.571
61	0.306	3.315	0.906	0.026
OL				
2	0.164	2.208	0.973	0.265
5	0.298	4.335	0.911	0.068
8	0.549	7.383	0.698	1.124
14	0.282	3.625	0.920	0.713
17	0.891	12.381	0.206	1.563
20	0.647	12.109	0.582	0.483
23	0.274	3.342	0.925	1.034
27	0.259	3.539	0.933	0.259
32	0.589	8.417	0.653	1.175
35	0.307	3.528	0.906	1.138
38	0.582	8.473	0.661	0.921
40	-0.018	-0.238	1.000	0.109
44	0.323	4.399	0.895	0.532
47	0.558	9.364	0.688	0.387
53	0.274	3.830	0.925	0.077
56	0.330	4.671	0.891	0.339
58	0.685	11.209	0.530	0.890
59	-0.034	-0.458	0.999	0.087
60	0.630	10.059	0.603	0.825
63	0.452	6.690	0.796	0.424
66	0.696	12.220	0.516	0.732
69	0.426	5.874	0.818	0.774

Table 3-16. Parameter estimates of the four-factor model in Pakistan data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	0.345	5.057	0.881	0.055
4	0.324	4.541	0.895	-0.217
7	0.442	6.167	0.805	0.523
13	0.358	5.129	0.872	0.188
16	0.412	6.04	0.830	0.377
19	0.153	1.969	0.976	-0.437
22	0.226	3.071	0.949	-0.063
26	0.416	5.988	0.827	0.205
33	0.291	4.035	0.915	-0.253
36	0.184	2.387	0.966	-0.289
39	0.246	3.308	0.940	-0.102
42	0.451	6.914	0.797	-0.163
46	0.378	5.268	0.857	0.383
49	0.247	3.385	0.939	-0.363
52	0.400	5.787	0.840	-0.102
57	0.429	6.479	0.816	-0.336
62	0.230	3.118	0.947	-0.238
65	0.157	2.124	0.975	0.146
67	0.225	3.091	0.949	-0.022
PM				
3	-0.191	-2.309	0.964	0.584
6	0.317	4.378	0.900	-0.110
9	0.585	8.26	0.658	-0.951
11	-0.045	-0.533	0.998	0.709
15	0.289	3.958	0.916	-0.044
21	0.304	4.098	0.907	-0.296
25	0.306	4.144	0.906	-0.224
31	-0.009	-0.107	1.000	0.288
34	-0.264	-3.601	0.930	-0.174
45	0.374	5.159	0.860	-0.358
48	0.637	9.059	0.594	-0.600
51	0.449	6.077	0.799	-0.761
64	0.392	5.308	0.846	-0.200
68	-0.077	-1.008	0.994	0.099

Table 3-16. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.459	5.775	0.789	-0.136
18	0.569	6.903	0.676	-0.294
28	0.407	5.113	0.834	-0.005
37	0.337	4.169	0.886	-0.118
43	-0.026	-0.307	0.999	-0.393
50	0.111	1.303	0.988	-0.187
55	0.228	2.758	0.948	0.429
61	0.655	7.884	0.571	0.166
OL				
2	0.164	2.262	0.973	0.471
5	0.234	3.577	0.945	-0.143
8	0.057	0.813	0.997	0.306
14	-0.104	-1.452	0.989	0.292
17	0.691	11.136	0.523	1.132
20	0.490	8.489	0.760	0.121
23	0.458	7.220	0.790	0.678
27	0.050	0.737	0.997	-0.121
32	0.547	9.195	0.701	0.607
35	0.178	2.496	0.968	0.481
38	0.534	9.665	0.714	0.393
40	-0.031	-0.450	0.999	-0.188
44	0.489	8.351	0.761	0.615
47	0.504	8.903	0.746	0.273
53	0.268	4.148	0.928	0.224
56	0.332	5.267	0.890	0.190
58	0.480	8.134	0.769	0.385
59	0.358	5.663	0.872	0.264
60	0.651	12.186	0.576	0.615
63	0.053	0.792	0.997	-0.213
66	0.718	15.599	0.485	0.609
69	0.370	5.657	0.863	0.642

Table 3-17. Parameter estimates of the four-factor model in Philippines data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	-0.042	-0.482	0.998	0.737
4	0.418	5.777	0.826	-0.431
7	0.211	2.453	0.956	0.704
13	0.300	3.868	0.910	-0.452
16	0.461	6.704	0.788	0.515
19	0.233	2.984	0.946	-0.142
22	0.279	3.573	0.922	0.336
26	0.550	8.471	0.698	0.079
33	0.426	5.607	0.819	-0.218
36	0.307	3.758	0.906	-0.641
39	0.500	7.260	0.750	0.193
42	0.442	5.478	0.805	-0.721
46	0.638	10.195	0.593	0.248
49	0.488	5.925	0.762	-0.573
52	0.188	2.316	0.965	0.618
57	0.518	6.510	0.732	-0.867
62	0.379	5.195	0.857	-0.218
65	0.139	1.743	0.981	0.110
67	0.269	3.594	0.928	0.290
PM				
3	-0.043	-0.501	0.998	0.349
6	0.407	4.270	0.834	-0.831
9	0.599	7.080	0.641	-0.983
11	0.010	0.115	1.000	0.174
15	0.239	2.818	0.943	-0.336
21	0.200	2.498	0.960	0.180
25	0.699	9.343	0.511	-0.596
31	-0.084	-1.002	0.993	0.199
34	0.239	3.010	0.943	0.316
45	0.572	7.957	0.673	-0.277
48	0.459	6.964	0.789	0.016
51	0.546	6.847	0.702	-0.831
64	0.290	3.295	0.916	-0.737
68	-0.196	-2.292	0.962	0.303

Table 3-17. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	-0.028	-0.287	0.999	0.270
18	0.826	7.308	0.318	-0.522
28	0.596	5.989	0.645	-0.456
37	-0.105	-0.995	0.989	0.904
43	0.144	1.371	0.979	-0.501
50	0.216	2.231	0.953	0.116
55	0.322	3.597	0.896	0.641
61	0.135	1.368	0.982	0.231
OL				
2	0.105	1.476	0.989	-0.079
5	0.770	15.564	0.408	0.696
8	0.311	4.617	0.903	0.091
14	0.228	2.857	0.948	0.787
17	0.593	7.406	0.648	1.238
20	0.469	7.598	0.780	0.206
23	0.112	1.147	0.987	1.421
27	0.569	9.721	0.676	0.257
32	0.503	5.781	0.747	1.173
35	0.527	7.250	0.722	1.068
38	0.636	11.046	0.595	0.551
40	0.217	2.880	0.953	0.551
44	0.582	9.044	0.661	0.762
47	0.575	9.710	0.669	0.316
53	0.516	7.847	0.734	0.522
56	0.287	3.728	0.917	0.688
58	0.401	6.192	0.840	-0.085
59	0.332	4.820	0.890	-0.167
60	0.749	14.504	0.440	0.618
63	0.014	0.180	1.000	-0.438
66	0.558	8.077	0.688	0.923
69	0.198	2.802	0.961	0.120

Table 3-18. Parameter estimates of the four-factor model in Poland data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	0.528	9.317	0.721	0.081
4	0.498	7.970	0.752	-0.366
7	0.277	4.040	0.841	-0.546
13	0.268	3.861	0.928	0.511
16	0.760	15.592	0.422	0.219
19	0.226	3.153	0.949	-0.219
22	0.488	7.817	0.761	0.273
26	0.698	13.914	0.513	0.055
33	0.500	8.089	0.750	0.049
36	0.129	1.813	0.983	-0.189
39	0.444	7.074	0.803	0.316
42	0.248	3.378	0.938	-0.526
46	0.610	11.374	0.628	0.151
49	0.484	7.869	0.766	-0.251
52	0.340	5.131	0.884	0.046
57	0.285	3.905	0.919	-0.612
62	0.025	0.353	0.999	-0.113
65	0.297	4.230	0.912	0.664
67	0.384	5.856	0.853	-0.287
PM				
3	0.439	5.976	0.807	0.209
6	0.399	5.198	0.841	-0.546
9	0.397	4.954	0.842	-0.848
11	0.403	5.271	0.838	0.257
15	0.665	9.732	0.557	-0.635
21	-0.051	-0.609	0.997	-0.333
25	0.386	5.077	0.851	-0.157
31	0.331	4.372	0.890	-0.055
34	0.223	2.724	0.950	0.348
45	0.248	3.138	0.938	0.287
48	0.573	7.815	0.671	0.348
51	0.383	4.972	0.853	-0.678
64	0.289	3.664	0.917	-0.168
68	-0.170	-2.035	0.971	0.532

Table 3-18. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.620	9.475	0.616	-0.04
18	0.652	9.782	0.575	-0.397
28	0.710	11.528	0.496	0.052
37	0.422	5.919	0.822	-0.114
43	0.250	3.277	0.937	-0.493
50	0.205	2.714	0.958	0.331
55	0.495	7.087	0.755	0.486
61	0.514	7.581	0.736	0.157
OL				
2	0.377	5.820	0.858	-0.215
5	0.487	7.213	0.763	-0.566
8	0.384	6.138	0.853	0.245
14	0.545	8.571	0.704	0.776
17	0.655	14.110	0.571	0.435
20	0.397	6.142	0.842	-0.605
23	0.255	3.790	0.935	0.339
27	0.287	4.362	0.918	-0.58
32	0.589	11.238	0.653	0.576
35	0.573	10.529	0.672	0.46
38	0.501	8.866	0.749	0.151
40	0.400	6.692	0.840	0.263
44	0.327	5.137	0.893	0.017
47	0.480	7.893	0.770	-0.385
53	0.277	4.299	0.923	-0.069
56	0.226	3.398	0.949	0.339
58	0.544	9.032	0.704	0.664
59	0.020	0.286	1.000	-0.573
60	0.712	15.952	0.492	-0.081
63	-0.022	-0.301	1.000	-0.865
66	0.647	13.677	0.582	0.116
69	0.492	8.480	0.758	0.287

Table 3-19. Parameter estimates of the four-factor model in Romania data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	-0.063	-0.789	0.996	0.433
4	0.578	8.891	0.666	-0.687
7	0.125	1.578	0.984	0.341
13	0.200	2.545	0.960	0.223
16	0.536	8.523	0.712	0.164
19	0.446	6.151	0.801	-0.569
22	0.228	2.925	0.948	0.617
26	0.631	10.610	0.601	-0.256
33	-0.429	-5.688	0.816	0.207
36	0.442	6.114	0.805	-0.361
39	0.366	4.989	0.866	0.077
42	0.365	4.671	0.866	-0.541
46	0.618	9.173	0.618	-0.003
49	0.539	7.853	0.710	-0.445
52	0.192	2.430	0.963	-0.371
57	0.515	7.226	0.735	-0.825
62	0.571	8.577	0.674	-0.323
65	0.116	1.486	0.987	-0.051
67	0.196	2.553	0.961	0.035
PM				
3	0.439	5.548	0.807	0.293
6	0.408	4.476	0.833	-0.623
9	0.322	3.564	0.897	-0.767
11	0.016	0.183	1.000	0.758
15	0.494	6.685	0.756	0.094
21	-0.064	-0.773	0.996	-0.190
25	0.522	6.633	0.728	-0.426
31	0.205	2.456	0.958	0.296
34	0.324	4.150	0.895	0.161
45	0.388	5.041	0.850	-0.010
48	0.575	7.063	0.670	-0.438
51	0.402	4.372	0.838	-1.097
64	-0.206	-2.481	0.958	-0.304
68	0.093	1.052	0.991	0.629

Table 3-19. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.198	1.926	0.902	-0.240
18	0.313	3.139	0.672	0.019
28	0.573	5.457	0.958	0.296
37	0.377	3.806	0.858	0.478
43	0.508	4.846	0.741	-0.336
50	0.168	1.574	0.972	0.183
55	0.382	3.545	0.854	0.795
61	0.427	4.135	0.817	0.119
OL				
2	0.412	6.433	0.830	0.296
5	0.616	11.152	0.621	0.155
8	0.374	5.443	0.860	0.491
14	0.455	6.978	0.793	0.662
17	0.750	15.052	0.438	0.770
20	0.678	13.051	0.540	-0.019
23	0.556	8.704	0.691	0.771
27	0.515	8.905	0.734	-0.410
32	0.824	18.742	0.320	0.745
35	0.476	7.444	0.773	0.594
38	0.611	11.582	0.626	0.371
40	0.091	1.281	0.992	0.016
44	0.421	6.681	0.823	-0.097
47	0.423	6.562	0.821	0.022
53	0.147	2.093	0.978	0.151
56	0.390	5.556	0.848	0.755
58	0.604	11.197	0.636	0.251
59	-0.011	-0.157	1.000	-0.273
60	0.634	11.598	0.598	0.257
63	0.042	0.579	0.998	-0.454
66	0.655	11.924	0.571	0.524
69	0.361	5.379	0.869	0.371

Table 3-20. Parameter estimates of the four-factor model in Samoa data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	0.016	0.185	0.999	1.200
4	0.365	6.041	0.867	-0.468
7	0.709	13.34	0.497	0.157
13	0.590	9.023	0.652	0.385
16	-0.020	-0.276	0.999	-0.063
19	0.510	9.732	0.740	-1.326
22	0.286	4.159	0.918	-0.496
26	0.214	2.878	0.954	-0.517
33	0.437	8.393	0.809	-0.682
36	0.414	6.735	0.828	-0.496
39	0.321	4.085	0.897	0.598
42	0.527	10.369	0.722	-0.824
46	0.540	9.785	0.708	-0.372
49	0.447	8.268	0.801	-0.454
52	0.435	7.275	0.811	0.824
57	0.245	3.51	0.940	-0.517
62	0.500	10.105	0.750	-0.636
65	0.363	6.556	0.868	0.789
67	0.176	2.581	0.969	-0.038
PM				
3	0.217	3.179	0.953	-0.228
6	0.201	2.923	0.960	-0.824
9	0.388	6.486	0.849	-1.115
11	-0.069	-1.016	0.995	0.419
15	0.017	0.232	0.999	-0.161
21	0.398	6.497	0.842	-0.598
25	0.332	4.728	0.890	-1.015
31	0.369	5.851	0.864	-0.524
34	-0.232	-2.969	0.946	-0.247
45	0.354	4.584	0.875	-1.282
48	0.519	7.989	0.731	-0.798
51	0.579	10.314	0.665	-0.954
64	-0.011	-0.142	0.999	-0.575
68	0.311	4.407	0.903	-0.392

Table 3-20. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.209	2.92	0.956	0.138
18	0.284	4.243	0.920	-0.461
28	0.231	3.21	0.947	0.132
37	0.410	4.666	0.832	1.555
43	-0.264	-3.915	0.930	0.399
50	0.450	7.424	0.797	-0.842
55	0.456	8.32	0.792	0.674
61	0.630	9.486	0.603	0.126
OL				
2	0.282	3.733	0.921	-0.496
5	0.432	4.323	0.813	0.944
8	0.276	4.842	0.924	0.706
14	0.226	2.468	0.949	1.457
17	0.492	8.005	0.758	1.576
20	-0.193	-2.981	0.963	-0.273
23	0.527	8.523	0.722	1.695
27	0.604	10.98	0.636	0.069
32	0.494	6.787	0.756	1.812
35	0.279	3.631	0.922	0.994
38	0.304	3.471	0.907	1.282
40	0.432	8.181	0.814	0.379
44	-0.255	-3.422	0.935	0.352
47	0.216	3.206	0.953	0.561
53	0.490	8.739	0.760	0.781
56	0.381	5.04	0.855	1.240
58	0.370	5.805	0.863	0.915
59	0.494	9.413	0.756	-1.058
60	0.463	7.534	0.786	1.047
63	0.363	6.221	0.868	0.365
66	0.499	10.2	0.751	1.069
69	0.231	3.534	0.947	0.532

Table 3-21. Parameter estimates of the four-factor model in Singapore data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	0.272	4.427	0.926	0.191
4	0.641	12.564	0.589	-0.644
7	0.377	6.184	0.858	0.378
13	0.508	9.000	0.742	-0.446
16	0.661	13.216	0.563	0.378
19	0.417	7.120	0.826	-0.408
22	0.508	9.815	0.742	-0.023
26	0.680	14.776	0.538	0.397
33	0.164	2.592	0.973	0.175
36	0.377	6.444	0.858	-0.062
39	0.487	9.172	0.763	-0.271
42	0.566	10.092	0.680	-0.686
46	0.497	9.016	0.753	0.352
49	0.495	9.245	0.755	-0.235
52	0.309	5.061	0.905	0.283
57	0.620	13.278	0.616	-0.288
62	0.434	7.618	0.812	-0.194
65	0.189	2.932	0.964	0.339
67	0.403	7.158	0.838	0.199
PM				
3	0.408	5.620	0.833	-0.013
6	0.325	4.294	0.894	-0.317
9	0.494	6.482	0.756	-1.014
11	0.386	5.176	0.851	0.086
15	0.378	5.185	0.857	-0.120
21	0.186	2.416	0.965	-0.016
25	0.428	5.911	0.817	-0.185
31	0.338	4.523	0.886	-0.003
34	0.188	2.411	0.965	0.271
45	0.182	2.292	0.967	-0.287
48	0.272	3.569	0.926	-0.276
51	0.627	8.740	0.606	-0.753
64	0.240	3.085	0.942	-0.086
68	0.230	3.012	0.947	0.062

Table 3-21. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.567	8.595	0.678	-0.068
18	0.469	6.324	0.780	-0.609
28	0.640	9.976	0.590	-0.060
37	0.460	6.309	0.788	0.578
43	0.073	0.846	0.995	-0.746
50	0.238	3.185	0.944	-0.096
55	0.446	5.195	0.801	1.004
61	0.593	9.024	0.649	0.279
OL				
2	0.371	5.787	0.862	0.534
5	0.596	12.167	0.645	0.419
8	0.324	4.680	0.895	0.699
14	0.383	3.934	0.854	1.456
17	0.623	8.206	0.612	1.357
20	0.646	13.530	0.583	0.425
23	0.339	4.805	0.885	1.005
27	0.356	5.659	0.874	0.049
32	0.671	10.358	0.550	1.250
35	0.406	4.718	0.836	1.309
38	0.685	13.941	0.530	0.753
40	0.164	2.403	0.973	0.454
44	0.403	6.207	0.838	0.641
47	0.356	5.958	0.873	0.112
53	0.333	5.323	0.889	-0.013
56	0.300	4.636	0.910	0.435
58	0.502	8.476	0.748	0.510
59	0.226	-3.432	0.949	-0.068
60	0.823	19.564	0.323	0.827
63	0.214	3.275	0.954	0.089
66	0.689	12.767	0.525	0.845
69	0.540	9.277	0.708	0.691

Table 3-22. Parameter estimates of the four-factor model in U.S. data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	0.428	7.525	0.817	0.217
4	0.575	7.966	0.670	-1.146
7	0.207	3.159	0.957	0.524
13	0.362	5.665	0.869	-0.490
16	0.619	13.517	0.617	0.111
19	0.274	4.485	0.925	-0.015
22	0.476	8.681	0.773	0.233
26	0.666	15.751	0.556	0.264
33	0.296	4.669	0.912	0.496
36	0.258	4.041	0.934	-0.429
39	0.241	3.821	0.942	-0.105
42	0.377	6.220	0.858	-0.396
46	0.627	13.648	0.607	0.337
49	0.488	9.156	0.762	-0.238
52	0.380	6.342	0.856	0.311
57	0.569	10.346	0.677	-0.706
62	0.215	3.447	0.954	-0.166
65	0.332	5.503	0.890	0.327
67	0.405	7.196	0.836	-0.187
PM				
3	0.318	4.471	0.899	0.305
6	0.440	6.498	0.806	-0.300
9	0.344	4.804	0.881	-0.524
11	0.008	0.111	1.000	0.565
15	0.550	8.381	0.697	0.111
21	0.539	8.113	0.709	0.116
25	0.430	6.287	0.815	0.050
31	0.153	2.111	0.977	-0.030
34	0.455	6.745	0.793	0.181
45	0.358	5.026	0.872	-0.300
48	0.496	7.423	0.754	0.327
51	0.473	6.829	0.776	-0.479
64	0.449	6.655	0.798	-0.111
68	0.087	1.185	0.993	0.402

Table 3-22. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.535	8.502	0.714	-0.187
18	0.677	11.983	0.542	-0.269
28	0.674	11.732	0.546	0.005
37	0.471	7.331	0.778	0.342
43	0.166	2.245	0.972	-0.429
50	0.384	5.873	0.853	0.065
55	0.424	6.350	0.820	0.674
61	0.696	12.552	0.516	0.010
OL				
2	0.540	10.764	0.709	-0.161
5	0.463	8.560	0.786	0.085
8	0.414	6.831	0.829	0.650
14	0.353	5.453	0.875	0.745
17	0.643	12.977	0.587	0.650
20	0.603	12.760	0.636	-0.166
23	0.362	5.431	0.869	0.863
27	0.317	5.365	0.899	0.040
32	0.659	13.599	0.566	0.559
35	0.418	6.445	0.825	0.856
38	0.606	12.432	0.633	0.202
40	0.327	5.326	0.893	0.380
44	0.397	6.833	0.842	-0.290
47	0.329	5.368	0.892	-0.440
53	0.292	4.799	0.915	0.085
56	0.359	6.162	0.871	0.342
58	0.573	11.894	0.671	0.060
59	0.328	5.591	0.892	-0.025
60	0.765	18.592	0.414	0.402
63	0.324	5.531	0.895	-0.070
66	0.534	10.532	0.715	0.161
69	0.408	6.975	0.834	0.451

Table 3-23. Parameter estimates of the modified four-factor model in Venezuela data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	0.317	4.752	0.900	-0.156
4	0.558	7.413	0.689	-0.997
7	0.072	0.865	0.995	0.787
13	0.313	4.614	0.902	0.132
16	0.465	7.231	0.784	0.444
19	0.287	4.112	0.917	-0.129
22	0.271	3.803	0.927	0.037
26	0.518	8.377	0.731	0.250
33	0.421	6.532	0.823	-0.380
36	0.438	5.154	0.808	-1.022
39	0.388	5.233	0.850	-0.733
42	0.512	7.147	0.738	-0.861
46	0.365	5.555	0.867	0.138
49	0.314	4.497	0.902	-0.332
52	0.427	6.098	0.818	-0.525
57	0.827	15.439	0.315	-0.931
62	0.482	7.459	0.768	-0.402
65	0.106	1.441	0.989	0.408
67	0.418	6.552	0.825	-0.260
PM				
3	0.437	5.735	0.809	0.285
6	0.399	4.189	0.841	-1.011
9	0.033	0.359	0.999	-0.674
11	0.237	2.748	0.891	-0.268
15	0.134	1.607	0.982	-0.339
21	0.275	3.252	0.925	-0.623
25	0.208	2.545	0.957	-0.288
31	0.249	3.234	0.938	0.034
34	-0.182	-2.256	0.967	0.073
45	0.091	1.133	0.992	-0.028
48	0.315	3.973	0.901	-0.272
51	0.153	1.787	0.976	-0.625
64	0.087	1.069	0.992	-0.289
68	-0.152	-1.863	0.977	0.971

Table 3-23. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.330	3.735	0.891	-0.268
18	0.416	4.779	0.827	-0.555
28	0.679	7.01	0.539	0.003
37	0.108	0.856	0.988	1.370
43	0.117	1.293	0.986	-0.203
50	0.151	1.613	0.977	0.229
55	0.483	4.611	0.767	0.676
61	0.487	5.527	0.763	0.125
OL				
2	0.300	4.691	0.910	0.119
5	0.857	25.896	0.266	0.485
8	0.357	5.377	0.872	0.587
14	0.090	1.074	0.992	0.981
17	0.680	13.419	0.538	0.951
20	0.589	11.302	0.654	-0.018
23	0.270	3.949	0.927	0.680
27	0.293	4.415	0.910	-0.442
32	0.704	14.269	0.504	0.895
35	0.142	1.89	0.980	0.921
38	0.426	6.656	0.818	-0.260
40	0.172	2.501	0.971	0.307
44	0.441	7.545	0.806	-0.067
47	0.591	11.344	0.651	-0.162
53	0.405	6.503	0.836	0.340
56	0.424	6.772	0.821	0.490
58	0.481	8.038	0.768	0.667
59	0.259	3.801	0.933	-0.451
60	0.818	22.63	0.332	0.510
63	0.425	7.142	0.819	0.179
66	0.674	13.225	0.545	0.743
69	0.572	10.729	0.672	0.490

Table 3-24. Parameter estimates of the modified four-factor model in Zimbabwe data

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
EI				
1	-0.100	-1.284	0.990	0.296
4	0.418	5.057	0.825	-0.782
7	0.264	3.228	0.930	0.670
13	0.143	1.818	0.979	0.416
16	0.104	1.358	0.989	-0.028
19	0.151	1.949	0.977	-0.323
22	0.114	1.472	0.987	-0.067
26	0.162	2.161	0.974	0.033
33	0.029	0.363	0.999	-0.342
36	0.288	3.617	0.917	-0.406
39	0.285	3.783	0.919	-0.154
42	0.323	4.164	0.896	-0.412
46	0.047	0.593	0.998	0.269
49	0.163	2.146	0.974	-0.157
52	0.217	2.774	0.953	0.191
57	0.612	8.838	0.625	-0.480
62	0.562	7.875	0.684	-0.482
65	-0.001	-0.010	0.999	0.202
67	0.219	2.553	0.952	0.746
PM				
3	0.226	3.200	0.949	-0.294
6	0.243	2.946	0.941	-0.813
9	0.343	4.562	0.882	-0.710
11	-0.270	-3.288	0.927	0.783
15	0.084	1.147	0.993	-0.152
21	0.114	1.564	0.987	-0.236
25	0.278	4.005	0.923	-0.246
31	0.502	6.967	0.748	-0.727
34	0.231	3.267	0.947	-0.036
45	0.300	4.247	0.910	-0.416
48	0.421	6.211	0.823	-0.519
51	0.340	4.609	0.885	-0.501
64	0.300	4.047	0.916	-0.235
68	0.423	6.129	0.821	0.431

Table 3-24. Continued

Trait/Item	Factor loading	Estimate/SE	Residual Variance	Threshold
TF				
12	0.194	1.897	0.962	-0.227
18	0.431	4.284	0.814	-0.495
28	0.390	3.832	0.848	-0.350
37	0.221	2.024	0.951	0.505
43	0.116	1.148	0.987	-0.196
50	0.218	2.117	0.953	-0.350
55	0.409	3.759	0.833	0.584
61	-0.003	-0.034	0.999	-0.074
OL				
2	-0.010	-0.149	0.999	-0.275
5	0.123	1.832	0.985	-0.183
8	0.380	4.619	0.856	1.121
14	0.300	3.456	0.915	0.910
17	0.687	8.529	0.527	1.579
20	0.485	7.235	0.765	0.811
23	0.694	11.235	0.518	1.015
27	0.324	4.908	0.895	0.461
32	0.450	6.359	0.797	1.031
35	0.493	7.234	0.757	0.922
38	0.641	10.835	0.589	0.989
40	0.128	1.891	0.984	0.204
44	0.563	9.773	0.683	0.768
47	0.388	6.254	0.849	0.472
53	0.311	4.875	0.904	0.333
56	0.062	0.936	0.996	0.232
58	0.512	8.237	0.738	0.826
59	-0.054	-0.807	0.997	0.113
60	0.594	9.981	0.647	0.925
63	0.443	7.407	0.803	0.425
66	0.585	9.683	0.657	0.916
69	0.370	5.333	0.863	0.790

CHAPTER 4
DISCUSSION

The purpose of this study was to examine cross-nationally the four-factor structure of temperament as measured by the SSQ-- a somewhat popular measure of children's temperament used internationally. Since SSQ's development in 1996, results of the measure's psychometric properties have been supported largely by U.S. data. Although SSQ data on children from other countries are available, there have been few attempts to examine their factor structure.

The validity of SSQ is grounded on both theoretical and empirical evidences. The development of SSQ relies on a well-established psychological type theory that was introduced by Jung (1971/1921) and extended by Briggs and Myers (MBTI: 1976). This theory defined the framework of SSQ as a measure of four bipolar temperament traits.

As to SSQ's internal structure, the four-factor model of temperament has been consistently supported through exploratory and confirmatory factor analyses (Callueng, de Carvalho, Isobe, & Oakland, 2012; Benson, Oakland, & Shermis, 2009; Strafford & Oakland, 1996; Oakland, Glutting, & Horton, 1996). Validity of the SSQ is further supported by other sources. At the item level, response patterns generally were similar among Hispanic, African American, and White students (Stafford, 1994). External validation procedures indicated that temperament preferences may influence career choices, class preferences, and involvement in school activities (Oakland, Glutting, & Horton, 1996). Temperament preferences are associated to personal values and similar traits measured by the MBTI. Temperament preferences are independent of cognitive traits, including achievement and intelligence (Oakland, Glutting, & Horton, 1996).

Prior studies on the SSQ's factor structure used factor analysis (e.g. Benson, Oakland, & Shermis, 2009) and Rasch measurement modeling (Mpofu, Oakland, & Gwirayi, 2010) that reported stability of the four-factor model. The current study attempted to extend the validity evidence of the four-factor model by applying factor analytic procedures at the item level. Additionally, this study is a large scale cross-national investigation that includes data from 17,867 children in 21 countries that represent major world regions. This large data has implications for the use of chi-square (χ^2) statistics.

Overall model fit of the data from each country was assessed using root mean square error approximation (RMSEA) as primary index of fit, and comparative fit index (CFI) and Tucker-Lewis index (TLI) as secondary indices of fit. Although χ^2 values also are reported, they are not considered to be a primary fit criterion because they tend to be less useful with larger sample size (Hopwood & Bonnelan, 2010).

General findings from confirmatory factor analyses (CFA), using data from each country, revealed RMSEA values that were within acceptable limits (i.e., ≤ .06), thus suggesting a good fit of the four-factor model of temperament. However, all CFI and TLI indices failed to meet the cut-offs (i.e., ≤ .90) and thus suggesting a poor fit to the model in both the initial and modified solutions. As expected, the χ^2 values did not confirm the hypothesized factor model. Taken as a whole, the overall fit indices indicate that the data from the 21 countries did not fit the four-factor model of temperament reasonably well. Consequently, a test of invariance was not implemented.

The inadequate fit of the four-factor model of temperament in the 21 countries is not consistent with previous findings (Callueng, de Carvalho, Isobe, & Oakland, 2012;

Benson, Oakland, & Shermis, 2009; Strafford & Oakland, 1996; Oakland, Glutting, & Horton, 1996). They reported a good fit of the four-factor model in data from children in the U.S. as well as children in other countries. The discrepancy between the findings of the current study and the earlier studies may be attributed to the use of different methodologies to assess the test's factor structure. Previous studies used item parcels as indicators of extroverted-introverted (EI), practical-imaginative (PM), thinking-feeling (TF), and organized-flexible (OL) traits by combining single dichotomous items into parcels in order to produce greater variance (Brown, 2006; Zwick, 1987). The use of item parceling supports the four-factor model of temperament. Moreover, the bipolar traits were non-overlapping in data from children in the U.S. (Oakland, Glutting, & Horton, 1996; Strafford & Oakland, 1996) and children from other countries (Callueng, de Carvalho, Isobe, & Oakland, 2012; Benson, Oakland, & Shermis, 2009).

In contrast, the current study employed item-level factor analysis for categorical variable methodology (CVM). Some believe this analysis may be more appropriate for dichotomous test items (Bandalos, 2008; Ivanova et. al., 2007; Yu & Muthen, 2002).

However, factor analysis for CVM is robust and stable when dealing with a maximum of 25-30 variables (Muthen & Kaplan, 1992). The 63 SSQ items constitute 63 non-overlapping variables used in the CFA. Thus, this figure exceeds the limits set for the use of a factor analysis with CVM. This large number of variables used as indicators of a relatively small number of latent factors has been unable to meet the requirements for model fit with other data sets. Thus, this issue is not unique to SSQ data. For example, similar findings are reported in a study that examined the factor structure of the Myers-Briggs Type Indicator, also a measure of four bipolar temperaments albeit in adults (Bess, Harvey, & Swartz, 2003). Findings and the results from simulation work on CVM (Flora & Curran, 2004) when used with robust approaches to evaluate fit indices of models with categorical and non-normal manifest variables suggest item parceling may be an acceptable alternative strategy (Brown, 2006). Item parceling seemingly works well with SSQ data and may be more suitable than CVM that is prone to model misspecifications when used with large number of items for a simple factor structure.

Moreover, there is little support for the accuracy and stability of fit indices for factor solution estimates when using the mean and variance-adjusted weighted least-squares estimator (WLSMV) (Bauducel & Herzberg, 2009). More specifically, use of CFI and TLI as overall fit indices in CFA for categorical data is found to be less stable and robust compared to RMSEA (Yu and Muthen, 2002). Following this cautionary note, the current study and other studies have considered CFI and TLI as secondary fit indices to RMSEA (e.g., Ivanova et. al., 2007).

Aside from overall fit indices, item factor loadings impact model fit. Item reliability in turn, influences factor loadings (Nunnally & Berstein, 1994; Gorsuch, 1983). The findings of the current study indicate considerable variability in the number of items that have appreciable loadings (i.e., ≥ .30) among the countries, ranging from 31 (Zimbabwe) to 54 (U.S.), with an average of 43 item (68%) indicators distributed across the four temperament traits. Items that measure EI and OL traits have higher factor loadings than those that measure PM and TF traits. Factor loadings of items measuring EI and OL have been enhanced by their higher reliability coefficients. Conversely, lower factor loadings for items that measure PM and TF have been attenuated by their lower reliability coefficients.

Factor intercorrelations provide important data to assist us when evaluating the fit of the four-factor model of temperament. Previous studies reported that the bipolar traits measured in the SSQ are not overlapping and thus provide evidence that they are somewhat distinct constructs (Benson, Oakland, & Shermis, 2009; Strafford & Oakland, 1996; Oakland, Glutting, & Horton, 1996). In contrast, the current study found mixed results on whether the bipolar traits overlap. In general, data from 15 (71%) of the 21 countries displayed independence of the bipolar traits and the remaining 6 (29%) countries (i.e., China, Egypt, Gaza, Hungary, Iran, & Samoa) displayed multicollinearity and thus overlap on what the bipolar traits measure. Furthermore, multicollinearity in China, Egypt, and Gaza was severe, resulting to a non-positive correlation matrix or an improper solution.

Notably, PM and OL are associated more frequently in countries with traditions and value orientations that are consistent with a collectivist culture. Nigeria, Pakistan, Philippines, Romania, Samoa, Venezuela, and Zimbabwe are known for their collectivistic orientations. This finding concurs with the belief that culture may influence children’s temperament.

The non-positive solution displayed in data from Egypt and Gaza may be associated with turbulent situations in these countries at the time the data were gathered as well as test adaptation factors since the SSQ in both countries used the Arabic language.

The non-positive solution displayed in data from China may be associated with some cultural qualities common to Chinese people such as filial piety, harmony, *Ren Qing* (relationship orientation), modernization, thrift vs. extravagance, *Ah-Q Mentality* (defensiveness), and face (Cheung et al., 1996). A failure to find fit of the four-factor model for Chinese children as well as for the big five personality factor model for Chinese adults may be due to similar qualities (Cheung et al., 2001). Culture may shape temperament and personality qualities in Chinese people.

Lastly, the model misspecifications in the CFA solutions may have been influenced by sample sizes. The use of asymptotic distribution free estimator (e.g., the robust weighted least square estimator, WLS) requires an unbiased and large sample sizes (Byrne, 2012). Moreover when using WLSMV, guidelines for sample size requirements have not been fully established (Brown, 2006). Rules of thumb for minimum to moderate (i.e., 5 to 10 participants per indicator) sample sizes may not be sufficient for CFA when using robust WLS estimator (Byrne, 2012). Future studies that examine the factor structure of the SSQ at the individual item level can use power analysis as a viable method to determine the appropriate and adequate sample sizes. Power analysis is statistically defined as 1 minus the probability to commit a Type 2 error, with adequate power set at .80 (Cohen, 1988). Determining a sample size based on adequate power will help ensure precise parameter estimates and generalizability of model specifications (Schmitt, 2011).

Implications

International Research

Interest in multicultural issues has spawned research and testing practices within and between countries (Byrne et. al., 2009). Temperament data utilized in the current study were collected through collaborative efforts of researchers in 21 countries.

As a non-pathological measure of children's temperament, research on the SSQ has added to our understanding of children's temperament preferences together with their development and gender differences. Scholars and practitioners in some countries (e.g., Kuwait, Republic of China, Romania, and South Korea) that have adapted and normed the SSQ use it in research and school-based practices, including classroom learning, instructional design, and career exploration. The current study is a product of well-established international collaboration that was forged through common interest, scholarship, and professional commitment.

Cross-national studies are important when attempting to validate theoretical models. The five factor model of personality (McCrae & Costa, 2000; McCrae & Costa, 1977) and the 8-syndrome taxonomic model for youth psychopathology (Ivanova et. al., 2007) are based on theoretical models that successfully transcended linguistic and cultural differences. In a similar vein, the assessment of temperament as examined in the current study was able to transcend differences in language, religion, ethnicity, and sociopolitical orientation. The use of a common model of temperament facilitates communication and collaboration for educational and mental health professionals from different countries.

Structural and Measurement Equivalence

Structural and measurement equivalence constitutes a primary strand in cross-national research. An array of approaches has been introduced to establish different but complimentary forms of equivalence: construct, functional, linguistic, and metric (Poortinga, 1995; van de Vijver & Tanzer, 1997; van de Vijver & Leung, 2000). Construct equivalence is achieved when there is high degree of factorial structure invariance across groups (Schmitt, Allik, McCrae, & Benet-Martinez, 2007). Functional invariance is attained when factors or scales are associated with other variables in similar ways (Schmitt, Allik, McCrae, & Benet-Martinez, 2007). Linguistic equivalence is demonstrated when the scales or dimensions of a test are internally consistent across languages (Caprara, Barbaranelli, Bermudez, Maslach, & Ruch, 2000). Metric equivalence is achieved when items administered in different languages function in the same way (Ramirez-Esparza, Gosling, Benet-Martinez, Porter, & Pennebaker, 2006).

The SSQ's four-factor model of temperament has been assessed through exploratory factor analysis (Oakland, Glutting, & Horton, 1996), differential item functioning (Strafford & Oakland, 1996), Rasch measurement modeling (Mpofu, Oakland, & Gwirayi, 2010), and confirmatory factor analysis using item parcels (Benson, Oakland, & Shermis, 2009). These analytical strategies confirmed the stability of temperament as constituting four independent bipolar constructs: extraverted-introverted, practical-imaginative, thinking-feeling, and organized-flexible.

The current study examined the factor structure of temperament through the use of factor analysis that focused on dichotomous items. Because of the restricted variance inherent in the use of dichotomous items, the mean and variance-adjusted weighted least-squares (WLSMV) estimation was used along with tetrachoric correlations as input to confirmatory factor analysis. This analytic approach is robust in correcting observed covariances in binary data.

Current methods to assess construct validity of tests with dichotomous items have established the relationship between factor analysis and item response theory (IRT) (Kamata & Bauer, 2008). As shown in the above results, the limited variance inherent in the dichotomous item format of the SSQ may not be viable for the requirements of the CFA. As an alternative, SSQ's factor structure may be examined using IRT, an approach that is more suitable for dichotomous items (Wirth & Edwards, 2007) such as those found on the SSQ. IRT implements a two-parameter logistic model as a likelihood-based estimation model for logistic approximation to the normal ogive (Wirth & Edwards, 2007). The use of threshold as a difficulty index provides important statistical information of the latent response model that takes account of the dichotomous nature of the items (Kamata & Bauer, 2008). When applied to SSQ data, preference for temperament quality (i.e., EI, PM, TF, or OL) is conceptualized as a function of the probability of endorsing an item response that is plotted in an item characteristic curve or trace line.

Test Validity

Test validity, an essential test characteristic, is reflected in its theoretical and empirical support (AERA, APA, & NCME, 1999). Evidences of validity can be obtained through various sources that, in their combined use, can complement each other to provide an accurate judgment of the test's overall validity.

The increasing popularity of SSQ in other countries as evidenced in the current study indicates interest in the possible usefulness of temperament data in promoting self-understanding, counseling, assessing learning styles, instructional matching, exploring vocational needs, and facilitating research and program evaluation.

In addition to factor structure, the validity of the SSQ adapted versions can be enhanced by establishing a network of variables that are associated with temperament (i.e., convergent validity) and the variables that are not associated with temperament (i.e., discriminant validity). Validity evidence of the SSQ adapted versions also can be demonstrated through group comparisons (i.e., gender, age, clinical diagnoses, etc.).

Limitations

Several limitations of this study need to be acknowledged. First, sample sizes varied, ranging from 7,902 for the U.S. to 253 for Israel. DiStefano (2002) suggested using large sample sizes when using the weighted least squares estimator for CFA to ensure accuracy and generalizability of results. In addition, larger sample sizes increase statistical power and precision of parameter estimates in CFA (Schmitt, 2011). Hence, results from countries with relatively small sample sizes (e.g., Israel) may produce biased estimates and are not considered conclusive.

Second, samples in all countries except that of the U.S. were drawn from a specific geographical area and thus do not represent the general population of children in those countries. Whether responses of these are fully representative of children nationally cannot be verified.

Third, the SSQ English version was used to collect temperament data in children from Nigeria, Pakistan, and Singapore. Although children in these countries attended English speaking schools, it is not known whether responses in the English language and in a native language differ for bilingual children.

Fourth, only the four-factor model of temperament was tested in all countries and no competing model was introduced. Other measurement models that have theoretical and substantive support should have been used to determine the best fitting model in as much as the four-factor model of temperament in all the countries did not provide a good fit.

Fifth, a modified solution was performed for the countries to address misspecifications of the CFA model. Modifications in a CFA model are based on theoretical and empirical evidence. Since no previous studies were available to guide model modifications, content meaning at the item level provided the basis on which to carry out a modified solution.

Conclusions

Despite these limitations, results from this study contribute to the literature on the factorial validity of the SSQ. Although item level CFA did not support the cross-national generalization of the four-factor structure of temperament, the results can be considered necessary but not a sufficient gauge of the SSQ's validity. The SSQ's dichotomous item format restricts variance that can affect observed covariances and produce biased CFA estimates. Moreover, the relatively small sample sizes in most countries may not have met the large sample size requirements when using the WLSMV estimator. Thus, attempts to find a model fit using CFA can be problematic. In previous studies that used item parcels in factor analysis, differential item functioning, and IRT provided theoretical support for the four bipolar temperament model. Hence, we can conclude that the lack of support for the factorial structure of the SSQ as reported in this study is tentative and future studies can examine the SSQ's factor structure using larger sample size.

Future studies also may focus on collecting additional empirical evidences on the SSQ's adapted versions. These validity studies can be pattered from research done with children from the U.S. that examine whether there is support on the internal structure and functional validity of the SSQ.

Future research also may examine factorial validity of the SSQ's adapted versions using differential item functioning, item response theory, and factor analysis using item parcels. Utilizing these other methods with data from other countries may find support for the construct validity of the SSQ and add to existing literature.

The four temperament traits are known to function independently. Yet the findings from this study indicate these traits highly overlap in data from China, Egypt, and Gaza. Subsequent research can examine the factor structure of SSQ in these countries using larger sample sizes.

To overcome restricted variance in dichotomous items, SSQ item format could be modified into a continuous scale to eliminate its dichotomous nature. Factor structure of the SSQ with this continuous scale format can be examined to determine if the four-factor model of temperament fits well with this item format.

LIST OF REFERENCES

American Psychological Association, American Educational Research Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Bandalos, D. L. (2008). Is parceling really necessary? A comparison or results from item parceling and categorical variable methodology. *Structural Equation Modeling*, 15, 211-240.

Bandalos, D. L., & Finney, S. J. (2001). Item parceling issues in structural equation modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 269–296). Mahwah, NJ: Lawrence Erlbaum Associates.

Banducel, A. & Herzberg, P. Y. (2009). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling: A Multidisciplinary Jo*urnal, 13 (2), 186-203.

Barkley, R. A. (1997). *Defiant children: A clinician's manual for assessment and parent training*. New York: Guilford Press.

Bassett, K. & Oakland, T. (2009). Temperament preferences for children ages 8 through 17 in a nationally represented sample. In J. Kaufman (Ed.). *Intelligent testing: Integrating psychological theory and clinical practice* (pp. 30-52). Boston: Cambridge University Press.

Beckner, P.A. (1990). *The learning styles of achieving and underachieving gifted and talented tenth grade students*. Unpublished master's thesis, California State University, Sacramento, CA.

Benson, N., Oakland, T., & Shermis, N. (2009) Cross-national invariance of children's temperament. *Journal of Psychoeducational Assessment*, 27, 3-16.

Berry, J.W. Poortinga, Y.H., Segall, M.H., & Dasen, P.R. (1992*). Cross-cultural psychology: Research and applications*. Cambridge: Cambridge University Press.

Bess, T. L., & Harvey, R. J. (2002). Bimodal score distributions and the Myers-Briggs Type Indicator: Fact or artifact? *Journal of Personality Assessment*, 78, 176-186.

Bireley, M. (1991). Learning styles: One way to help gifted adolescents understand and choose life styles. In M. Bireley & J. Genshaft (Eds.), *Understanding the gifted adolescent: Educational, developmental, and multicultural issues* (pp. 189-200). New York: Teacher Collage Press.

Briggs, K. C. & Myers, I. B. (1976). *Myers-Briggs Type Indicator: Form F*. Palo Alto, CA: Consulting Psychologists Press.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford.

Brown, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Buss, A. H. (2011). *Pathways to individuality: Evolution and development of personality traits.* Washington, DC: APA.

Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS-Windows: Basic concepts, applications, and programming.* Thousand Oaks, CA: Sage.

Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York: Routledge.

Byrne, B. M., Oakland, T., Leong, F. T. L., van de Vijver, F. J. R., Hambleton, R. K., Cheung, F. M., et al. (2009). A critical analysis of cross-cultural research and testing practices: Implications for improved education and training. *Training and Education in Professional Psychology*, 3, 94-105.

Byrne, B. M., Shavelson, R. J., & Muthen, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance*. Psychological Bulletin*, 105, 456-466.

Callueng, C. M., de Carvalho, M., Isobe, M., & Oakland, T. (2012). Temperament styles of children from Japan and the United States: A cross-national study. *Educational Measurement and Educational Review*, 3, 3-23.

Callueng, C. M., Lee Hang, D. M., Gonzales, R. C., Ling-So'o, A. C., & Oakland, T.D. (2011). Temperament styles of children from Samoa and the United States. *Educational Measurement and Educational Review*, 2, 18-34.

Caprara, G. V., Barbaranelli, C., Bermúdez, J., Maslach, C., & Ruch, W. (2000). Multivariate methods for the comparison of factor structures. *Journal of Cross-Cultural Psychology,* 31 (4)*,* 437–464.

Cattell, R. B., Cattell, A. K., & Cattell, H. E. P. (1993). *Sixteen Personality Factor Fifth Edition Questionnaire.* Champaign, IL: Institute for Personality and Ability Testing.

Chess, S. & Thomas, A. (1984). *Origins and evolution of behavior disorders*. New York: Brunner/Mazel.

Chess, S. & Thomas, A. (1986). Temperament in clinical practice. New York: Guilford Press.

Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233-255.

Cheung, F. M., Leung, K., Fan, R., Zang, I. X., Sun, H. F., Gan, Y. Q., Song, W. Z., et al. (2001). Indigenous Chinese personality construct: Is the five-factor model complete? *Journal of Cross-Cultural Psychology*, 32, 407-433.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. (2nd ed.). Hillsdale, NJ: Erlbaum.

Cole, M., & Cole, S. R. (1996). *The development of children (3rd ed.).* New York: W. H. Freeman and Company.

Costa, P. T., Jr., & McCrae, R. R. (2001). A theoretical context for adult temperament. In T. D. Wachs & G.A. Kohnstamm, (Eds.), *Temperament in context* (pp. 1–19). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Cross, T.L., Speirs Neumeister, K.I., & Cassidy, J. C. (2007). Psychological types of academically gifted adolescents. *Gifted Child Quarterly*, 51 (3), 285-294.

CTB/McGraw-Hill (1985). *California Achievement Test*. Monterey, CA: Author.

Danielsen, S. (1991). *Personality type preferences of children with and without learning difficulties.* Unpublished master's thesis, University of Alberta, Edmonton.

Darby, L. K. (2009). *Effects of shared personality types of parents and children with high functioning autism spectrum disorders on parental stress.* Unpublished doctoral dissertation, Texas Woman's University, Denton.

Davis, S.E. (2007). *Effects of motivation preferred learning styles, and perception of classroom climate on achievement in ninth and tenth grade math students. Unpublished doctoral dissertation*, University of Florida, Gainesville.

Dempsey, R.D. (1975). *Cognitive style differences between gifted and average students*. Unpublished doctoral dissertation, United States International University, Ann Arbor, MI.

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, 43 (2), 121-149.

DiStefano, C. (2002). The impact of categorization with confirmatory factor analysis. *Structural Equation Modeling,* 9, 327–346.

Eysenck, H. J. & Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach*. New York: Plenum.

Fan, X. & Thompson, B. (2001). Confidence intervals about score reliability coefficients, please: An EPM guidelines editorial. *Educational and Measurement Journal*, 61 (4), 517-531.

Flora, D. B. & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9 (4), 466- 491

Fourgurean, J., Meisgeier, C., & Swank, P. (1988). The Murphy-Meisgeier Type Indicator for Children: Exploring the link between psychological type preferences of children and academic achievement. *Journal of Psychological Type*, 16, 42-46.

Galen (1916). *Galen on the natural forces*. (A. J. Brook, Trans.). Cambridge, MA: Harvard University Press. (Original work published date unknown)

Galen (1992). *The art of cure*-Extracts from Galen: Maimonides' medical writings. (U.S. Barzel, Trans.). Haifa, Israel: Maimonides Research Institute. (Original work published date unknown)

Gallager, S.A. (1990). Personality patterns of the gifted. *Understanding Our Gifted*. 3, 10-13.

Goldsmith, H. H., Buss, A. H., Plomin, R., Rothbart, M. K., Thomas, A., Chess, S., et al. (1987). Roundtable: What is temperament? Four approaches. *Child Development, 58,* 505-529.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Hambleton, R. K. (2002). The next generation of the ITC test translation and adaptation guidelines. *European Journal of Psychological Assessment*, 17 (3), 164-172.

Harrier, L. (2005). *The learning styles of children diagnosed with attention deficit/hyperactivity Disorder (ADHD) as determined by the Student Styles Questionnaire (SSQ)*. Unpublished doctoral dissertation, Texas Woman's University, Denton.

Hergenhahn, B. R. (2001). *An introduction to the history of psychology (4th ed.)*. Belmont, CA. Brooks/Cole.

Hippocrates. (1939). *Hippocrates, 1*. (W.H.S. Jones, Trans). Cambridge, MA: Harvard University Press. (Original work published date unknown).

Hippocrates (1988). *Hippocrates, 5*, (P. Potter, Trans.). Cambridge, MA: Harvard University Press. (Original work published date unknown).

Hippocrates (1994). *Hippocrates, 7*, (W. D. Smith, Trans.). Cambridge, MA: Harvard University Press. (Original work published date unknown).

Hopwood, C. J. & Donnelan, M.B. (2010). How should the internal structure of personality inventories be evaluated?. *Personality and Social Psychology Review*, 14 (3), 332-346.

Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55.

Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Dumenci, L., Almqvist, F., Bilenberg, N., et al. (2007). The generalizability of the Youth Self-Report syndrome structure in 23 societies. *Journal of Consulting and Clinical Psychology*, 75 (5), 729-738.

Jennings, D.K. (2005*). Temperament-based cognitive styles of children with serious Emotional disturbance*. Unpublished doctoral dissertation, Texas Woman's University, Denton.

Joyce, D. (2010). *Essentials of temperament assessment*. New York: Wiley.

Joyce, D. & Oakland, T. (2005). Temperament differences among children with conduct disorder and oppositional defiant disorder. *The California School Psychologist*, 10, 125-136.

Jung, C. G. (1926). *Psychological types (H. G. Baynes, Trans)*. New York: Harcourt, Brace, & Co. (Original work published 1920)

Jung, C. G. (1945). *Contributions to analytical psychology (H. G. Baynes & C. F. Baynes, Trans.)*. New York: Kegan Paul, Trench, Trubner, & Co. LTD. (Original work published 1928)

Jung, C. G. (1971*). Psychological types (R.F.C. Hull, Revision of Trans. by H. G. Baynes)*. Princeton, NJ: Princeton University Press. (Original work published 1921)

Kamata, A. & Bauer, D. J. (2008). A note on the relation between factor analytic and item response theory models*. Structural Equation Modeling: A Multidisciplinary Journal*, 15 (1), 136-153.

Katona, N. & Oakland, T. (2000). The development of temperament in Hungarian children. *Hungarian Journal of Psychology*.1, 17-29.

Kaufman, A. S., McLean, J. E., & Lincoln, A. (1996). The relationship of the Myers-Briggs Type Indicator (MBTI) to IQ level and the fluid and crystallized IQ discrepancy of the Kaufman Adolescent and Adult Intelligence Test (KAIT). *Assessment,* 3 (3), 225-239.

Keirsey, D. (1998). *The Keirsey Temperament Sorter – II*. Del Mar, CA: Prometheus Nemesis.

Keirsey, D., & Bates, M. (1978). *Please understand me: Character and temperament types*. Del Mar, CA: Prometheus Nemesis Book Co.

Leon, C., Oakland, T., Wei, Y., & Berrios, M. (2009) Venezuelan children's temperament styles and comparison with their United States peers. *Revista Interamericana de Psicología/Interamerican Journal of Psychology, 43,* 407-415.

Loehlin, J.C. (1998). *Latent variable models: An introduction to factor, path, and structural analysis* (3rd ed.). Mahwah, NJ: Erlbaum.

Little, T.D., Cunningham, W.A., Sharar, G., & Widaman, K.F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, 9 (2), 151-173.

Marsh, H.W. & Hau, K.T. (1996). Assessing goodness of fit: Is parsimony always desirable? *The Journal of Experimental Education*. 64, 364-390.

McCrae, R. R. (2001). Trait psychology and culture: Exploring intercultural comparisons. *Journal of Personality.* 69, 819-846.

McCrae, R. R., & Costa, P. T., Jr. (1997). Conceptions and correlates of openness to experience. In R. Hogan, J. Johnson, & S. Briggs (Eds.). *Handbook of personality psychology* (pp. 825-847). San Diego, CA: Academic Press.

McGhee, R. L., Ehrler, D. J., & Buckhalt, J. A. (2007). Five Factor Personality Inventory-Children. Austin, TX: Pro-Ed.

Meredith, W. (1993). Measurement equivalence, factor analysis and factorial equivalence. *Psychometrika,* 58, 525–543.

Miller, J. G. (1999). Cultural psychology: Implications for basic psychological theory. *Psychological Science*, 10 (2), 85-91.

Mills, C. J. (1993). Personality, learning style, and cognitive style profile of mathematically talented students. *European Journal of High Ability*, 4, 70-75.

Mpofu, E., Oakland, T., & Gwirayi, P. (2010). Temperament styles of Zimbabwean early adolescents: A Rasch measurement model analysis. *South African Journal of Psychology*, 40 (1), 44-53.

Muthen, B. & Kaplan, D. (1992). A comparison of some methodologies for the factor analysis of non-normal likert variables. *British Journal of Mathematical and Statistical Psychology,* 38, 171-189.

Muthén, L. K., & Muthén, B. O. (2006). *MPlus user's guide version 4.1*. Los Angeles, CA: Muthén & Muthén.

Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *MBTI Manual: A guide to the development and use of the Myers-Briggs Type Indicator (3rd ed.).* Palo Alto, CA: Consulting Psychologists Press.

Myers, I. B. & McCauley, M. H. (1985). *Manual: A guide for the development and use of the Myers-Briggs Type Indicator.* Palo Alto, CA: Consulting Psychologists Press.

Myers, I. B., & Myers, P. B. (1980). *Gifts differing: Understanding personality type*. Palo Alto, CA: Consulting Psychological Press.

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

Oakland, T., Alghorani, M. A., & Lee. D. H. (2006). Temperament-based learning styles of Palestinian and US children. *School Psychology International*, 28, 110-128.

Oakland, T., Banner, D., & Livingston, R. (2000) Temperament-based learning styles of visually impaired children. *Journal of Visual Impairment and Blindness,* January, 26-33

Oakland, T. Faulkner, M., & Bassett, K. (2005). Temperament styles of children from Australia and the United States. *Brazil Educational and Developmental Psychologist,* 19, 35- 51.

Oakland, T., Stafford, M., Horton, C., & Glutting, J. (2001). Temperament and vocational preferences: age, gender, and racial-ethnic comparisons, *Journal of Career Assessment, 9,* 297-314.

Oakland, T. & Hatzichristou, S, (2010) Temperament styles of children from Greece and the United States. School Psychology International, 31(4), 422-437

Oakland, T. Illiescu, D., Dinca, M., & Dempsey, A. (2009). Temperament styles of Romanian children*. Psihologia Sociala*, 22, 70-84.

Oakland, T., Joyce, D., Horton, C. & Glutting J. (2000). Temperament-based learning styles of male and female gifted and non-gifted children. *Gifted Child Quarterly*, 44, 183-189.

Oakland, T., & Lee, D. H. (2010) Temperament styles of children from South Korea and the United States. *School Psychology International*, 77-94.

Oakland, T. & Lu, L. (2006). Temperament styles of children from the People's Republic of China and the United States. *School Psychology International*, 27, 192-208.

Oakland, T., & Mata, A. (2007) Temperament styles of children from Costa Rica and the United States. *Journal of Psychological Type*, 67, 91-102.

Oakland, T., Mogaji, A., & Dempsey, J. (2006). Temperament Styles of Nigerian and U.S. children. *Journal of Psychology in Africa*, 16, 27-34.

Oakland, T., Mpofu, E., & Sulkowski, M. (2007). Temperament styles of Zimbabwe and U.S. children. *Canadian Journal of School Psychology*, 21, 139-153.

Oakland, T. & Pretorius, J. (2009). Temperament styles of children from South Africa and the United States. *School Psychology International*, 29, 627-639.

Oakland, T., Rizwan, M, Aftab, S., & Callueng, C. (2011). Temperament styles of children from Pakistan and the United States. *School Psychology International*, 33 (2), 207-222.

Oakland, T., Singh, K, Callueng, C., & Goen, A. (2010). Temperament styles of Indian and U.S. children. School Psychology International, 32 (6), 655-670.

Parker, L.D. & Robinson, D.C. (1989). Type and academically gifted adolescents. *Journal of Psychological Type*, 17, 66-89.

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.

Pinner, R. & Forlano, G. (1943). Personality tests of partially sighted children. *Journal of Applied Psychology*, 27, 283-287.

Poortinga, Y. (1995). Use of tests across cultures. In T. Oakland & R. K. Hambleton (Eds.), *Interpersonal perspectives on academic assessment* (pp. 187-206). Norwell, MA: Kluwer.

Prior, M., Sanson, A., Smart, D., & Oberklaid, F. (2000). *Pathways from infancy to adolescence*: Brazil temperament project 1983-2000. Melbourne, Australia: Brazil Institute of Family Studies.

Quenk, N. L. (2009). *Essentials of Myers-Briggs Type Indicator Assessment* (2nd ed.). New York: John Wiley & Sons.

Ramírez-Esparza, N., Gosling, S. D., Benet-Martínez, V., Potter, J., & Pennebaker, J. W. (2006). Do bilinguals have two personalities? A special case of cultural frame switching. *Journal of Research in Personality*, 40, 99-120.

Rothbart, M. K. (1989). Temperament and development. In G.A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 187–247). Chichester, UK: Wiley.

Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, 29 (4), 304-321.

Schmitt, D. P., Allik, J., McCrae, R. R., & Benet-Martinez, V. (2007). The geographic distribution of big five personality traits: Patterns and profile of human self-description across 56 nations. *Journal of Cross-Cultural Psychology*, 38 (2), 173-212.

Schuerger, J. M. (2001). *16PF Adolescent Personality Questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.

Storr, A. (1991). *Jung*. New York: Routledge.

Strafford, M. E. (1994). *Ethnic differences in responding to the student temperament assessment record*. Unpublished doctoral dissertation. The University of Texas, Austin.

Stafford, M. E. & Oakland, T. (1996). Racial-ethnic comparisons of temperament constructs for three age groups using the Student Styles Questionnaire. *Measurement and Evaluation in Counseling and Development, 19* (2), 100-110.

Teglasi, H. (1998a). Introduction to the mini-series: Implications of temperament for the practice of school psychology. *School Psychology Review*, 27, 475-478.

Teglasi, H. (1998b). Temperament constructs and measures. *School Psychology Review*, 27, 564-585.

Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.

Thomas, A., & Chess, S. (1989). *Temperament and personality*. In G.A. Kohnstamm, J.E. Bates, & M. K. Rothbart (Eds.), Temperament in childhood (pp. 249–261). New York: Wiley.

Tobacyk, J. J., Hearn, R. E., & Wells, D. H. (1990). Jungian type and California Achievement Test performance in junior high school students at high risk for dropout. *Journal of Psychological Type*, 19, 13-20.

Urbina, S. (2004). *Essentials of psychological testing*. Hoboken, NJ:John Wiley & Sons.

Vandenberg, R.J. & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4-70.

van de Vijver, F. J. R. & Leung, K. (2000). *Methodological issues in psychological research*. Thousand Oaks, CA: Sage.

van de Vijver, F. J. R. & Tanzer, N. K. (1997). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, 47, 263-279.

Wachs, & G. A. Kohnstamm (Eds.). *Temperament in Context.* Mahwah, NJ: Lawrence Erlbaum Associates.

Wehr, G. (1971). *Portrait of Jung: An illustrated biography*. (W. A. Hargreaves, Trans.). New York: Herder and Herder. (Original work published 1969).

Wirth, R. J. & Edwards, M. C. (2007). Item factor analysis: Current approaches and future directions. *Psychological Methods*, 12 (1), 58-79.

Yang, C., Nay, S., Hoyle, R. H. (2010). Three approaches to using lengthy ordinal scales in structural equation models: Parceling, latent scoring, and shortening scales. *Applied Psychological Measurement*, 34 (2), 122-142.

Yu, C.Y. & Muthen, B.O. (2002). *Evaluation of model fit indices for latent variable models with categorical and continuous outcome*s (Technical Report). Los Angeles: UCLA, Graduate School of Education and Information Studies.

Zafran, H.A.S. (1965). A study of personality differences between blind and sighted children. *British Journal of Education Psychology*, 35, 329-338.

Zwick, R. (1987). Assessment of dimensionality of year 15 reading data. In A.E Beaton (Ed.), *Implementing the new design: The NAEP 1983-1984 technical report.* Princeton, NJ: Educational Testing Service.

BIOGRAPHICAL SKETCH

Carmelo M. Callueng is a native of the Philippines. He obtained his Bachelor of Science in Psychology at St. Paul University Philippines (Cum Laude) in 1988 and a Master of Science in Measurement and Evaluation at De La Salle University-Manila in 1993. Before his admission to the School Psychology Program at the University of Florida, Carmelo served as a faculty member and held administrative offices, including academic chair, in the Department of Psychology at De La Salle University-Manila. He also served as principal investigator or co-investigator in externally funded grants that included a project sponsored by the World Bank on Philippine elementary education as well as other supported research on corporate values of Asian employees in a multi-national company and cognitive styles of Filipino students. He also held an elected position on the Philippine Psychological Association's executive committee.

Carmelo completed his internship in Psychology in the academic year 2011-2012 at the Shands Hospital's Behavioral Health Unit and at the Southeastern Health Psychology in Gainesville, Florida. He anticipates receiving his Master in Education and Doctor of Philosophy in School Psychology in 2012 from the University of Florida.

While at the University of Florida, Carmelo authored or co-authored articles in quality peer-reviewed journals and presented several papers at national and international conferences. He holds memberships in the American Psychological Association, National Association of School Psychologists, Florida Association of School Psychologists, International Association of School Psychology, International Association of Cross-Cultural Psychology, and the Psychological Association of the Philippines.

He has received the following research-related awards/scholarships: American Psychological Association Division 33 Research Excellence Award (2009 and 2012), the American Academy of School Psychology Irwin Hyman/Nadine Lambert Memorial Scholarship (2011), the Florida Association of School Psychology Doctoral Graduate Studies Award (2011), the University of Florida College of Education Graduate Research Award (2012), International School Psychology Association Cal Catterall Award (2012), American Psychological Association Division 42 Graduate Research Recognition Award (2012), and the American Psychological Association Division 52 Student Poster Award (2012).