BROCK UNIVERSITY LIBRARY

3 9157 00834744 8

Beyond g: Putting Multiple Intelligences Theory to the Test

by

Beth Visser

A thesis
submitted in partial fulfillment
of the requirements for the degree
Master of Arts

Department of Psychology
BROCK UNIVERSITY
St. Catharines, Ontario

August, 2005

©Beth Visser, 2005

**JAMES A GIBSON LIBRARY
BROCK UNIVERSITY
ST. CATHARINES ON**

Abstract

The purpose of this study was to investigate Howard Gardner's (1983) Multiple Intelligences theory, which proposes that there are eight independent intelligences: Linguistic, Spatial, Logical/Mathematical, Interpersonal, Intrapersonal, Naturalistic, Bodily-Kinesthetic, and Musical. To explore Gardner's theory, two measures of each ability area were administered to 200 participants. Each participant also completed a measure of general cognitive ability, a personality inventory, an ability self-rating scale, and an ability self-report questionnaire. Nonverbal measures were included for most intelligence domains, and a wide range of content was sampled in Gardner's domains. Results showed that all tests of purely cognitive abilities were significantly correlated with the measure of general cognitive ability, whereas Musical, Bodily-Kinesthetic, and one of the Intrapersonal measures were not. Contrary to what Multiple Intelligences theory would seem to predict, correlations among the tests revealed a positive manifold and factor analysis indicated a large factor of general intelligence, with a mathematical reasoning test and a classification task from the Naturalistic domain having the highest *g*-loadings. There were only minor sex differences in performance on the ability tests. Participants' self-estimates of ability were significantly and positively correlated with actual performance in some, but not all, intelligences. With regard to personality, a hypothesized association between Openness to Experience and crystallized intelligence was supported. The implications of the findings in regards to the nature of mental abilities were discussed, and recommendations for further research were made.

Acknowledgements

This thesis came about with tremendous support from my advisor, Dr. Michael Ashton, who provided advice on grammar, statistics, organization, and the importance of parsimony. He was a constant source of calm support throughout this process, from research design and data collection to reporting results and making conclusions.

My committee, Dr. Drew Dane and Dr. Nancy DeCourville, provided valuable input into this project, both at the planning stage, and in the drafting of a final document.

The author wishes to acknowledge support from SSHRC for the collection of data in this study.

Rob took on far more than his share to make this thesis happen, and Craig, Ian, and Sean were always there to remind me that the best insights sometimes come after a day off.

TABLE OF CONTENTS

ABSTRACT	2
ACKNOWLEDGEMENTS	3
LIST OF TABLES	6
INTRODUCTION	8
Multiple Intelligences	9
Assessment of Multiple Intelligences	14
The Intelligences.....	17
Linguistic	19
Spatial	18
Logical/Mathematical	20
Interpersonal	21
Intrapersonal	23
Naturalistic.....	24
Bodily-Kinesthetic	25
Musical	26
Self-Estimated Multiple Intelligences.....	28
Personality and Intelligence	30
Sex Differences.....	32
Multiple Intelligences	32
Self-Estimated Intelligences.....	34
Personality	34
The Current Study.....	36
Hypotheses.....	37
Measured Multiple Intelligences.....	37
Self-Estimated Ability.....	38
Personality	38
Sex Differences.....	38
METHOD	40
Participants	40
Measures.....	40
Demographic Information.....	40
General Cognitive Ability.....	40
Linguistic Intelligence.....	42
Spatial Intelligence.....	43
Logical/Mathematical	44
Interpersonal Intelligence.....	44
Intrapersonal Intelligence.....	45
Naturalistic Intelligence	46
Bodily-Kinesthetic Intelligence.....	46
Musical Intelligence.....	47
Self-Estimates of Ability.....	49
Personality	50
Procedure.....	50

TABLE OF CONTENTS (Cont.)

RESULTS & DISCUSSION	53
Missing Data	53
Measured Multiple Intelligences.....	56
Descriptive Statistics.....	56
Reliabilities	59
Measured Abilities and <i>g</i>	61
Coherence of Abilities	64
Self-Estimated Intelligences	67
Descriptive Statistics.....	67
Reliabilities	70
Within-Instrument Correlations	70
Correlations Between Direct Self-Ratings and MAQ Scales	73
Correlations Between Self-Estimated Abilities & Measured Multiple Intelligences...	74
Personality	79
Descriptive Statistics.....	79
Reliabilities	79
Correlations Among Personality Factors	80
Correlations with Measured Multiple Intelligences	82
Correlations with Self-Estimated Multiple Intelligences.....	84
Sex Differences.....	85
Measured Multiple Intelligences.....	85
Self-Estimated Multiple Intelligences	87
Personality	88
GENERAL DISCUSSION	89
Measured Multiple Intelligences.....	89
Self-Estimated Multiple Intelligences.....	93
Personality	96
Sex Differences.....	97
CONCLUSIONS	98
REFERENCES	100
APPENDICES	108
A. Brock University Ethics Approval	108
B. Multiple Ability Questionnaire.....	110
C. Self-Estimated Ability Inventory	113
D. Personality Self-Rating Form	115
E. Descriptive Statistics of Measures of Each Intelligence Domain and <i>g</i> Prior to Expectation Maximization.....	117

LIST OF TABLES

Table	Page
1. Demographic Data	41
2. Percentage of Missing Data by Variable	54
3. Descriptive Statistics of Measures of Each Intelligence Domain and g	57
4. Split-Half Reliability Coefficients of Ability Tests	60
5. g Loadings of Tests and Correlations of Tests with Wonderlic Personnel Test	62
6. Correlations Among Ability Measures	65
7. Residual Correlations within Each Intelligence Domain after Extraction of g Factor	66
8. Descriptive Statistics of Direct Self-Ratings and MAQ Self-Estimates	68
9. Internal Consistency Reliabilities of Multiple Ability Questionnaire Intelligence Domains	69
10. Within-Instrument Correlations for Ability Self-Ratings and Multiple Ability Questionnaire Scales	71
11. Loadings of Direct Self-Ratings of Multiple Intelligences in First Unrotated Factor	72
12. Loadings of Multiple Ability Questionnaire Scales in First Unrotated Factor	73
13. Correlations Between Ability Self-Ratings and Multiple Ability Questionnaire Scales	74
14. Correlations Between Self-Ratings and Multiple Ability Questionnaire Scales	76
15. Correlations Between Multiple Ability Questionnaire Scales and Intelligence Measures	78
16. Descriptive Statistics and Internal Consistency Reliabilities of Personality Self-Rating Scales	79

17. Correlations Between Self-Reported Personality Scales	80
18. Correlations Between Self-Reported Personality Scales and Ability Measures and Academic Average	81
19. Correlations Between Direct (Self-Rating) and Indirect (MAQ) Self-Estimated Multiple Intelligences and Personality	83
20. Means and Standard Deviations for Each Ability Test by Sex and Results of <i>t</i> -tests of Sex Differences for Each Variable	85
21. Means and standard deviations for Direct Self-Ratings and Indirect MAQ Self-Estimates by Sex, and Results of <i>t</i> -tests of Sex Differences for Each Variable	86
22. Means and Standard Deviations for Self-Reported Personality by Sex and Results of <i>t</i> -tests of Sex Differences for Each Variable	87

Beyond *g*: Putting Multiple Intelligences Theory to the Test

In 1904, Charles Spearman set out to investigate the structure of mental abilities and discovered that correlations between several ability measures, including visualization, vocabulary, mathematical ability, and judging musical pitch, were always positive. Spearman referred to this phenomenon as a positive manifold, and his conclusion that it was attributable to a factor of general intelligence (*g*) underlying all cognitive tasks has stood the test of time. Intelligence testing and statistical methods have been refined immensely since Spearman's day, but the existence of *g* remains fundamental to many modern theories of the structure of mental abilities.

Louis Thurstone (1938), on the other hand, suggested that intelligence was best understood in terms of group factors that were relatively independent of each other. Thurstone posited nine primary mental abilities: verbal relations, perceptual speed, spatial ability, numerical facility, word fluency, memory, induction, deduction, and restriction. He suggested that it was the combination of these primary abilities that would determine an individual's performance on any complex cognitive task.

Hierarchical views of mental abilities (e.g., Vernon, 1961) dominate the intelligence literature today. These theories, grounded in empirical support, have successfully incorporated the findings of both Spearman and Thurstone, by including group factors in addition to *g*. While the number of levels can vary from one hierarchical model to another, *g* is always at the top of the hierarchy, and narrow, specific factors at the bottom. Vernon's model was highly complex and included major group factors of verbal/educational and mechanical/spatial, yet the author maintained that most variance in human abilities in daily life could be attributed to *g*.

In 1993, Carroll put forth a “three-stratum model” based on his factor analysis of 461 data sets. His first stratum, or level, consists of about 60 narrowly-defined primary factors, such as visualization, visual memory, deductive reasoning, and recollection of tunes. Correlations among these first-stratum factors produce a set of eight second-stratum factors: fluid intelligence, crystallized intelligence, general memory and learning, broad visual perception, broad auditory perception, broad retrieval ability, broad cognitive speediness, and processing speed. These abilities are positively inter-correlated, but they reflect the fact that individuals vary somewhat in their cognitive strengths and weaknesses. When these second-stratum factors are analyzed, a single third-stratum factor of general intelligence emerges.

Multiple Intelligences

The notion of general intelligence had been broadly accepted by psychologists when Howard Gardner introduced Multiple Intelligences (MI) theory in his 1983 book, *Frames of Mind*, proposing that there are several independent ability areas. Gardner (1993) described intelligence as a biopsychological potential that could be influenced by experience, culture, and motivational factors. He defined intelligence as the ability to solve problems, or to fashion products, that are culturally valued. Gardner (1983) initially proposed that there were seven intelligences: Linguistic, Spatial, Logical/Mathematical, Interpersonal, Intrapersonal, Bodily-Kinesthetic, and Musical. He has more recently (Gardner, 1999) added Naturalistic intelligence and has suggested that an existential intelligence might exist, but that a hypothesized spiritual intelligence does not.

Gardner (1999) has stated that his choice of the word “intelligences” was a

deliberate one, noting that if he had written a book referring to “faculties” or “gifts,” it is unlikely that his theory would have garnered the attention that it has. Gardner has professed to be quite willing to refer to his eight intelligences as talents or abilities, but only if verbal and quantitative abilities are referred to as talents. Gardner has argued that there is no hierarchy of ability, and that verbal and logical/mathematical intelligence are of no greater real-life importance than any of the other intelligences.

Gardner (1999) noted that he reviewed hundreds of studies before publishing *Frames of Mind*, and that he assessed all candidate intelligences on the basis of eight criteria: the potential of isolation by brain damage; an evolutionary history and evolutionary plausibility; an identifiable core operation or set of operations; susceptibility to encoding in a symbol system; a distinct developmental history; the existence of savants, prodigies, and other exceptional people; support from experimental psychological tasks; and support from psychometric findings. Gardner (1983) noted that the criteria were somewhat flexible; some intelligences which might have met all criteria, such as face-recognition, were discarded because they were not culturally highly-valued. In addition, Gardner retained candidate intelligences which seemed promising even if they did not meet every criterion.

MI theory was enthusiastically welcomed by educators and parents. Here, Gardner’s message that children have unique and diverse abilities meshed well with educators’ intuitive sense that children learn in very different ways. Indeed, Gardner’s premise seemed far more egalitarian than did notions of *g*. MI theory seemed to say that a child who performed poorly in math and reading had just as great a chance as anyone else at being successful in music, art, physical education, or even geography: everyone is

smart in some way. At a time when standardized testing in U.S. schools had become highly controversial, Gardner claimed that intelligence tests measured nothing but verbal and logical/mathematical ability, the two intelligence domains most valued in school. Whereas IQ tests could predict academic success, he insisted that they had little utility in predicting success beyond school (1983, p. 3). This claim, however, is clearly contradicted by Gottfredson's (2002) finding that no meta-analysis has reported exceptions to the generality of *g* in predicting job performance, with brighter employees always performing better on average than less intelligent employees.

Although the educational community has embraced MI theory, psychologists have accused the theory of lacking scientific rigor (Alix, 2000; Klein, 1998). Gardner (Gardner & Connell, 2000) responded to this accusation by claiming that social scientists, unlike physical scientists, do not set out to create deductive and predictive theories. Instead, Gardner noted that he introduced MI theory with tentative language, aware of the fact that it would be unable to account for all phenomena. Similarly, Gardner (1999) has elsewhere identified his theory as empirically-based, and thus, fallible in the light of new findings: "The questions to ask of a new theory are whether it stimulates important questions and investigations, and whether its initial delineation of factors appears to be on the right track" (p. 86). Chen (2004) maintained that MI theory qualified as scientific theory, but not because it met any particular set of scientific principles. Chen argued that first, intelligence is not a tangible, measurable object, and that theoretical hypotheses and methodologies will differ from those of physical sciences; and second, that the theory's value is in its contribution to educational understanding and practice.

While Gardner (1999) now acknowledges the existence of *g*, he continues to

question its explanatory power. In addition, he continues to maintain that each intelligence has unique processing resources, and that there are no horizontal capacities, such as memory or creativity, that cut across all intelligences. Instead, he views creativity as an operation performed within a domain, rather than a general, cross-cutting ability. Gardner (1999) has stated that executive functioning likely emerges from Intrapersonal intelligence rather than constituting an intelligence of its own or a horizontal capacity. While stating that he has no objection to others invoking an executive function, Gardner said that for the purposes of modeling mental abilities, “it is useful to see whether one can explain human behavior in the absence of such hierarchical considerations, or whether the hierarchy can emerge naturally, as part of everyday functioning, rather than by invoking a separate executive intelligence” (p. 106).

Gardner’s acknowledgement that *g* can exist, along with his argument that *g* does not compromise MI theory, has led to criticism that the theory lacks falsifiability (e.g., Klein, 1998). In his review of *Frames of Mind*, Messick (1992) stated that Gardner had selectively drawn upon research that supported his view while ignoring or downplaying results of factor analyses that are unexplained by MI theory. In response to accusations that Multiple Intelligences theory could not be disproved, Gardner (1993, p. 38) stated that there were several experiments which could disconfirm his theory, including tests of the universality of intelligences, the developmental stability of the intelligences, or the independence of the intelligences. In the same publication, however, Gardner claimed that it was not essential to his theory that intelligences be entirely independent (1993, p. 41-42) and Torff and Gardner (1999) said of MI theory, “It is not the kind of theory that can be proved or disproved by crucial experiment” (p. 140). Klein (1998) has suggested

that the problem of falsifiability is compounded by Gardner's circular definitions of his intelligences: "Why are some people good at composing or interpreting music? Because they have high musical intelligence. What is musical intelligence? It is the ability to compose and interpret music" (p. 107).

Carroll (1993) has pointed out that Gardner's intelligences bear a striking similarity to the second-stratum factors of his hierarchy. For example, Carroll noted that Linguistic intelligence corresponded to the factor of crystallized intelligence, Musical intelligence to auditory perception ability, Logical/Mathematical intelligence to fluid intelligence, and Spatial intelligence to visual perception. Interpersonal intelligence, said Carroll, was represented to some extent in one of the knowledge factors, knowledge of behavioral content. Carroll stated that only Bodily-Kinesthetic and Intrapersonal intelligences appeared to have no counterpart in second-stratum factors. However, psychomotor ability is not typically recognized as an important component of cognitive ability and, thus, Bodily-Kinesthetic intelligence would not be represented in Carroll's model. Carroll noted that adequate measures of intrapersonal ability have never been included in factor analytic studies of cognitive structure. Gardner had not introduced his eighth ability, Naturalistic intelligence, at the time of Carroll's writing, but the categorization of objects would seem to be related to logical reasoning.

Carroll (1993) has argued that a human ability is revealed by differences in levels of task difficulty on a defined class of tasks. The broader the class of tasks, the more general the ability is. Presumably, the converse would be true, that performance on a narrower class of tasks would be better predicted by a more specific ability that is conceptually relevant to the same task. Gardner's Musical intelligence should, thus, be a

better predictor of musical performance than would g , even if musical ability is itself g -loaded. In contrast with this “specific is better” argument is the abundance of evidence that g is a good predictor of job performance. For many years, the pervading belief among personnel psychologists was that different specific aptitudes and abilities were required to predict performance in different jobs, but the accumulation of evidence has shown that g is an important predictor of performance in any job (Humphreys, 1986).

While the eight intelligence domains are theoretically independent, Gardner (1993) has acknowledged that two or more could overlap. He cautioned, however, that such correlations would have to be found and replicated across cultures before they could be assumed to be valid, rather than the result of statistical or cultural artifact. He has speculated that correlations among subtests of standardized intelligence tests occur because the tasks all rely on rapid responses to items which are heavily based on logical/mathematical and linguistic abilities. However, Messick (1992) noted that variability in reading ability should not influence intelligence test performance as long as all participants are able to easily understand the task instructions. While Gardner expressed concern about the “verbal lens”, Messick claimed that the reasoning component of Logical/Mathematical intelligence is far more ubiquitous. Messick noted that the reasoning that a person employs to solve a novel task in a different intelligence domain appears more similar to a horizontal, cross-cutting ability than to method variance.

Assessment of Multiple Intelligences

Gardner (1999) has claimed that Linguistic and Logical/Mathematical

intelligences are the domains most highly valued in school and most heavily represented on traditional intelligence tests. The challenge, he has said, is to develop assessment tools that are “intelligence fair”; that is, measures that do not assess intelligence through the lens of verbal, or any other, ability. A difficulty with the measurement of the eight intelligence domains is that Gardner believes that not only should assessment be intelligence-fair, but that it should reflect an individual’s success in completing culturally-valued tasks that rely on particular intelligences (Gardner, 1999), rather than the completion of paper-and-pencil tests. Given Gardner’s (1999) contention that the performance of real-life activities necessarily employs the combination of two or more intelligences, assessment would seem to be an impossible task. In addition, Gardner has emphasized the importance of having tests administered with familiar materials. For example, Gardner (1999) suggested that Musical intelligence could be tested by exposing a person to a new melody, then exploring how quickly they learn to sing, recognize, and transform it. He has also noted that in working with preschoolers, he has had children take apart and reassemble pencil sharpeners and doorknobs to assess Logical/Mathematical, Spatial, and Bodily-Kinesthetic intelligence. It is unclear as to how this assessment of the combination of three abilities is intelligence-fair or how each ability could be assessed independently by such a test.

Gardner (1993) stated that he and colleagues were engaged in efforts to create operational definitions and diagnostic procedures for each intelligence area, although he admitted that it could be difficult to define and assess Intrapersonal and Interpersonal intelligences. In 2005, there still do not appear to be standardized testing instruments for the multiple intelligences. Existing measures, such as Teele Inventory of Multiple

Intelligences (TIMI; Teele, 1992) and the Multiple Intelligences Developmental Assessment Scales (MIDAS; Shearer, 1996) are self-report inventories rather than maximum-performance tests, and typically include items related to interests and learning preferences rather than abilities alone. Gardner (1993, 1999) has repeatedly stated that MI theory is quite distinct from learning style, in that learning styles are horizontal, cutting across intelligence domains, whereas each intelligence is content-specific.

It is unclear whether MI self-report measures are sound predictors of the intelligences, in the sense that the measures share strong convergent and weak discriminant correlations with measured ability. McMahon, Rose, and Parks (2004) reported that only TIMI Logical/Mathematical intelligence was a significant predictor of whether 288 fourth graders were below versus at or above grade level in reading achievement, suggesting either psychometric issues with the measure and/or an overlap in the intelligences. Shearer (1994) correlated full-scale IQ and ability measures with the Hillside Assessment of Perceived Intelligence (HAPI; Shearer, 1991), and concluded that the results supported HAPI's validity, despite the fact that estimated full-scale IQ was correlated with all HAPI measures; the lowest correlation was between IQ and HAPI Kinesthetic ($r = .23$). Additionally, Shearer admitted that some correlations were problematic, such as an Expressional Fluency test correlating more highly with HAPI Spatial-Perceptual ($r = .59$) than with HAPI Linguistic ($r = .48$). While Shearer's results could be seen as reflecting psychometric problems with HAPI, it seems equally likely that the results reflect reality—that is, a positive manifold emerged because all the ability measures tap into general intelligence. Similarly, when Shearer validated the MIDAS-KIDS (personal communication, October 7, 2003), correlations between the self-reported

abilities ranged from .39 (Logical and Musical) to .72 (Interpersonal and Intrapersonal), suggesting that children regard their own abilities as somewhat consistent.

The Intelligences

Below, the eight “intelligences” as conceptualized by Gardner (1993) are described in detail, with the aim of identifying the range of abilities subsumed by each domain and of examining the cognitive demands of tasks assessing these abilities. Based on this analysis, the tasks to be assessed as indicators of the various intelligences will be identified.

Linguistic Intelligence. Gardner has described Linguistic intelligence as sensitivity to spoken and written language and ability to use language to accomplish goals, as well as the ability to learn new languages. Lawyers, public speakers, writers, and poets all possess high levels of linguistic intelligence, according to Gardner.

The Linguistic intelligence domain as described by Gardner seems to encompass a large number of primary abilities. Thurstone, for example, differentiated between word fluency and verbal ability, two of his seven primary mental abilities, whereas Gardner would include both under the domain of Linguistic intelligence. Carroll (1993) noted that the line between the language domain and other domains, such as reasoning, is a difficult one to draw. Factors identified by Carroll in the language domain include language development, printed language comprehension, lexical knowledge, reading comprehension, reading speed, spelling ability, grammatical sensitivity, foreign language aptitude, oral production, writing ability, and others. Carroll noted that measures of language skills are substantially correlated with other cognitive abilities and that

vocabulary measures are among the best predictors of general intelligence. Gardner would likely argue that this relationship exists because measures of general intelligence tend to be heavily verbal.

Two important aspects of linguistic ability are vocabulary and associational fluency. Vocabulary would seem to be part of Thurstone's verbal ability construct. Interestingly, measures of vocabulary are typically highly correlated with full-scale IQ. For example, of the 11 subtests that constitute the full-scale Wechsler Intelligence Scale for Adults, 3rd edition (WAIS-III; Wechsler, 1997), the manual reports that Vocabulary has the largest *g* loading. While there is clearly a crystallized component of vocabulary, with its acquisition depending on exposure to words, Jensen (1991) argued that the broad heritability of vocabulary is convincing evidence of its value as an indicator of *g*. Carroll (1993) suggested that vocabulary measured verbal ability, acculturation, and crystallized intelligence.

Associational fluency, on the other hand, would seem to be part of Thurstone's word fluency primary ability. Associational fluency involves the association of words and falls into Guilford's (1967) factor of Divergent Production of Semantic Units (DMU), a category involving the generation of ideas based upon relationships. Associational fluency involves a search of long-term memory and the use of semantic flexibility (Ekstrom, French, Harman, & Dermen, 1976). The ability to generate several words that are associated by meaning should involve crystallized intelligence as well as divergent thinking. Carroll (1993) described Associational Fluency as a first-order factor in the domain of Idea and Language Production, that loaded on the second-stratum factor of Broad Retrieval Ability.

To measure Linguistic intelligence, the current study will include measures of associational fluency as well as vocabulary.

Spatial Intelligence. Gardner defined Spatial intelligence as the ability to recognize both large and small visual patterns. He has noted that navigators and pilots would possess spatial intelligence, as would sculptors, surgeons, chess players, and architects.

Gardner (1983) suggested that spatial ability comprised several underlying capacities, such as the ability to accurately perceive visual stimuli, to transform and modify one's initial perceptions, and to re-create visual stimuli in the absence of such stimuli. He described spatial capacities as being important to the recognition of objects and scenes, orienting oneself, and working with graphic depictions, such as diagrams and maps. While spatial ability is primarily perceived and expressed through the visual modality, Gardner noted that spatial tasks can be tactile as well.

Gardner (1983) described a basic distinction between static and active forms of spatial knowledge. Manipulating a mental object would be an active form, while merely holding a mental image was static. Gardner's static form of spatial knowledge seems to be related to visual memory, which Carroll (1993) reported to be a part of the memory and learning domain, while visualization is clearly active. Carroll (1993) identified several first-order factors as falling within the visual perception or spatial domain, including visualization, spatial relations, closure speed, closure flexibility, and perceptual speed.

Two important elements of Spatial intelligence are spatial visualization and spatial scanning. Spatial visualization refers to the ability to imagine the movement of an

object and is typically measured with mental rotation tasks. Carroll (1993) noted that visualization tasks are often very similar to tests of Spearman's g , such as Raven's Progressive Matrices. Hegarty and Waller (2004) determined that, although correlated, spatial visualization was distinct from perspective taking. Spatial scanning is the ability to scan a field quickly, visually follow paths, and reject false leads (Ekstrom et al., 1976). Carroll (1993) tentatively identified this capacity as a first-order factor, but stated that further research was necessary before it could be considered independent and interpreted accordingly.

To evaluate Spatial intelligence, the current study will include both a visualization and a spatial scanning measure, in order to include both an active and a static spatial task. These two tasks also address two important elements in Gardner's description of Spatial intelligence: the capacity to mentally manipulate objects, and the ability to work with maps.

Logical/Mathematical Intelligence. Gardner described Logical/Mathematical intelligence as the ability to study problems, carry out mathematical operations logically and analytically, and conduct scientific investigations. Gardner identified mathematicians, logicians, and scientists as professionals who would possess high levels of this intelligence.

Logical/Mathematical abilities are not clearly distinguishable from other domains in factor analytic studies. Reasoning, the domain most comparable to Gardner's "logical" abilities, corresponds to six first-stratum factors according to Carroll (1993). General reasoning, verbal reasoning, induction, quantitative reasoning, syllogistic reasoning and classification ability all fall into Carroll's category of reasoning. Quantitative reasoning,

typically measured in tasks requiring individuals to reason with concepts involving mathematical relations to arrive at correct conclusions, seems to be most related to Gardner's Logical/Mathematical intelligence domain. According to Carroll, however, quantitative reasoning is highly *g*-loaded. To maximize the validity of quantitative reasoning, Carroll suggesting minimizing demands for speed, vocabulary, syntax, spatial visualization, and complexity of numerical computations in the tasks.

Numerical facility, on the other hand, appears to measure a different aspect of mathematical ability, and one that is less *g*-loaded. Numerical facility is measured with tasks requiring participants to quickly perform simple arithmetic computations, such as addition, subtraction and multiplication. Carroll suggested that numerical facility is also related to working memory as well as perceptual speed.

To best represent the breadth of Gardner's Logical/Mathematical domain, the current study will include measures representing both quantitative reasoning and numerical facility.

Interpersonal Intelligence. According to Gardner (1983), an individual who is high in Interpersonal intelligence understands the intentions, motivations, needs, and desires of others, and is capable of working effectively with them. Gardner stated that teachers, clinicians, salespeople, politicians, and religious leaders all use Interpersonal intelligence.

Interpersonal intelligence appears to be closely related to the construct of emotional intelligence. Indeed, Gardner (1999) claimed that emotional intelligence was a combination of intrapersonal and interpersonal abilities. Davies, Stankov, and Roberts (1998) reported that emotional intelligence questionnaires were problematic in that they

appeared to be tapping into personality, but that objective measures currently in use have poor reliability. The authors concluded that emotion perception might be the only aspect of emotional intelligence that could meet the criteria to be considered a trait. In a study incorporating self and other ratings of emotional intelligence, academic intelligence and personality, Van der Zee, Thijs, and Schakel (2002) found three emotional intelligence factors they labeled as Empathy, Autonomy, and Emotional Control. The authors reported little relationship between emotional intelligence and academic intelligence but, similar to the findings of Davies et al. (1998), there were strong relationships between emotional intelligence factors and the Big Five personality factors, particularly Extraversion and Emotional Stability. In this study, emotional intelligence predicted academic and social success over and above academic intelligence and personality, lending support to the notion that it is a unique construct.

Whether the construct of emotional intelligence is associated with intelligence or with personality seems to depend on how it is measured. O'Connor Jr. and Little (2003) reported that an ability-based measure of emotional intelligence was correlated highly with cognitive ability, but not with personality. A self-report inventory of emotional intelligence, on the other hand, was highly correlated with personality but not with cognitive ability. Schulte, Ree, and Carretta (2004) reported that performance on an ability-based measure of emotional intelligence was predicted by general cognitive ability, Agreeableness, and sex, with women scoring higher than men in emotional intelligence. The authors concluded that the notion of emotional intelligence as a unique construct was not supported.

In the present study, Interpersonal intelligence will be evaluated with maximum

performance tests. One measure will assess participants' ability to understand differences in the meanings and intentions of statements made in varying social contexts, which O'Sullivan and Guilford (1976) called cognition of behavioral transformations. The authors described people who were high in this ability as being able to interpret expressions and situations accurately, and be quick to correct social errors. The second test will measure the ability to anticipate correctly the social consequences of various courses of action, using an entirely nonverbal format.

Intrapersonal Intelligence. Gardner described Intrapersonal intelligence as the ability to understand and have an effective working model of oneself. Intrapersonal intelligence includes the awareness of one's own desires, fears, and abilities, and also using this information to make sound life decisions.

From Gardner's description, it appears that having a clear conception of oneself is a key component of Intrapersonal intelligence. In an investigation of the nature of self-esteem, Campbell (1990) operationalized self-concept clarity by having participants make "me/not me" decisions for a 48-item list of adjectives, which contained terms representing opposite poles of various personality traits. Campbell was then able to examine the inconsistency of participants' self-descriptions by determining to what extent they endorsed opposites to describe themselves. Campbell reported that this measure effectively distinguished two groups who were hypothesized to differ in self-concept clarity. Therefore, this task might serve as an indicator of Intrapersonal intelligence.

Intrapersonal intelligence, as described by Gardner, seems to be somewhat related to metacognition in general and to self-monitoring in particular. That is, the individuals with high intrapersonal intelligence should be aware of what they know as well as what

they do not know. However, Stankov (2000) reported that his research has found very little correlation between self-monitoring, as measured by the difference between a confidence score and the actual percentage of correctly solved items, and intelligence. These findings could be interpreted as support for Gardner's contention that intrapersonal ability is an independent area of intelligence. Self-confidence is a reliable trait which is moderately correlated with measures of both intelligence and personality, according to Stankov. The author concluded that self-confidence, as a stable trait, straddles the borderline of personality and ability. Stankov stated that, in the search for a trait broader than self-confidence or self-monitoring, there has been some evidence that self-evaluation, the awareness of the quality of one's work following performance, is distinguishable from self-confidence. This broad metacognitive factor, if uncovered, might be a construct compatible with Gardner's Intrapersonal intelligence. Thus, measures of the extent to which individuals can accurately judge their relative strengths and weaknesses might serve as an index of Intrapersonal intelligence.

There will be two measures of Intrapersonal intelligence in the current study. The first will examine consistency of self-described personality, while the second will be a measure of the accuracy of participants' self-ratings of their abilities in each intelligence domain.

Naturalistic Intelligence. Gardner (1999) described a naturalist as one who is able to recognize and classify objects, suggesting that hunters, farmers and gardeners would have Naturalistic intelligence, as would artists, poets, and social scientists who are adept at pattern-recognition. He stated that a marketing professional who promotes the small differences between competing products is applying Naturalistic intelligence, as is the

individual who can recognize cars from the sounds of the engines.

Gardner stated that dramatic examples of naturalistic ability and its impairment come from clinical and experimental studies of brain-damaged individuals who are able to recognize and name inanimate objects although they have lost the ability to identify living things. Conversely, he said, other patients have the reverse impairment, being able to name only inanimate objects. He concluded that it should be possible to identify which brain regions play the crucial role in this naturalistic perception. Given Gardner's criterion that an intelligence must have the potential of isolation by brain damage, it is unclear as to why he has stated that a naturalist's abilities could just as easily be utilized with artificial items, such as cars or sneakers. Moreover, Carroll (1993) identified the classification of objects as an inductive reasoning task which, according to previous research, is highly *g*-loaded.

The crucial abilities in Naturalistic intelligence involve the capacity to categorize and make distinctions, so the naturalistic tasks in the current study will test these abilities. The first test will require participants to make logical groups from a list of objects. Ekstrom et al. (1976) categorize this test as a Flexibility of Use task, which includes elements of semantic flexibility and semantic redefinition. The second test requires participants to select which of five diagrams best describes the relationship between three items. This test is considered a logical reasoning task by Ekstrom et al., but the identifying relationships aspect clearly falls into Gardner's Naturalistic domain.

Bodily-Kinesthetic Intelligence. Gardner (1999) described this intelligence as the potential of using the whole body or parts of the body in problem-solving or the creation of products. Gardner identified not only dancers, actors, and athletes as those who excel

in bodily-kinesthetic intelligence, but also craftspeople, surgeons, mechanics, and other technicians. Thus, Gardner does not appear to differentiate between gross motor and fine motor abilities in describing Bodily-Kinesthetic intelligence. Shearer (1997) attributed the weak inter-item reliability of the Kinesthetic scale of the MIDAS self-report inventory of multiple intelligences to the mix of fine and gross motor items, suggesting the possibility of two distinct constructs.

Francis Galton investigated sensory-motor abilities as a possible basis of intelligence, but modern psychologists have typically made a point of focusing only on higher mental processes when operationalizing intelligence. There are still some who have included sensory processes within the realm of cognitive structure (e.g, Carroll, 1993), but the general agreement has been that psychomotor processes should not be included. In the current study, measures of both fine and gross motor abilities will be employed to assess Bodily-Kinesthetic intelligence.

Musical Intelligence. Gardner suggested that musical intelligence was parallel in structure to Linguistic intelligence, and that it would be reflected in the performance, composition, and appreciation of musical patterns. Gardner (1999) cited the existence of music prodigies and savants as evidence that musical ability is an intelligence in its own right, distinct from *g*. Ruthsatz and Detterman (2003), however, noted that many musical savants have unusual memory, which might explain their skillful musical performance. These authors reported that performance on Gordon's (1986) Intermediate Measures of Music Audiation was significantly related to short-term memory ($r = .57$) in musical novices, but not in students at a music institute ($r = .04$). The authors concluded that Gardner's Musical intelligence might be exceptional short-term memory transferred into

the musical domain. Brandler and Rammsayer (2003) reported that the only significant differences in intelligence between musicians and non-musicians were in verbal memory and reasoning, with musicians higher in verbal memory and non-musicians performing significantly better in reasoning. The authors, however, suggested that it was the long-term musical training that improved the verbal memory of the musicians, rather than superior verbal memory leading to high musical ability.

Lynn, Wilson, and Gault (1989) found positive correlations between musical ability and general intelligence ranging from .27 to .40 in a sample of 10-year-olds, suggesting that musical ability is not independent of *g*. Similarly, people with mental retardation have been found to have low musical ability (DiGiammarino, 1990). Of the often-noted connection between mathematical and musical abilities, Gardner (1999) maintained that the connection might be more about interest than expertise, speculating that while mathematicians and scientists seem to be attracted to music, musicians are not particularly attracted to mathematics. Root-Bernstein (2001) called musical and scientific abilities “correlative talents”, by which he meant that skills and abilities in one area inform skills in the other. The author claimed that the pattern-forming and recognition, kinesthetic ability, imaging, aesthetic sensitivity, analogizing, and analysis, developed through musical training, can be transferred to the scientific domain. In discussing research suggesting that musical training is associated with superior performance on spatial-temporal tasks, Gardner (1999) stated that if research consistently establishes that musical intelligence is primary, stimulating development in other domains, then he would need to make revisions to his list of intelligences.

Carroll (1993) noted that factor analytic research had largely had little success in

delineating the dimensions of musical ability, but Gardner (1983) has claimed that the two most central constituent elements of music are rhythm and pitch. Carroll identified 12 narrow factors relative to Auditory Receptive Ability. Of the 12, eight would seem to be relevant to Gardner's Musical intelligence: discrimination of tones and sequences of tones with respect to basic attributes such as pitch, intensity, duration and rhythm; auditory cognitive relations (judgments of complex relations among tonal patterns); tonal imagery; discrimination and judgment of tonal patterns in musicality; temporal tracking; ability to recognize and maintain mentally an equal-time beat; ability to retain on a short-term basis, images of tones, tonal patterns, and voices; and absolute pitch ability. Because rhythm and tone would appear to be core aspects of these narrow factors, sensitivity to rhythm and tone within a musical context, will be evaluated in the current study.

Self-Estimated Multiple Intelligences

Although the relations among actual abilities in the various intelligences are of primary interest, it is also of some value to examine (a) the relations among self-estimated abilities in each domain and (b) the relations between self-estimated and actual abilities. For example, will a *g* factor of self-estimated abilities emerge? Will self-estimated abilities correlate strongly with actual abilities?

When Furnham, Tang, Lester, O'Connor, and Montgomery (2002) factor analyzed participants' self-estimates of Gardner's multiple intelligences, they found two factors. Self-estimated verbal, logical, spatial, and naturalistic abilities loaded on one factor, while self-reported bodily-kinesthetic, interpersonal, and intrapersonal abilities loaded on the second. The authors suggested that the two factors represented fluid and

crystallized intelligence, but this solution would seem to ignore the fact that verbal abilities are typically viewed as crystallized, whereas bodily-kinesthetic abilities are not cognitive. Another explanation of the two factors could be that they represent cognitive and non-cognitive abilities. Furnham et al. admitted that the findings were clearly limited by the absence of any measures of actual ability.

When self-estimates of intelligence have been compared to actual performance, correlations tend to be small but significant (Paulhus, Lysy, & Yik, 1998). Furnham and Rawles (1999) reported a significant correlation of .16 between self-estimated general intelligence and performance on a mental rotation task, with men ($r = .27$) more accurate than women ($r = .09$). However, those participants who might have evaluated their intelligence on the basis of their verbal abilities could have appeared inaccurate only because the criterion variable was not, in fact, a measure of intelligence in general, but rather of spatial ability in particular. Borkenau and Liebler (1993) reported a correlation of .32 between self-rated and measured intelligence, whereas other-rated intelligence correlated .29 with measured intelligence. In the same study, strangers watched a film of each participant reading a standard text, then rated the participant's intelligence. The mean zero-order correlation between the stranger estimate and measured intelligence was .25, but when the estimates of the six stranger judges were aggregated, the correlation rose to .38. Furnham and Fong (2000) reported a mean correlation of .19 between self-rated and measured intelligence. Paulhus et al. found that participants' self-ratings correlated from .20 to .23 with measured intelligence, but when a composite rating scale of four items ("Is intelligent"; "Is smart"; "Is ingenious, a deep thinker"; and "Is not exceptionally gifted at academic things" [reverse-keyed]) was used, the correlation of the

aggregated measure with measured intelligence in the same samples increased to between .24 and .26. Scales with indirect items (e.g., one item was “As a child I was always reading”) had correlations ranging from .17 to .27. The authors concluded that self-report measures were not acceptable substitutes for IQ tests in a college sample, and that even with aggregation, the validity limit seemed to be .30. They found that direct items about broad cognitive abilities as well as those referring to enjoyment and frequency of reading were the most valid.

Two measures of self-estimated ability were included in the current study to reflect direct and indirect methods of assessing self-reported intelligence. The first measure was a direct self-rating scale in which participants received a brief description of each ability, and ranked themselves in comparison to other university students. This direct method is similar to that used by Furnham and colleagues (e.g., Furnham, Clark, & Bailey, 1999; Furnham et al., 2002) who showed participants a normal distribution with a mean of 100 and requested that participants indicate their self-estimated IQ for each ability. The second measure of self-estimated ability used in the current study was indirect. A self-report questionnaire was developed to be comparable to many of the multiple intelligence questionnaires currently in use in which respondents indicate their agreement or disagreement with statements describing specific, domain-relevant activities, corresponding to Gardner’s eight intelligences.

Personality and Intelligence

Recent studies have explored the association of personality traits with intelligence and achievement. Results have generally supported the notion that the “Big Five” (see

McCrae & John, 1992) factor of Openness to Experience is positively and significantly correlated with intelligence (Ackerman & Heggstad, 1997; Holland, Dollinger, Holland, & McDonald, 1995). Recent studies have suggested that the nature of the relationship is that Openness to Experience is substantially associated with crystallized (acquired knowledge, such as vocabulary) but not fluid (reasoning and the ability to see relationships) aspects of intelligence (Ashton, Lee, Vernon, & Jang, 2000; Bates & Shieles, 2003), presumably because individuals who are high in Openness to Experience seek out information about their worlds, thus, accumulating greater vocabularies and knowledge. Some studies have reported significant positive correlations between Openness to Experience and academic achievement (De Fruyt & Mervielde, 1996; Farsides & Woodfield, 2003; Hair & Graziano, 2003) while others have reported no association between the two (Busato, Prins, Eshout, & Hamaker, 2000).

Two recent studies have found that Conscientiousness was significantly and negatively correlated with intelligence (Moutafi, Furnham, & Crump, 2003; Moutafi, Furnham, & Paltiel, 2004) with Moutafi et al. (2004) reporting that Conscientiousness was negatively correlated only with fluid intelligence. These findings would have to be replicated before any conclusions can be drawn. The literature has been consistent, on the other hand, in reporting significant, positive correlations between Conscientiousness and academic achievement (e.g., Chamorro-Premuzic & Furnham, 2003; De Fruyt & Mervielde, 1996).

For the current study, personality will be assessed with a list of 48 adjectives related to six personality factors, as conceptualized by Lee and Ashton (2004): Honesty-Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness, and Openness

to Experience. The Honesty-Humility factor includes tendencies to be genuine in interpersonal relations, to avoid fraud and corruption, to be uninterested in lavish wealth and privilege, and to be modest and unassuming. Emotionality involves tendencies to experience fear and worry, to need emotional support from others, and to feel strong emotional bonds with others. The Extraversion factor assesses tendencies to be excitable and dramatic, to be confident in a variety of social situations, to enjoy social interaction, and to be enthusiastic and energetic. Agreeableness includes a willingness to compromise and cooperate with others, and to be trusting and forgiving of those who have hurt one. Agreeableness also involves tendencies to be mild and lenient with others, and to remain calm, rather than angry. The Conscientiousness factor includes tendencies to seek order, to work hard, to be thorough and detail-oriented, and to deliberate carefully and inhibit impulses. Openness to Experience assesses enjoyment of beauty in art and nature, a preference for innovation and experiment, and tendencies to seek information and experience with the natural and human world and to accept the unusual.

To assess self-reported personality, an objective self-rating form will be used. Participants will be asked indicate on a 9-point scale how closely each adjective describes their personality.

Sex Differences

Multiple Intelligences. While some research has shown sex differences in favour of men in general intellectual ability (Lynn, 1999), many studies have reported no such sex differences (Colom & Garcia-Lopez, 2002) or negligible differences (Jensen, 1998). Colom, Juan-Espinosa, Abad, and Garcia (2000) reported a mean correlation of .00 in a

study of sex differences in *g* with independent samples of over 10,000 adults. The method of correlated vectors used in the latter study has, however, been criticized by Ashton and Lee (2005) as potentially resulting in vector correlations of zero even when real differences in *g* exist between two groups.

While sex differences in general intelligence might be minimal, it is possible that there are sex differences in the various domains of intelligence. Gardner (1983) noted that sex differences seem to be most pronounced in spatial ability. More specifically, in Halpern's (1997) review of sex differences in intelligence, she noted that men tend to perform better on tasks that require transformations in visual working memory, such as mental rotation. Gardner (1983) suggested that men's superior performance in spatial tasks could be explained by the evolutionary advantage of spatial ability to men, who were traditionally responsible for hunting, a task that involved wandering over long distances.

With respect to sex differences in Linguistic intelligence, women tend to outperform men in verbal fluency, spelling, reading comprehension, writing, synonym generation, and knowledge of foreign languages (Halpern, 1997). In the Logical/Mathematical domain, men, on average, outperform women on tests of mathematical and scientific reasoning (Halpern, 1997). Women, however, obtain higher average scores than do men on tasks involving mathematical calculations (e.g., Hyde, Fennema, & Lamon, 1990).

In the Bodily-Kinesthetic domain, Halpern (1997) noted that whereas men tend to perform better at motor tasks that involve aiming, such as accuracy at throwing darts, women score higher in fine motor tests, such as pegboard tasks. One aim of the present

research will be to examine sex differences in the various domains.

Self-Estimated Intelligences. Related to the use of self-report measures, there is a burgeoning literature in self-estimated multiple intelligences (see Furnham, 2001, for a review). Hogan's (1978) finding that men's self-estimates of their general intelligence were higher than women's has been replicated (Bennett, 1996; Furnham & Rawles, 1995). Hogan reported that not only were women's self-estimates lower than men's, but that women significantly underestimated their own intelligence scores. Recent studies (Furnham, Clark, & Bailey, 1999; Rammstedt & Rammsayer, 2000), however, have found no sex differences in self-estimated general intelligence. Reilly and Mulhern (1995) initially found that men's self-estimates of general intelligence were higher than women's, but the effect disappeared when outliers were removed. When participants have been asked to estimate their multiple intelligences rather than their general intelligence, men's self-estimates of mathematical and spatial abilities have been higher than women's (Furnham, Clark, & Bailey, 1999; Furnham & Ward, 2001; Rammstedt & Rammsayer).

Personality. Studies have consistently found that women report higher levels of the Big Five trait of Neuroticism (Loehlin, McCrae, & Costa, 1998; Rammsayer & Rammstedt, 2000), with Costa, Terracciano, and McCrae (2001) reporting that in reviewing studies of 1,000 American adults, there were consistent sex differences in Neuroticism ($d = .51$). Sex differences in Neuroticism in the 25 other cultures from which Costa et al. reviewed data ranged from $d = -.02$ in a Zimbabwean sample to $d = .70$ in an Italian sample. Lee and Ashton (2004) found that women reported significantly higher levels of the related HEXACO factor of Emotionality ($d = 1.08, p < .01$).

Some studies have found that women report higher levels of Big Five Agreeableness (Budaev, 1999). Costa et al.'s (2001) secondary analyses of independent samples from 26 cultures reported that American women reported significantly higher levels of Agreeableness ($d = .59$), and sex differences in 25 other cultures ranged from $d = -.05$ in a Zimbabwean sample and $d = .05$ in a black South African sample to $d = .55$ in a Belgian sample. Lee and Ashton found no sex differences in self-reports of the HEXACO version of this factor ($d = -.11$). This discrepancy in results could be attributed to the fact that the Big Five Agreeableness factor includes some aspects of HEXACO Emotionality; traits such as sentimentality and sensitivity, which would be found in the Big Five Agreeableness factor, are a part of Emotionality in the HEXACO inventory. Thus, the Agreeableness factor as conceptualized in the HEXACO model, does not appear to be one in which sex differences are found.

In addition to the anticipated sex differences in Neuroticism and Agreeableness, Costa et al. (2001) found that American women reported higher levels of Openness to Feelings ($d = .28$) and Openness to Aesthetics ($d = .34$), whereas men reported greater Openness to Ideas ($d = -.32$). Lee and Ashton (2004) found that men reported significantly higher levels of Openness to Experience ($d = -.22$), based on greater self-reported inquisitiveness ($d = -.62$), creativity ($d = -.25$), and unconventionality ($d = -.31$), although women reported greater aesthetic appreciation ($d = .39$). Costa et al., on the other hand, found no significant differences in overall Openness to Experience.

Lee and Ashton found that women reported greater Conscientiousness ($d = .26$), whereas Costa et al. (2001) found no sex differences. Sex differences in Extraversion vary according to the content that is emphasized, with men scoring higher in dominance

and women in nurturance. Costa et al. noted that the NEO-PI-R Extraversion factor weights warmth more heavily than assertiveness, so women tend to score higher on this particular scale. Lee and Ashton found no overall sex differences in Extraversion, although women reported greater expressiveness ($d = .30$) and sociability ($d = .31$), and men reported greater social boldness ($d = -.30$).

The Current Study

The current study will investigate Gardner's Multiple Intelligences theory by examining the independence of the eight intelligences, as well as their correlations with a measure of general cognitive ability. While Gardner has diluted MI theory somewhat by incorporating the existence of g and suggesting that intelligences might not be entirely independent, the theory would still seem to predict that tests of the eight intelligences should be relatively independent of each other.

Given the substantial research support for hierarchical models of intelligence and the current lack of empirical support for Multiple Intelligences theory, the current study will take the position that hierarchical theories will best predict and explain the pattern of results that appears in this investigation. The current study will hypothesize a positive manifold among all cognitive tests. Given the relatively small number of ability tests, it is not anticipated that Carroll's three strata will emerge, although support is expected for the existence of group factors in addition to g , as would be revealed by substantial residual correlations between tests of the same intelligence domain even after the extraction of the g factor. Each intelligence domain will be measured with two maximum performance tasks. In addition, a measure of general intelligence will be employed.

Gardner (1999) cautioned that there is a tendency to measure non-verbal abilities with verbal measures, leading to artificially high correlations among the ability domains. Whenever possible, at least one non-verbal measure was selected for any given domain, in order to avoid the problem of the “verbal lens”.

In addition, the relationship of personality and abilities will be explored in this study, as well as the association of direct and indirect self-estimates with each other and with measured intelligence.

Hypotheses

Measured Multiple Intelligences

1. Because each cognitive ability test should measure *g* in addition to a specific ability, all of the tests except those in the non-cognitive domains (Bodily-Kinesthetic, Intrapersonal, and Musical) will be correlated with each other and with an independent measure of general intelligence. Factor analysis should yield a single factor of general intelligence, with all cognitive tests having substantial *g*-loadings, whereas non-cognitive tests should have lower, near-zero *g*-loadings.
2. Due to associations within domains of narrow ability factors independent of *g*, it is hypothesized that the two within-domain tests for each intelligence, with the exception of Bodily-Kinesthetic, will be significantly and positively correlated above and beyond their association with the other measures, as indicated by *g*-loadings. For the Bodily-Kinesthetic domain, the measures of gross- and fine-motor skills are expected to be nearly uncorrelated.

Self-Estimated Ability

3. It is expected that self-estimates will be positively and significantly correlated with each other, and that factor analysis of the self-estimates (both direct and indirect) of cognitive abilities will yield a large general factor, similar in size to that of the measured abilities. It is also expected that self-estimated intelligences will correlate moderately with these corresponding measured intelligences.

Personality

4. It is anticipated that Openness to Experience will be significantly and positively correlated with academic average and with measures of crystallized intelligence (Linguistic and Interpersonal measures) but not with measures of fluid intelligence (Spatial, Naturalistic, and Logical/Mathematical tasks) or with the non-cognitive tasks (Musical, Intrapersonal, and Bodily-Kinesthetic measures). Conscientiousness will, in keeping with previous research, be significantly and positively correlated only with academic average.

Sex Differences

5. It is anticipated that sex differences in measured abilities will be minor. However, it is predicted that men will perform better on a visualization (Spatial) and a mathematical reasoning (Logical/Mathematical) task, whereas women will outperform men in a task requiring them to generate opposites (Linguistic) as well as in a test of fine motor ability (Bodily-Kinesthetic), and in both Interpersonal measures. For the self-report measures of abilities, it is expected that, in keeping with previous research, men's self-estimates of Logical/Mathematical ability, Spatial ability, and overall intelligence will be higher than women's, and that

women's self-estimates of Musical and Interpersonal abilities will be higher than those of men. With regards to personality, it is predicted that women's self-reported Emotionality and Honesty will be higher than will men's.

Method

Participants

Two hundred participants (116 women) were recruited through posters, in-class presentations, and sign-up sheets on two university campuses. To recruit participants, presentations were made in a variety of classes, including chemistry, accounting, sports management, psychology, and geography. Participants were required to speak English as their first or best language and to be at least 18 years of age. Seventeen (8.5%) of the participants were recruited and tested at the University of Western Ontario, while the remaining 183 (91.5%) participated at Brock University (see Table 1 for demographic data). While the vast majority of participants reported being undergraduate students, there were also university graduates, graduate students, high school graduates, as well as a college student, a high school student, and some who did not report their student status. Participants ranged in age from 17 to 66 years ($M = 22.73$, $SD = 6.14$). The 186 participants who reported an academic major represented 30 different academic departments, with self-reported academic averages ranging from 56% to 94% ($M = 75.91$, $SD = 7.59$) for the previous year.

Measures

Demographic Information. Participants provided the following demographic information: age, sex, faculty in which the student is enrolled, subject in which the student is majoring, year of study, and previous year's average mark, expressed as a percentage.

General Cognitive Ability. The Wonderlic Personnel Test (WPT; Wonderlic, 2002) Form A was used as a measure of general intelligence. The WPT is a 50-item,

Table 1

Demographic Data

Variables	Number	Percent	Missing (%)
Sex			0.0
Male	84	42.0	
Female	116	58.0	
Age			0.0
18 to 25	171	85.5	
26 to 33	16	8.0	
34 to 41	6	3.0	
42 to 49	6	3.0	
50 or more	1	0.5	
Student Status			3.5
Completed High School	4	2.0	
University Undergraduate Student	171	85.5	
Completed University	10	5.0	
University Graduate Student	6	3.0	
Other	2	1.0	
Year			12.5
1st	21	10.5	
2nd	59	29.5	
3rd	54	27.0	
4th	39	19.5	
5th	2	1.0	
Faculty			8.0
Social Science	74	37.0	
Business	38	19.0	
Humanities	26	13.0	
Math & Science	21	10.5	
Applied Health Science	18	9.0	
Other	7	3.5	
Department			8.5
Psychology	33	16.5	
Finance/Accounting	29	14.5	
Communications/Film	13	6.5	
Phys. Ed./Kinesiology/Recreation	11	5.5	
Child Studies	10	5.0	
Biology	10	5.5	
Business	8	4.0	
English	7	3.5	
Sociology	7	3.5	
Undeclared	7	3.5	

(continued...)

Table 1 (Continued)

Variables	Number	Percent	Missing (%)
Political Science	5	2.5	
Visual Art	5	2.5	
Tourism	4	2.0	
History	3	1.5	
Geography	3	1.5	
Music	3	1.5	
Linguistics	3	1.5	
Chemistry	3	1.5	
Other	19	9.5	

12-minute omnibus test that has been extensively normed on 126,000 individuals. WPT items represent a broad range of problem types (e.g., analogies, mathematical problems, proverbs, geometric figures, number series, and word definitions) which are presented in order of increasing difficulty. The score is the number of correct responses. The manual reports alternate form reliabilities of .73 to .95, internal consistency ranging from .88 to .94, and short-term test-retest reliabilities of .82 to .94. In a longitudinal study (Dodrill, 1983), the five-year test-retest WPT reliability was .94 in 30 normal adolescents and adults. The manual reports that the WPT correlates .91 to .93 with the Wechsler Adult Intelligence Scale full-scale IQ.

Linguistic Intelligence. The Educational Testing Service (ETS) Advanced Vocabulary Test II is a 36-item, eight-minute test of verbal comprehension, appropriate for a university undergraduate population. For each target word, participants select one of four words which has the same or nearly the same meaning as the target word. The score is the number of correct responses less a fraction of incorrect responses. The ETS manual

reports that the reliability was .83 when this test was administered to 181 college students.

The ETS Opposites Test assesses associational fluency by requiring participants to generate words which are opposite or nearly opposite in meaning to a given word, for each of eight target words (four on each half of the test). Scores reflect the number of opposites that participants were able to generate, up to a maximum of six per target word, within the five-minute time limit for each half. The Opposites test had a reliability of .82 in a study of 625 male Naval recruits (Ekstrom, French, Harman, & Dermen, 1999).

Spatial Intelligence. The ETS Map Planning Test is a 40-item, six-minute test of spatial scanning. For each of two parts of the test, participants are presented with a map with streets, road-blocks and numbered buildings. Participants must find the shortest route between two lettered points, avoiding road-blocks. The correct response is the number of the building which is passed in traveling the shortest route between two given points. The score is the number of correct responses. The manual reports that the reliability of this test ranged from .75 to .80 when administered to high school and college students. Carroll (1993) found little evidence for spatial scanning as a factor, and suggested that these tasks might be related to perceptual speed or visualization. However, this test seems to fit well with Gardner's description of sense of direction and map-reading as aspects of spatial intelligence.

The ETS Paper Folding Test is a 20-item, six-minute measure of the ability to make serial mental rotations of spatial stimuli, which Carroll (1993) identified as a visualization task. Each item consists of drawings of a square of paper that is folded two or three times, in various ways, and then had a hole punched through it. Participants

choose which of five drawings accurately depicts how the sheet of paper and the pattern of holes in the paper would appear when fully opened. The score is the number of correct responses less a fraction of the number of incorrect responses. The manual reports reliabilities of .75 and .84 in studies with high school students and with college students/Army enlistees, respectively.

Logical/Mathematical Intelligence. The ETS Subtraction and Multiplication Test is a 120-item, four-minute measure of the ability to perform basic arithmetic operations with speed and accuracy. The ETS manual notes that this factor is not a major component in mathematical reasoning or higher mathematical skills. The score is the number of correct responses. The manual reports that the reliability of this test was .92 when it was administered to 119 ninth-grade boys.

The ETS Necessary Arithmetic Operations Test is a 30-item, 10-minute measure of general arithmetic reasoning. The problems require the participant to determine which arithmetic operations are required to solve mathematical word problems, but do not require the participant to perform those computations. The score is the number of correct responses less a fraction of incorrect responses. The manual reports that the reliability of this test was .71 when it was administered to 119 ninth grade boys.

Interpersonal Intelligence. Social Translations (O'Sullivan & Guilford, 1976) is a 10-minute test with 24 items. Participants are given a statement and told who said the statement to whom. Participants choose one of three pairs of people with which the same verbal statement would have a different meaning. For example, "That's mine" said by a man to a parking lot attendant is different from the same phrase when said by an angry child to a playmate. The score is the number of correct responses plus $1/k$ times the

number of items left blank, k being the number of alternative answers. The authors described this test as a measure of the ability to recognize changes in behavioral meaning. They noted that an individual high in this ability should be good at interpreting expressions and situations. Internal consistency reliability has ranged from .84 (O'Sullivan, Guilford, & deMille, 1965) to .90 (O'Sullivan, 1973). Norms are available for Caucasian 10th grade and college students.

Cartoon Predictions (O'Sullivan & Guilford, 1975) is a 10-minute test with 30 items, each consisting of one cartoon showing people's reactions in a situation. Participants must choose one of three related cartoons to determine which would most logically happen next. For example, the sample item shows a man clinging to the roof of the house, while a child looks on. The man appears to be frightened and asking for help. The correct response shows the boy with his mother bringing a ladder to help the man down. The two incorrect responses show the man trying to climb on the roof, and the boy and his mother laughing at the man. The score is the number of correct responses plus $1/k$ times the number of items left blank, k being the number of alternative answers. This test was chosen because it seems to tap into social ability without drawing on verbal ability. Sullivan and Guilford noted that the test measures the ability to predict social consequences. Internal consistency reliability for this test has ranged from .31 (O'Sullivan, 1973) to .70 (O'Sullivan et al., 1965). Norms are available for Caucasian 10th grade and college students.

Intrapersonal Intelligence. One Intrapersonal measure involved the consistency of self-ratings of personality adjectives belonging to the same factor. Recall that participants completed a personality self-rating form (Appendix D), in which they rated, on a 9-point

scale, how closely each adjective described their typical personality. The standard deviation of scores was calculated for each individual across his or her self-ratings on adjectives within each factor (after recoding of reverse-keyed items). The average standard deviation across factors, each of which can range from 0 to 4, indicates inconsistency of self-description; this average value, when multiplied by -1, was used as an index of consistency of self-description.

The second intrapersonal measure was an accuracy coefficient, which reflected how closely participants' self-ratings in each ability domain matched their actual performance. As previously described, participants rated their ability on each of nine domains corresponding to Gardner's eight intelligences as well as overall intelligence. Recall that for each ability, participants indicated which percentage of their fellow-students would have a lower level of that ability than themselves. A within-person correlation between self-estimated ability and actual ability was calculated across the eight abilities after conversion of percentiles to standard scores.

Naturalistic Intelligence. ETS Making Groups is a 4-item, 10-minute test of the ability to combine objects into logical groups. Each item consists of a list of seven items, for example: *trout, robin, frog, car, boat, bat, and airplane*. The participant is asked to generate groups of three or more objects into a maximum of 10 groups, and to provide a rationale for each grouping. In the previous example, one group might be *trout, robin, frog, and bat*, with the explanation that all are animals. The score is the number of correct groups. The manual reports that the reliability of this test was .60 when it was administered to 625 male Naval recruits.

ETS Diagramming Relationships is a 30-item, 8-minute test of the ability to

diagram relationships among objects. For example, if given the items, *animals*, *cats*, and *dogs*, the correct diagram would be the one with two non-overlapping circles (representing *cats* and *dogs*) inside another circle (representing *animals*). For each item, participants are asked to select one of five diagrams that best illustrates the relationships among the sets of objects. The score is the number of correct choices less a fraction of the number of incorrect choices. The manual reports that this test had reliability of .79 when it was administered to 542 male Naval recruits.

Bodily-Kinesthetic Intelligence. For the Bodily-Kinesthetic domain, tests of both fine and gross motor skills were used to fully represent the domain as Gardner conceptualized it, and to determine whether there are differences in the relations of the two sub-abilities with the other intelligences. A test of balance was chosen to represent gross motor ability, because it is a key element in assessing potential in gymnastics, diving, skating, and many other athletic activities (Johnson & Nelson, 1986). In addition, balance seemed to be a measure that would be less susceptible to training effects than would many other potential measures, such as strength, endurance, or speed.

The Stork Stand (Johnson & Nelson, 1986) was used to measure balance. Participants were timed while standing with hands on hips, one leg raised and bent with its foot on the inside of the other knee. On a signal, participants raised their heels so that only the ball of one foot remained on the floor. Each participant's score was the average of the two longest times in seconds between raising the heel and losing balance, from three trials. The Stork Stand has high test-retest reliability (.87) and inter-rater reliability is reported to be as high as .99 (Johnson & Nelson, 1986).

The General Aptitude Test Battery (GATB; U.S. Department of Labor, 1970)

Mark Making test was used to measure fine motor ability. Participants were required to make marks consisting of two parallel lines with another line underscoring the two. Participants were asked to make one such mark in each of many squares on the printed page, filling as many squares as possible in 60 seconds. The three-month test-retest reliability of the Mark Making test was .76 in a sample of 1,000 boys (Cronbach, 1960). The test score is the number of marks completed in the allotted time.

Musical Intelligence. GIA Publications' Advanced Measures of Music Audiation (AMMA; Gordon, 1989) is a 30-item, 16-minute test, consisting of sets of musical statements played to participants on a CD player. Three practice items are included. For each set, the participant must decide whether the two musical statements are the same, tonally different, or rhythmically different.

No musical background is required for the AMMA. The test is based on Edwin Gordon's (1989) contention that audiation—the ability to hear mentally and to comprehend musical patterns when the sound is no longer present—is the basis of musical aptitude. The AMMA assumes that being able to hold a musical pattern in memory and to make a comparison and judge the presence and type of discrepancy is a better measure of musical aptitude than is musical achievement. Each set consists of an equal number of tones. Tonal differences (10 items) can be due to changes in individual pitch, in mode tonal center (the musical key of the melody), or in a combination of these. Rhythm differences (10 items) consist of items with differences in durations, meters, tempos, or combinations of these. Ten item pairs are identical.

Tonal and Rhythm scores were calculated. The Tonal score (40 points possible) was calculated by summing the number of correct tonal responses (10 points possible)

and the number of correct identifications of sameness (10 points possible) plus 20 points. The number of erroneous tonal identifications was then subtracted from the score, with omitted items counted neither for nor against. Rhythm (40 points possible) was calculated similarly. For the purposes of all analyses in the current study, however, only the first half of the Rhythm test and the second half of the Tonal test were used, with scores potentially ranging from 0 to 20. This approach was taken to correct for the fact that Tonal and Rhythm total scores were artificially correlated due to overlapping content in incorrect responses.

The AMMA was normed with 872 high school students as well as 5,336 college and university students, including music majors and non-majors, across the United States. Split half reliability ranged from .80 to .88 and test-retest reliability ranged from .80 to .89. Correlations with Gordon's Musical Aptitude Profile range from .43 to .85. Gordon reported validity of .82 when total AMMA scores were used to predict the performance of 114 university students enrolled in orchestra, choir, and band.

Self-Estimates of Ability. Two instruments were used to obtain self-estimates of abilities. The Multiple Ability Questionnaire (MAQ; Appendix B) was developed for use in the present study, and consisted of 48 self-report true/false items representing the eight ability domains (e.g., "I can usually tell when someone is lying to me," representing Intrapersonal ability). Of the six items for each ability, two are reverse keyed (e.g., "I'm frequently off-key when I sing"). This test was scored with one point awarded for each *True* response to a statement representing an ability domain, with the exception of the reverse keyed items, for which a point was awarded for each *False* response. A score out of six was tallied for each domain. The MAQ did not include any items intended to

represent general intelligence.

The second measure of self-estimated ability was a self-rating scale (Appendix C). Participants rated their relative ability on each of nine items, which represented the eight intelligence domains as well as general intelligence (e.g., the item for Naturalistic Intelligence is “Classifying things [especially living things] and making distinctions between similar objects”). For each item, participants were instructed to compare themselves to other Brock University students by placing an *x* on a line, anchored by 0% and 100%, to estimate what percentage of students have a lower level of ability than the participant. The score on this measure was simply the percentile ranking that participants assigned themselves.

Personality. A list of 48 adjectives (Appendix D), which were found to be highly loaded (Ashton, Lee, & Goldberg, 2004) on the personality factors of Openness to Experience, Conscientiousness, Extraversion, Agreeableness, Emotionality, and Honesty-Humility, was used to measure self-reported personality. The adjectives represented the high-loading terms on these six factors of personality variation. For each personality variable, with the exception of Honesty-Humility, three of the six adjectives were reverse-keyed (e.g., *unimaginative* for Openness to Experience). For the Honesty-Humility trait, all six adjectives were reverse-keyed (e.g., *pompous*). For each adjective, participants rated themselves on a scale of 1 (*extremely inaccurate*) to 9 (*extremely accurate*). A composite score was calculated for each personality factor by averaging the ratings on the six relevant adjectives.

Procedure

Participants who met the inclusion criterion (18 years or greater, English as first

or best language) were tested in groups of 2 to 36 in various meeting rooms on a university campus. All participants were given information letters as well as an oral description of procedures prior to providing signed informed consent. The testing session took place over a 2.5 to 3 hour period, with a 10-minute break at approximately the halfway point, following completion of the music test.

The 15 measures were presented in two different orders. The first order was as follows: Multiple Ability Questionnaire (indirect self-estimate); self-rating (direct self-estimate); Opposites (Linguistic); Map Planning (Spatial); Subtraction and Multiplication (Logical/Mathematical); Cartoon Predictions (Interpersonal); Diagramming Relationships (Naturalistic); Stork Stand (Bodily-Kinesthetic); AMMA (Musical); 10-minute break; Making Groups (Naturalistic); Social Translations (Interpersonal); Necessary Arithmetic Operations (Logical/Mathematical); Paper Folding (Spatial); personality inventory; and Mark Making (Bodily-Kinesthetic). The second order reversed the presentation of the Linguistic, Spatial, Logical/Mathematical, Interpersonal, and Naturalistic tests. The second order was as follows: Multiple Ability Questionnaire (indirect self-estimate); self-rating (direct self-estimate); Making Groups (Naturalistic); Social Translations (Interpersonal); Necessary Arithmetic Operations (Logical/Mathematical); Paper Folding (Spatial); Stork Stand (Bodily-Kinesthetic); AMMA (Musical); 10-minute break; Opposites (Linguistic); Map Planning (Spatial); Subtraction and Multiplication (Logical/Mathematical); Cartoon Predictions (Interpersonal); Diagramming Relationships (Naturalistic); personality inventory; and Mark Making (Bodily-Kinesthetic). One hundred and one (50.5%) participants completed tests in the first order and the remainder completed them in the second.

Each participant was paid \$30 upon completion of the testing session and provided with a feedback letter.

Results and Discussion

Missing Data

The amount of missing data (Table 2) ranged from 0.0% (Sex, Age, Vocabulary, Paper Folding, Necessary Arithmetic Operations, Social Translations, Diagramming Relationships, Mark Making) to 12.5% (Academic Year) with an overall average missing data on all demographic and study variables of 1.65%. Variables related to student status (Academic Year, Major, Faculty, and Average) had the greatest amounts of missing data due to the fact that these categories were irrelevant for a number of participants who were not undergraduate students. Academic year, for example, tended to be left blank by graduate students as well as non-students.

Six and one half per cent of all participants had missing data on one or more of the ability, personality, and/or self-estimate variables, but for any given measure, the proportion of participants with missing data never exceeded 1.0%. The average data missing in the variables used in data analyses was only 0.70%. Cohen, Cohen, West, and Aiken (2003) noted that with three per cent or less data missing, the method of handling the missing data makes far less difference than in cases of larger amounts of missing data. In addition, Cohen et al. described samples of greater than 200 as being less problematic than smaller samples in terms of missing data.

Reasons for missing data on ability tests in the current study were varied and sometimes unknown. For example, the one individual who did not complete the Stork Stand (Bodily/Kinesthetic) had a broken leg. The two Opposites (Linguistic) tests coded as missing were due to participants providing synonyms rather than opposites in response to target words. The only Rhythm and Tone missing score was due to the music response

Table 2

Percentage of Missing Data by Variable

Variable	Percent	N
Sex	0.0	200
Academic Average	7.0	186
MAQ	0.0	200
Self-Ratings	1.5	197
Opposites	1.0	198
Vocabulary	0.0	200
Map Planning	0.5	199
Paper Folding	0.0	200
Subtraction and Multiplication	0.0	200
Necessary Arithmetic	0.0	200
Cartoon Predictions	0.5	199
Social Translations	0.0	200
Consistency	0.0	200
Accuracy	2.0	196
Diagramming Relationships	0.0	200
Making Groups	1.0	198
Stork Stand	0.5	199
Mark Making	0.0	200
Rhythm	0.5	199
Tone	0.5	199
Wonderlic	0.5	199
Personality	0.0	200

Overall Average Missing Data = 0.70%

card being absent from one participant's package of completed test materials.

Listwise deletion would have resulted in a loss of data from 13 participants (6.5%) of the sample of 200, a reduction that was considered too severe, particularly because this approach renders uncertainty as to whether the remaining sample is unbiased. Pairwise deletion, although largely preserving the sample size, is based on the assumption that data are missing completely at random, which could not be assumed for

the current study. Thus, a conservative approach to managing the missing data was employed, and estimation using Expectation Maximization (EM), an iterative procedure, was conducted in order to impute values. Enders (2001) reported that EM was superior to listwise deletion, pairwise deletion, and means substitution, in terms of bias and sampling variability. The EM procedure relies on the assumption that data are missing at random, a less stringent criterion than missing completely at random, and one which these data would appear to satisfy.

More problematic was the discovery that several very low Social Translations (Interpersonal) scores were potentially the result of a misinterpretation of test instructions. A histogram of the distribution of Social Translations scores revealed some departures from normality, with several scores clustering at the very low end of the distribution. In fact, seven participants scored 0.00 on both halves of the test. Further examination of the original test booklets revealed that some participants had responded in a way that suggested they had likely misunderstood the instructions. In the Social Translations test, participants were given a statement and told who made the statement to whom. They were then provided with three other pairs of speaker/listener and asked to identify the pair for whom the statement took on a different meaning or intention. It appears that some participants might have responded as though they were asked to determine which of four pairs made the statement most *similar* to the target, rather than most different. If this was, indeed, the case, it suggests that other low scorers might have responded in a similar, incorrect fashion.

Closer inspection of the data revealed 23 suspect scores on Social Translations. The criteria for identifying scores of participants who had likely misunderstood

instructions were an average score of 1.5 or less out of a possible 12 on the two halves of the test, indicating that participants performed at well below a chance level, or a score of zero on either half of the test. When *t*-tests were conducted to determine whether these 23 participants were different from the rest of the sample on other variables, only one statistically significant difference was found. The 23 participants with suspect Social Translations scores rated themselves higher in musical ability on a task in which they compared their own abilities to those of other university students ($M = 66.48$, $SD = 24.14$) than did the rest of the sample ($M = 52.81$, $SD = 24.26$), $t(196) = 2.54$, $p = .01$. Given that there was only one statistically significant difference in 40 *t*-tests, it was concluded that the 23 participants were not significantly different from the rest of the sample, and thus, it was considered reasonable to substitute the Social Translations mean for the actual, presumably inaccurate, score attained by these 23 participants. This imputation procedure avoided the loss of data from 23 participants, had their scores been coded as missing and listwise deletion been used. However, by imputing the mean, variance was reduced and correlations between Social Translations and other variables are likely to be attenuated, as are factors derived from correlations between Social Translations and other variables.

Measured Multiple Intelligences

Descriptive Statistics. Table 3 shows descriptive statistics for the variables in this study (similar descriptive statistics prior to the means substitution and EM are presented in Appendix E). The WPT mean in this study was within one standard deviation of the mean WPT score (21.06, $SD = 7.12$) on a 1992 normative study ($n = 118,549$; Wonderlic,

Table 3

Descriptive Statistics of Measures of Each Intelligence Domain and g

Measure	Intelligence Domain	Min.	Max.	Mean	SD
Wonderlic	g	9.00	44.00	26.62	5.85
Opposites	Linguistic	9.00	40.00	23.98	5.76
Vocabulary	Linguistic	5.00	33.50	19.36	5.32
Map Planning	Spatial	1.00	39.00	22.90	7.40
Paper Folding	Spatial	0.26	20.00	10.50	4.00
Subt/Mult	Logical/Math	5.00	115.00	41.60	15.82
Necessary Arithmetic	Logical/Math	0.34	28.68	14.32	5.92
Cartoon Predictions	Interpersonal	10.00	28.00	20.86	3.02
Social Translations	Interpersonal	4.00	23.00	17.36	3.86
Accuracy	Intrapersonal	-0.75	0.93	0.23	0.41
Consistency	Intrapersonal	-3.28	-0.91	-1.77	0.38
Diag. Rel.	Naturalistic	0.00	30.00	16.92	6.94
Making Groups	Naturalistic	5.00	29.00	18.86	4.58
Stork Stand	Bodily-Kin.	0.00	120.00	32.74	31.27
Mark Making	Bodily-Kin.	46.00	120.00	77.84	10.52
Rhythm	Musical	8.00	19.00	13.79	2.31
Tonal	Musical	7.00	17.00	11.78	2.39

Note. Min. = Minimum; Max. = Maximum; *SD* = Standard Deviation; Subt/Mult = Subtraction and Multiplication; Logical/Math = Logical/Mathematical; Diag. Rel. = Diagramming Relationships; Bodily-Kin. = Bodily-Kinesthetic.

N = 200

2002), and corresponds to a full-scale WAIS IQ score of 114. In addition, the mean WPT score in the present study seems to correspond well to that of male college students aged 20 to 30 ($M = 26.51$, $SD = 6.03$) and female college students ($M = 24.95$, $SD = 5.59$) in the normative study.

The minimum score of 0.00 for Diagramming Relationships could suggest that some participants misunderstood the task. Further investigation, however, shows that only one individual received scores of 0.00 on both halves of the Diagramming Relationships task. Removal of this score had no significant impact on subsequent factor

analyses or correlations between ability tests, so the score was retained. All other tests appear to have means that are reasonably close to the hypothesized midpoints, and standard deviations that are fairly large, suggesting sufficient variability in the tests.

A correlation can only have limits of -1 and +1 when the distributions of the variables are identical (Cohen et al., 2003). An examination of histograms revealed that the distributions of scores on most ability tests appeared to be relatively normal. Scores on the Stork Stand, however, tended to peak at the lower scores, rather than at the mean, although skewness (1.47) was still within the acceptable parameters of -2 to +2 (George & Mallery, 1999). Fourteen of the 17 ability tests had their skewness statistic within the excellent range of -1 to +1 (George & Mallery, 1999) with the remaining three (Multiplication and Subtraction, Social Translations, and the previously mentioned Stork Stand) within the acceptable range. In terms of kurtosis, distributions for all but four of the 17 ability tests fell into the excellent range of -1 to +1 (George & Mallery, 1999). The kurtosis of Mark Making and Stork Stand were both in the acceptable range of -2 to +2 (George & Mallery, 1999). The kurtosis statistics for Multiplication and Subtraction (2.93) and Social Translations (2.62), however, indicate that these distributions are leptokurtic, or too tall. Departures from normality in the distributions of variables tend to attenuate correlations with other variables, particularly when the two variables are skewed in opposite directions (Cohen et al., 2003). Thus correlations between the Stork Stand, Multiplication and Subtraction, and Social Translations, and other variables are likely to be underestimates of true relationships. In addition, factors derived from correlations between variables with differing distributions are also likely to be attenuated (Goldberg & Digman, 1994).

Independent samples *t*-tests were conducted to determine whether there were differences in performance between the two orders of presentation of ability tests. Only Making Groups (Naturalistic) was significantly different between the two orders, with participants receiving materials in order 1 ($n = 100$) scoring higher ($M = 10.09$, $SD = 2.11$) than those receiving materials in the second order ($n = 99$; $M = 8.44$, $SD = 3.60$), $F(1, 196) = 16.33$, $p < .001$. The Levene's test was not significant for any of the comparisons, suggesting that the assumption of homogeneity of variances was not contravened for any of the significance tests.

Reliabilities. Correlations can be attenuated by unreliability in either or both variables (Cohen et al., 2003). The split-half reliability coefficients for the ability tests (Table 4) ranged from .17 (Tonal) to .94 (Subtraction and Multiplication). DeVellis (1991) proposed the following guidelines for internal consistency coefficients of research instruments: below .60 is unacceptable; .60 to .65 is undesirable; .65 to .70 is minimally acceptable; .70 to .80 is respectable; .80 to .90 is very good; and, much above .90 suggests that consideration should be given to shortening the scale. Using these guidelines, both music tests, Consistency (Intrapersonal), and Cartoon Predictions (Interpersonal) had unacceptable split-half reliabilities, but all other tests had split-half reliabilities that were at least minimally acceptable according to DeVellis' guidelines.

The poor reliability of both music tests was surprising, given the high split-half reliabilities reported in the AMMA manual (.80 for both tonal and rhythm in undergraduate and graduate non-music majors). A very small part of this discrepancy might be attributed to methodological differences in calculating the split-half reliability. Gordon (1989) computed item difficulty and discrimination levels prior to dividing the

Table 4

Split-Half Reliability Coefficients of Ability Tests

Domain	Test	Reliability
Linguistic	Opposites	0.67
	Vocabulary	0.67
Spatial	Map Planning	0.80
	Paper Folding	0.76
Logical/Mathematical	Subraction and Multiplication	0.94
	Necessary Arithmetic	0.81
Interpersonal	Cartoon Predictions	0.45
	Social Translations	0.90
Intrapersonal	Consistency	0.53
	Accuracy	N/A
Naturalistic	Diagramming Relationships	0.82
	Making Groups	0.79
Bodily-Kinesthetic	Stork Stand	0.92
	Mark Making	0.93
Musical	Rhythm	0.28
	Tonal	0.17

test into two equal halves, matched by difficulty levels and discrimination values. Gordon did not provide difficulty and discrimination information in the manual, so the method could not be replicated with the present data set. For the current study, split-half reliability of Rhythm was determined by calculating the correlation between the first and second sets of 10 rhythm items on the AMMA. The split-half reliability of Tonal was calculated similarly with tonal items.

The split-half reliability of Cartoon Predictions was disappointing, and suggests that future investigations should employ a more reliable nonverbal measure of interpersonal intelligence. While the split-half reliability of Consistency also fell into the unacceptable range, it should be remembered that this measure was an attempt, perhaps

the first of its kind, at measuring intrapersonal intelligence via a consistency-based measure. The Consistency measure could be refined for future studies, perhaps through minor adjustments to the personality descriptors. No split-half reliability was calculated for Accuracy of self-ratings, the other Intrapersonal measure, because it did not include multiple items for which internal consistency or split-half reliability could be calculated.

All remaining split-half reliability coefficients were .67 or greater, with almost half (47%) of the ability tests having coefficients of .80 or greater.

Measured Abilities: Correlations with Overall Intelligence and g-Loadings. In this study, correlations with the WPT and g-loadings determined through factor analysis were considered to be indicators of saturation with general intelligence. All cognitive measures (i.e., Linguistic, Spatial, Logical/Mathematical, Interpersonal, and Naturalistic tests), as hypothesized, were significantly correlated with the WPT (Table 5). Note that if the 23 inaccurate Social Translations scores had been deleted, rather than preserved through means substitution, the correlation between Social Translations and the WPT would have been only slightly reduced, from .37 to .34. Only the Bodily-Kinesthetic and Musical tests, as well as one of the Intrapersonal tests (Accuracy), were not positively and significantly correlated with the WPT. Necessary Arithmetic Operations and Diagramming Relationships had the greatest correlations with the WPT. Using Cohen's (1988) guidelines in which a correlation of .10 constitutes a small effect size, .30 a medium effect size, and .50 a large effect size, the majority of the correlations with the WPT (Opposites, Vocabulary, Map Planning, Paper Folding, Subtraction and Multiplication, Social Translations, and Making Groups) would be considered medium effect sizes. Of the remaining significant correlations between ability tests and the WPT,

Table 5

g Loadings of Tests and Correlations of Tests with Wonderlic Personnel Test

Intelligence Domain	Test	g-Loading			Correlation with WPT
		PCA	PAF	ML	
Linguistic	Opposites	.56	.50	.48	.39**
	Vocabulary	.61	.54	.51	.46**
Spatial	Map Planning	.61	.55	.57	.46**
	Paper Folding	.56	.50	.53	.45**
Logical/Math	Subt/Mult	.29	.24	.26	.34**
	Necessary Arithmetic	.73	.70	.71	.66**
Interpersonal	Cartoon Predictions	.43	.36	.35	.16*
	Social Translations	.59	.53	.51	.37**
Intrapersonal	Consistency	.33	.27	.26	.21**
	Accuracy	.20	.16	.13	.13
Naturalistic	Diag. Rel.	.77	.75	.76	.57**
	Making Groups	.63	.57	.56	.35**
Bodily-Kinesthetic	Stork Stand	.04	.03	.03	-.04
	Mark Making	.08	.06	.04	.01
Musical	Rhythm	.22	.18	.18	.10
	Tonal	.13	.10	.11	.04

Note. PCA = Principal Components Analysis; PAF = Principal Axis Factoring; ML = Maximum Likelihood; WPT = Wonderlic Personnel Test; Logical/Math = Logical/Mathematical; Subt/Mult = Subtraction and Multiplication; Diag. Rel. = Diagramming Relationships.

N = 200

p* < .05. *p* < .01, two-tailed.

Cartoon Predictions, and Consistency are small effect sizes, and Necessary Arithmetic Operations and Diagramming Relationships are large.

As hypothesized, the correlations among the ability tests showed a positive manifold (Table 6), with over 70% of the 45 correlations between cognitive abilities being significant and positive. About two-thirds of these significant correlations fell into Cohen's (1988) category of medium effect sizes. Less than 20% of the correlations in

which at least one of the variables was non-cognitive reached statistical significance, however.

Factor analysis was conducted using exploratory factor analysis. Although a particular model was investigated in this study, confirmatory factor analysis (CFA) was not used to determine whether the data fit Gardner's MI model. There are two reasons why CFA was not applied. First, given the time constraints involved, only two tests of each of Gardner's intelligences were administered, meaning that there were an insufficient number of variables for Gardner's eight-factor solution to constitute a goodfit, even if MI theory accurately described the data. Second, with a complex data set such as the current one, CFA is likely to reject the model due to non-zero secondary loadings (McCrae, Zonderman, Costa, Bond, & Paunonen, 1996).

Three alternative methods of factor analysis were conducted: Principal Axis Factoring, Principal Components Analysis, and Maximum Likelihood Estimation. All three commonly used methods of factor extraction were employed to demonstrate that results were not specific to any one method; in fact, all three methods yielded similar results, with cognitive tests loading onto a single factor (Table 5). Factor loadings from the Principal Axis Factoring solution will be discussed, because this method is most conservative in that it extracts factors on the basis of a correlation matrix whose diagonal values are the squared multiple correlation of each variable and all the other variables. In this method, only the common variance of each variable is used in the extracted factor.

Eigenvalues for the first seven factors were: 3.74 (23.39% of total variance), 1.64 (10.26%), 1.41 (8.81%), 1.19 (7.42%), 1.13 (7.03%), 0.96 (6.01%), and 0.90 (5.63%). A clear break in eigenvalue sizes is observed between the second and first factors only, thus

suggesting only one large factor in this data set.

The *g*-loadings on the first unrotated factor ranged from .03 (Stork Stand) to .75 (Diagramming Relationships). In addition to the Stork Stand, Mark Making also had a factor loading under .10, suggesting that the Bodily-Kinesthetic tests do not share this factor's common element of general intelligence. Tests with loadings between .10 and .19 include Tonal (Musical), Accuracy (Intrapersonal), and Rhythm (Musical). Tests with loadings between .20 and .39 include Subtraction and Multiplication (Logical/Mathematical), Consistency (Intrapersonal), and Cartoon Predictions (Interpersonal). Those tests with factor loadings between .40 and .60 included Opposites (Linguistic), Paper Folding (Spatial), Social Translations (Interpersonal), Vocabulary (Linguistic), Map Planning (Spatial), and Making Groups (Naturalistic). Necessary Arithmetic Operations (Logical/Mathematical) and Diagramming Relationships (Naturalistic) both had loadings of .70 and greater. Note that if the 23 inaccurate Social Translations scores had been removed entirely, the pattern of *g*-loadings would have remained quite similar. Consistency and Social Translations would be most greatly affected by removing these scores, with their *g*-loadings being reduced from .27 to .20 and .53 to .47 respectively.

Coherence of Abilities within Domains, Beyond g. The coherence of the tests within each intelligence domain was examined through an inspection of zero-order correlations within and between domains (Table 6). Each of the Linguistic tests had its highest correlation with the other within-domain test. Subtraction and Multiplication had

Table 6

Correlations Among Ability Measures

Measure	Vocab.	Map Plan.	Paper Fold.	Subt/ Mult	Nec. Arith.	Cart. Pred.	Social Tran.	Accur. Cons.	Diag. Rel.	Mak. Grps	Stork Stand	Mark Rhythm	Tonal Mak.		
Opp.	.55**	.10	.07	.06	.32**	.22**	.24**	.18**	.10	.38**	.40**	.01	-.05	.03	.03
Vocab		.10	.08	.10	.36**	.32**	.39**	.18*	.17*	.37**	.27**	-.05	-.08	.18*	.03
Map Plan			.45**	.26**	.46**	.09	.23**	.05	.21**	.44**	.40**	.04	.17*	.04	.10
Paper Fold				.09	.45**	.20**	.13	.00	.10	.43**	.31**	.07	.01	.11	-.02
Subt/Mult					.32**	-.09	.09	.05	.12	.18**	.06	-.02	.18*	-.01	.05
Nec. Arith.						.23**	.31**	-.02	.18*	.55**	.32**	.02	-.07	.14	.06
Cart. Pred.							.28**	.23**	.11	.22**	.19**	-.01	-.01	.07	.02
Soc. Tran.								.22**	.19**	.39**	.28**	-.09	.02	.14	.04
Accuracy								.01	.11	.11	.08	.09	.11	.00	-.06
Consist.									.15*	.19**	.19**	-.05	.06	-.02	-.03
Diag. Rel.										.42**	.04	.03	.15*	.13	
Mak. Grps											.05	.10	.04	.04	
Stork Stand												.13	.18*	.16*	
Mark Mak.													.06	.14*	
Rhythm														.19**	

Note. Opp. = Opposites; Vocab. = Vocabulary; Map Plan. = Map Planning; Paper Fold. = Paper Folding; Subt/Mult = Subtraction and Multiplication; Nec. Arith. = Necessary Arithmetic Operations; Cart. Pred. = Cartoon Predictions; Soc. Tran. = Social Translations; Consist. = Consistency; Diag. Rel. = Diagramming Relationships; Mak. Grps = Making Groups; Mark Mak. = Mark Making. Correlations in bold type are those between tasks within the same domain.

N = 200.

* $p < .05$. ** $p < .01$, two-tailed.

its highest correlation with Necessary Arithmetic Operations, the other Logical/Mathematical test, but Necessary Arithmetic Operations was correlated similarly or higher with four other tests. Making Groups had its highest correlation with Diagramming Relationships, the other Naturalistic test, but Diagramming Relationships was more highly correlated with three other tests from the Spatial and Logical/Mathematical domains. The two Interpersonal tests, Social Translations and Cartoon Predictions, were significantly correlated with each other, but each had higher correlations with other tests. Only the Intrapersonal and Bodily-Kinesthetic intelligences had their two within-domain measures uncorrelated.

Table 7

Residual Correlations within Each Intelligence Domain after Extraction of g Factor

Ability	Residual Correlation
Linguistic (Opposites, Vocabulary)	.29**
Logical/Mathematical (Necessary Arithmetic, Subtraction and Multiplication)	.15*
Spatial (Map Planning, Paper Folding)	.18*
Interpersonal (Cartoon Predictions, Social Translations)	.09
Intrapersonal (Accuracy, Consistency)	-.04
Naturalistic (Making Groups, Diagramming Relationships)	-.01
Bodily-Kinesthetic	.13
Musical	.17*

Note. $N = 200$.

* $p < .05$. ** $p < .01$, two-tailed.

In addition to examining the zero-order correlations between tests within an intelligence domain, the residual correlations following factor analysis (i.e., with the removal of the g factor) were examined (Table 7). This allows an evaluation of the extent to which measures within each domain are related beyond the influences of g . The residual correlations ranged from $-.04$ (Intrapersonal) to $.29$ (Linguistic). The Intrapersonal and Bodily-Kinesthetic tests were unrelated to each other after the removal of the g factor. While the two Naturalistic tests were substantially correlated (Table 6), their shared variance would appear to be largely due to their g -loadings as there was only a small residual correlation. The small residual correlation for Bodily-Kinesthetic, identical in size to the zero-order correlation between these two variables, suggests that the little variance these variables share is likely domain-specific and unrelated to g . Spatial, Logical/Mathematical, Musical, and Linguistic tests had residual correlations that suggest shared variance above and beyond g . The large residual correlation of the Linguistic tests, in particular, suggests that despite their substantial g -loadings, these two tests shared a significant amount of variance unrelated to g .

Self-Estimated Intelligences

Descriptive Statistics (Table 8). Recall that participants completed direct self-ratings in which they ranked themselves in comparison to other university students. In addition, they completed the MAQ self-report questionnaire, an indirect method of having participants provide self-estimates.

In the direct self-ratings, participants rated themselves as being superior to most university students in every intelligence domain. On average, participants rated

Table 8

Descriptive Statistics of Direct Self-Ratings and MAQ Self-Estimates

Measure	Min.	Max.	Mean	SD
Self-Ratings				
Linguistic	10.00	100.00	67.60	18.85
Logical/Mathematical	.00	100.00	55.32	24.16
Musical	.00	100.00	54.37	24.47
Spatial	.00	100.00	61.32	21.17
Bodily-Kinesthetic	.00	100.00	64.21	21.61
Intrapersonal	10.00	100.00	71.75	18.59
Interpersonal	10.00	100.00	70.05	17.32
Naturalistic	.00	100.00	60.99	20.24
Overall Intelligence	10.00	100.00	69.76	17.11
Multiple Ability Questionnaire				
Linguistic	1.83	5.00	3.94	.62
Logical/Mathematical	1.33	5.00	3.52	.76
Spatial	1.83	5.00	3.51	.66
Intrapersonal	1.83	5.00	3.54	.61
Interpersonal	2.00	5.00	3.93	.58
Bodily-Kinesthetic	1.50	5.00	3.54	.74
Musical	1.33	5.00	3.65	.88
Naturalistic	1.00	5.00	2.91	.77

Note. Min. = Minimum; Max. = Maximum.
N = 200.

themselves highest in Intrapersonal (greater than 72% of university students) and Interpersonal (70%) intelligence and lowest in Musical (54%) and Logical/Mathematical (55%) intelligence. Standard deviations for the intelligence domains ranged from 17 for Interpersonal to 25 for Musical, suggesting there was sufficient variability in responses.

An examination of the distributions of the self-ratings revealed that Musical and Logical/Mathematical self-ratings appear to be quite normally distributed. The remaining self-ratings appeared to be somewhat negatively skewed, in keeping with the finding that

participants tended to rate themselves as well above average in these domains. Skewness and kurtosis statistics, however, were in the excellent range of -1 to +1 for all self-ratings except Interpersonal, for which both skewness and kurtosis statistics fell into the satisfactory range of -2 to +2.

Participants responded to MAQ items on a 5-point scale, from 1 (*Very Inaccurate*) to 5 (*Very Accurate*). Participants estimated themselves lowest in Naturalistic ability and highest in Linguistic ability (Table 8). Standard deviations ranged from .58 to .88, suggesting sufficient variability in MAQ responses.

An examination of the distributions of MAQ responses indicated that all distributions approximated normality. Distributions for Musical, Interpersonal, and Linguistic, appeared slightly negatively skewed but skewness and kurtosis statistics for all MAQ variables fell into the excellent range of -1 to +1.

Table 9

Internal Consistency Reliabilities of Multiple Ability Questionnaire Intelligence Domains

MAQ Domain	Cronbach's Alpha
Linguistic	0.58
Logical/Mathematical	0.72
Spatial	0.57
Interpersonal	0.55
Intrapersonal	0.51
Naturalistic	0.32
Bodily-Kinesthetic	0.64
Musical	0.82

Note. $N = 200$.

Reliabilities. Cronbach's alpha internal consistency reliability coefficients for the self-reported abilities of the Multiple Ability Questionnaire (see Table 9) ranged from .32 (Naturalistic) to .82 (Musical) with a mean of .59. It seems possible that the Naturalistic items, reflecting awareness of and distinctions between clouds, birdcalls, car models, planets, trees, and the signs of fall, do not reflect a unitary underlying ability, although they all seem to be relevant to the Naturalistic domain as Gardner (1999) described it. The MAQ music items—clapping a beat, detecting pitch, recalling a tune, singing off-key, tapping one's foot, and being a musical person—might simply reflect a more unitary construct.

Within-Instrument Correlations. Correlations among the direct self-ratings (Table 10) ranged from .09 (Musical and Naturalistic) to .62 (Linguistic and Overall). It is noteworthy that, given the oft-cited association between math and music, participants clearly did not perceive their own abilities in these two domains as similar ($r = .10$). Self-rated abilities that were correlated at .40 or greater were as follows: Interpersonal and Intrapersonal; Spatial and Naturalistic; Linguistic and Intrapersonal; Spatial and Logical/Mathematical; Spatial and Bodily-Kinesthetic; Linguistic and Interpersonal; Spatial and Interpersonal; Logical/Mathematical and Naturalistic; Interpersonal and Naturalistic; Interpersonal and Bodily-Kinesthetic; Intrapersonal and Bodily-Kinesthetic; and Overall Intelligence with every other ability. All self-rated abilities, with the exceptions of the Naturalistic and Musical and Logical/Mathematical and Musical, were correlated at .22 or greater, suggesting that most people report somewhat consistent levels of ability across domains.

Correlations among the MAQ scales (Table 10) were generally weaker than

Table 10

Within-Instrument Correlations for Ability Self-Ratings and Multiple Ability Questionnaire Scales

	Ling.	Spat.	Math	Inter.	Intra.	Nat.	B-K	Mus.	Overall
Self-Ratings									
Ling.		.32**	.36**	.44**	.48**	.37**	.30**	.26**	.62**
Spat.			.47**	.41**	.29**	.51**	.47**	.22**	.47**
Math				.24**	.28**	.41**	.31**	.10	.60**
Inter.					.51**	.42**	.40**	.25**	.46**
Intra.						.25**	.40**	.23**	.47**
Nat.							.26**	.09	.44**
B-K								.28**	.41**
MAQ									
Ling.		.18*	.28**	.31**	.35**	.08	.10	.17*	
Spat.			.46**	.23**	.22**	.32**	.43**	.08	
Math				.15*	.13	.22**	.34**	.05	
Inter.					.26**	.06	.28**	.26**	
Intra.						.02	.23**	.15*	
Nat.							.06	.08	
B-K								.26**	

Note. Ling. = Linguistic; Spat. = Spatial; Math = Logical/Mathematical; Inter. = Interpersonal; Intra. = Intrapersonal; Nat. = Naturalistic; B-K = Bodily-Kinesthetic; Mus. = Musical; Overall = Overall Intelligence (there were no items for Overall on the MAQ).

N = 200.

p* < .05. *p* < .01, two-tailed.

correlations among the direct self-ratings. Correlations ranged from .02 (Intrapersonal and Naturalistic) to .46 (Logical/Mathematical and Spatial), with 12 of the 28 correlations less than .20. The lower inter-correlations for the MAQ abilities are likely due to the fact that self-estimates of several facets of each ability were combined to arrive at an overall mean self-estimate for each intelligence domain. The less complex, overall estimate of ability provided on the self-ratings might be more likely to be influenced by self-esteem and response biases, which would be common to all the self-ratings.

Table 11

Loadings of Direct Self-Ratings of Multiple Intelligences in First Unrotated Factor

	Factor Analytic Method		
	PCA	PAF	ML
Linguistic	.69	.62	.62
Spatial	.73	.68	.67
Logical/Mathematical	.61	.54	.54
Interpersonal	.72	.67	.67
Intrapersonal	.67	.60	.60
Naturalistic	.65	.59	.60
Musical	.41	.34	.34
Bodily-Kinesthetic	.66	.59	.59

Note. PCA = Principal Components Analysis; PAF = Principal Axis Factoring; ML = Maximum Likelihood.
N = 200.

Factor analysis of the single-item direct self-ratings (Table 11) using Principal Axis Factoring yielded a single factor, with loadings ranging from .34 (Musical) to .68 (Spatial). This result is not specific to the Principal Axis Factoring method, as the results were similar using Principal Components Analysis and Maximum Likelihood methods.

When the same procedure was applied to the MAQ (Table 12), all self-estimates again loaded onto a single factor, with loadings ranging from .27 (Naturalistic) to .65 (Spatial), again suggesting that people's self-reported abilities are in large part influenced by a general factor of self-perceived intelligence.

Table 12

Loadings of Multiple Ability Questionnaire Scales in First Unrotated Factor

MAQ Domain	Factor Analytic Method		
	PCA	PAF	ML
Linguistic	.54	.43	.38
Spatial	.71	.65	.69
Logical/Mathematical	.63	.55	.59
Interpersonal	.58	.46	.42
Intrapersonal	.52	.41	.37
Naturalistic	.35	.27	.30
Musical	.40	.29	.26
Bodily-Kinesthetic	.65	.57	.58

Note. PCA = Principal Components Analysis; PAF = Principal Axis Factoring; ML = Maximum Likelihood.
N = 200.

Correlations Between Direct Self-Ratings and MAQ Scales. Correlations between direct ratings of multiple intelligences and the corresponding indirect MAQ self-estimates (Table 13) were all positive and significant, ranging from .28 (Naturalistic) to .72 (Music), with a mean correlation of .52, suggesting that the two scales contained substantial overlap.

The 56 correlations between non-similar domains ranged from -.11 (MAQ Music

Table 13

Correlations Between Ability Self-Ratings and Multiple Ability Questionnaire Scales

	Self-Ratings								
	Ling.	Spat.	Math	Inter.	Intra.	Nat.	B-K	Mus.	Overall
MAQ									
Ling.	.47**	.20**	.15*	.38**	.27**	.16*	.09	.17*	.33**
Spat.	.15*	.51**	.30**	.17*	.04	.28**	.20**	.00	.27**
Math	.24**	.41**	.70**	.10	.16*	.29**	.26**	.08	.42**
Inter.	.19**	.21**	.05	.33**	.20**	.08	.20**	.10	.18*
Intra.	.26**	.12	.11	.30**	.52**	.10	.10	.13	.18*
Nat.	.19**	.21**	.13	.11	.04	.28**	.07	.09	.16*
B-K	.06	.36**	.20**	.26**	.19**	.12	.59**	.18**	.19**
Music	.11	.10	-.11	.17*	.09	-.01	.19**	.72**	.06

Note. Ling. = Linguistic; Spat. = Spatial; Math = Logical/Mathematical; Inter. = Interpersonal; Intra. = Intrapersonal; Nat. = Naturalistic; B-K = Bodily-Kinesthetic; Mus. = Musical; Overall = Overall Intelligence (there were no items for Overall on the MAQ).

N = 200.

p* < .05. *p* < .01, two-tailed.

and self-rated Logical-Mathematical) to .41 (MAQ Logical/Mathematical and self-rated Spatial). Thirty-eight of these 64 correlations between non-similar domains were positive and significant, suggesting that even when using different methods to obtain self-report measures of different abilities, there is much shared variance.

Correlations Between Self-Estimated Abilities and Measured Multiple Intelligences. To determine the accuracy of self-estimated abilities, correlations between

performance on ability tests and both direct self-ratings of ability and MAQ self-estimates were calculated. Direct self-ratings were correlated with actual performance only for some domains (Table 14). Self-rated Linguistic, Logical/Mathematical, and Musical abilities were positively and significantly correlated with both tests from the corresponding relevant domains. It seems likely that these three domains are ones in which people have received feedback regarding their performance. Self-rated Bodily-Kinesthetic ability was significantly correlated with the Stork Stand, a gross motor task, but not with Mark Making, a fine motor measure, although this self-rating item was worded to include both kinds of physical ability (“Coordinating your physical movements [for example, in sports, dancing, crafts, technical work]”).

Self-rated Intrapersonal ability was uncorrelated with Accuracy, and significantly but *negatively* correlated with Consistency. Correlations between Consistency and self-ratings of the other abilities were all negative as well, with several reaching statistical significance.

Self-rated Interpersonal ability had small correlations with both Interpersonal measures, but only its correlation with Cartoon Predictions reached significance. Self-rated Naturalistic and self-rated Spatial abilities were both uncorrelated with their associated ability measures.

Self-rated intelligence was significantly correlated with the WPT, but at $r = .19$, it was a smaller correlation than the .39 value that Chamorro-Premuzic, Furnham, and Moutafi (2004) reported between self-estimated intelligence and performance on the WPT in their study.

Table 14

Correlations Between Self-Ratings and Intelligence Measures

Self- Rating	Opp.	Voc.	Map	Paper	Subt/ Mult	Nec.	Cart.	Social	Accur.	Cons.	Diag.	Mak.	Stork	Mark	Rhythm	Tonal	WPT
			Plan.	Fold.		Arith.	Pred.	Tran.			Rel.	Grps	Stand	Mak.			
Ling.	.25**	.29**	-.09	-.12	-.06	.09	.17*	.23**	.13	-.17*	.18**	.09	.01	.05	.16*	.09	.15*
Spatial	.06	.10	.04	.05	.02	.05	.02	-.02	.09	-.19**	.04	.07	.20**	-.02	-.09	.08	.08
Math	.11	.09	.15*	.16*	.28**	.34**	.24**	.09	.00	-.14	-.01	.05	.02	-.03	.06	.09	.29**
Inter.	.44**	.01	-.09	-.11	-.04	-.15*	.15*	.10	.03	-.21**	-.09	.07	.03	.10	.02	.00	-.12
Intra.	.13	.06	-.04	-.10	.02	.02	.17*	.02	.06	-.20**	.06	.32**	-.03	-.02	.07	.07	.06
Nat.	-.01	.08	-.05	-.05	-.03	-.05	.08	-.04	-.03	-.07	-.04	.02	.00	-.01	.00	.05	-.06
B-K	-.08	-.12	.05	.04	.03	-.04	-.06	-.07	.01	-.22**	-.05	-.02	.31**	-.00	.01	.11	-.03
Mus.	.09	.05	-.22**	.05	-.17*	-.13	-.08	-.04	.04	-.17*	-.09	-.06	.20**	.06	.21**	.18*	-.10
Overall	.19**	.09	.01	.01	.07	.12	.13	.13	.05	-.19**	.13	.01	.04	.05	.02	.08	.19**

Note. Opp. = Opposites; Vocab. = Vocabulary; Map Plan. = Map Planning; Paper Fold. = Paper Folding; Subt/Mult = Subtraction and Multiplication; Nec. Arith. = Necessary Arithmetic Operations; Cart. Pred. = Cartoon Predictions; Soc. Tran. = Social Translations; Diag. Rel. = Diagramming Relationships; Mak. Grps = Making Groups; Mark Mak. = Mark Making; Ling. = Linguistic; Math = Logical/Mathematical; Inter. = Interpersonal; Intra. = Intrapersonal; Nat. = Naturalistic; B/K = Bodily-Kinesthetic; Mus. = Musical; Overall = Overall Intelligence.

Correlations in bold indicate those tests within the same ability domain.

N = 200.

* $p < .05$. ** $p < .01$, two-tailed.

Correlations between MAQ self-estimates and test performance (Table 15) were similar to correlations between direct self-ratings and measured ability. As with direct self-ratings, MAQ self-estimates of Linguistic, Logical/Mathematical, and Musical abilities were significantly and positively correlated with their corresponding ability measures.

MAQ self-estimated Bodily-Kinesthetic ability was correlated with the Stork Stand but not Mark Making, in keeping with results from the self-ratings. Because only two of the six Bodily-Kinesthetic items on the MAQ related specifically to fine motor ability, correlations between these two items and Mark Making were examined. Both items, "I'm good at making things with my hands", and "I have some difficulty shuffling a deck of cards," had near zero correlations with Mark Making performance ($r = .01$ and $r = .00$ respectively).

One test from each of Spatial, Interpersonal, and Bodily-Kinesthetic domains was significantly and positively correlated with the corresponding MAQ self-estimate, whereas Intrapersonal and Naturalistic measures were uncorrelated with their relevant MAQ self-estimates. Consistency was correlated negatively with all MAQ self-estimates, just as it was correlated negatively with all direct self-ratings. Consistency's negative correlations with MAQ Logical/Mathematical, Interpersonal, and Musical self-estimates were of statistical significance.

Paulhus et al. (1998) found that items with content related to enjoyment and frequency of reading were highly associated with measured intelligence, so the correlation between the MAQ item, "I don't care for reading" (reverse-keyed), and the WPT was examined, but it was found to be small in effect size at $r = .18$.

Table 15

Correlations Between Multiple Ability Questionnaire Scales and Intelligence Measures

MAQ	Opp.	Vocab.	Map	Paper	Subt/	Nec.	Cart.	Social	Cons.	Accur.	Diag.	Mak.	Stork	Mark	Rhythm	Tonal	WPT
		Plan.	Fold.	Mult.	Arith.	Pred.	Tran.				Rel.	Grps	Stand	Mak.			
Ling.	.36**	.36**	-.06	-.06	.09	.07	.19**	.17**	-.07	.09	.12	.20**	.04	.07	.01	-.04	.20**
Spatial	.06	.01	.20**	.13	.07	.09	-.04	-.10	-.09	-.06	.13	.13	.10	.05	-.07	.00	.11
Math	.20**	.14	.17*	.22**	.29**	-.01	.07	.07	-.20**	-.01	.25**	.12	.07	-.05	.02	.02	.32**
Inter.	.02	.01	-.09	-.11	-.01	-.08	.02	.11	-.26**	-.11*	.01	.07	.02	-.04	-.05	.02	-.06
Intra.	.11	.07	-.09	-.15*	-.03	-.10	.09	.04	-.06	-.01	.02	.15	.02	.15*	.03	-.03	-.03
Nat.	.14	.19**	.11	.08	-.05	.05	.17*	.09	-.03	.17*	.10	.09	-.05	-.10	-.01	.04	.08
B-K	-.14	-.13	.12	.21**	-.03	-.02	-.11	-.10	-.12	-.08	-.03	.03	.28**	.02	.11	.05	-.02
Mus.	.03	.03	-.18**	.05	-.19**	-.21**	.03	.02	-.19**	.00	-.11	-.06	.19**	.08	.26**	.20**	-.12

Note. Opp. = Opposites; Vocab. = Vocabulary; Map Plan. = Map Planning; Paper Fold. = Paper Folding; Subt/Mult = Subtraction and Multiplication; Nec. Arith. = Necessary Arithmetic Operations; Cart. Pred. = Cartoon Predictions; Soc. Tran. = Social Translations; Diag. Rel. = Diagramming Relationships; Mak. Grps = Making Groups; Mark Mak. = Mark Making; WPT = Wonderlic Personnel Test; Ling. = Linguistic; Math = Logical/Mathematical; Inter. = Interpersonal; Intra. = Intrapersonal; Nat. = Naturalistic; B-K = Bodily-Kinesthetic; Mus. = Musical.

Correlations in bold indicate those tests within the same ability domain.

N = 200.

* $p < .05$. ** $p < .01$, two-tailed.

Personality

Descriptive Statistics (Table 16). Recall that participants indicated on a 9-point scale how accurately several adjectives described their personality. Ratings on these adjectives were then averaged to calculate a score from 0 to 9 for each of six personality factors. On average, participants reported themselves highest in Openness to Experience and lowest in Agreeableness. All means were above the scale midpoint, but standard deviations were sufficiently large to suggest there was adequate variability in responses.

Visual inspection of the distributions of the six personality variables was conducted, with all six distributions approximating normality. Skewness and kurtosis statistics for all six variables fell into the excellent range of -1 to +1.

Reliabilities. Cronbach's alpha reliability coefficients for the six personality traits ranged from .66 to .86 with a mean of .76 (Table 16), meaning that all were at least minimally acceptable according to DeVellis' (1991) guidelines. Openness to Experience

Table 16

Descriptive Statistics and Internal Consistency Reliabilities of Personality Self-Rating Scales

Personality Factor	Min.	Max.	Mean	SD	Cronbach's Alpha
Openness to Experience	4.25	9.00	6.75	0.98	.66
Conscientiousness	2.75	9.00	6.59	1.21	.81
Agreeableness	2.75	9.00	5.85	1.20	.75
Extraversion	2.63	9.00	6.20	1.37	.86
Emotionality	2.88	8.63	5.93	1.18	.71
Honesty-Humility	2.43	9.00	5.89	1.47	.78

Note. Min. = Minimum; Max. = Maximum.
N = 200.

was the only trait with a reliability coefficient of less than .70, which is DeVellis' cut-off for respectable reliability. Item-total correlations for the Openness to Experience scale ranged from .20 for *shallow* to .51 for *philosophical*. Interestingly, the item *complex* had a .32 item-total correlation, substantially higher than its opposite, *shallow*.

Correlations Among Personality Factors. There were several significant correlations among the six personality constructs (see Table 17). The correlations between constructs tended to be only slightly higher than those reported by Lee and Ashton (2004) in their examination of the psychometric properties of the HEXACO Personality Inventory. For example, the average of the absolute value of correlations Between personality factors was .11 in Lee and Ashton's study, whereas in the current study, the average correlation was .16. Lee and Ashton reported three correlations of a

Table 17

Correlations Between Self-Reported Personality Scales

	Openness	Conscient.	Agree.	Extra.	Emotion.	Hon.-Hum.
Openness		.18**	.11	.21**	.18*	.10
Conscient.			.18*	.11	.19**	.12
Agree.				-.03	.14*	.37**
Extra.					.10	.03
Emotion.						.29**

Note. Openness = Openness to Experience; Conscient. = Conscientiousness; Agree. = Agreeableness; Extra. = Extraversion; Emotion. = Emotionality; Hon.-Hum. = Honesty-Humility.

N = 200.

p* < .05. *p* < .01, two-tailed.

magnitude greater than .20 (Conscientiousness and Honesty-Humility; Openness to Experience and Extraversion; and Agreeableness and Honesty-Humility). In the current study, there were also three correlations above .20 (Table 17), with two of the same correlations found in Lee and Ashton's highest three correlations (Openness to Experience and Extraversion; and Agreeableness and Honesty-Humility).

Table 18

Correlations Between Self-Reported Personality Scales and Ability Measures and Academic Average

Measure	Openness	Conscient.	Agree.	Extra.	Emotion.	Hon.-Hum.
Opposites	.26**	.04	.07	-.02	.00	.06
Vocabulary	.33**	.01	.16*	-.12	.00	.19**
Map Planning	-.09	-.10	.05	-.12	-.22**	-.24*
Paper Folding	-.03	.01	.04	-.10	-.13	-.05
Subt/Mult	-.08	.03	.11	-.16*	-.06	-.04
Nec. Arith.	-.09	-.01	.07	-.02	-.19**	-.04
Cartoon Predictions	.17*	.03	.04	.03	.05	.16*
Social Translations	.16*	.01	.08	.02	.02	.02
Accuracy	.12	-.20**	-.06	-.09	-.11	-.08
Consistency	-.05	-.05	-.01	-.09	.04	.01
Diag. Rel.	.11	.07	.12	-.12	-.12	-.07
Making Groups	.15*	.00	.09	.11	.09	-.09
Stork Stand	.02	.09	.02	.01	-.03	.02
Mark Making	.00	-.01	-.02	.00	.05	-.07
Tone	.01	.07	.06	-.08	.06	-.03
Rhythm	.05	.00	-.01	.08	.05	-.01
WPT	.05	.03	.13	-.05	-.15*	-.03
Academic Average	.14*	.23**	.09	-.11	.11	.08

Note. Openness = Openness to Experience; Conscient. = Conscientiousness; Agree. = Agreeableness; Extra. = Extraversion; Emotion. = Emotionality; Hon.-Hum. = Honesty-Humility; Subt/Mult = Subtraction and Multiplication; Nec. Arith. = Necessary Arithmetic Operations; Diag. Rel. = Diagramming Relationships; WPT = Wonderlic Personnel Test.

N = 200.

p* < .05. *p* < .01, two-tailed.

Correlations with Measured Multiple Intelligences. Correlations between the six personality constructs and ability tests are presented in Table 18. It was hypothesized that Openness to Experience would be significantly and positively correlated only with measures of crystallized intelligence (Linguistic and Interpersonal tests) and not measures of fluid intelligence (Spatial, Naturalistic, and Logical/Mathematical tasks) or with non-cognitive tasks (Musical, Intrapersonal, and Bodily-Kinesthetic measures). As hypothesized, Openness to Experience was significantly and positively correlated with both Linguistic tests (Opposites and Vocabulary) and both Interpersonal tests (Cartoon Predictions and Social Translations). Openness to Experience was uncorrelated with all other ability tests, with the exception of a small but significant correlation with Making Groups (Naturalistic).

It was hypothesized that Conscientiousness would, in keeping with previous research (Busato, Prins, Elshout, & Hamaker, 2000), be significantly and positively correlated only with academic average. This hypothesis was supported, in that Conscientiousness was correlated with self-reported academic average, and also had a significant negative correlation with Accuracy (Intrapersonal). In contrast to two recent studies reporting that Conscientiousness was significantly and negatively correlated with intelligence (Moutafi et al., 2003; Moutafi et al., 2004), Conscientiousness was uncorrelated with the WPT in the current study.

In previous research, measures of emotional intelligence have been associated with the Big Five factors of Extraversion and Emotional Stability. In the current study, both Intrapersonal and Interpersonal measures were uncorrelated with Extraversion.

Table 19

Correlations Between Direct (Self-Rating) and Indirect (MAQ) Self-Estimated Multiple Intelligences and Personality

Ability	Openness	Conscient.	Agree.	Extra.	Emotion.	Hon.-Hum.
Self-Rating						
Linguistic	.35**	.17*	.05	.17*	-.13	-.03
Spatial	.27**	.20**	.12	.13	-.11	-.06
Logical/Mathematical	.00	.23**	.02	-.03	-.28**	-.12
Interpersonal	.32**	.24**	.10	.28**	.05	.02
Intrapersonal	.17*	.36**	.08	.14*	-.11	.03
Naturalistic	.12	.14	-.01	.08	-.07	-.02
Bodily-Kinesthetic	.06	.12	.17*	.21**	-.19**	-.11
Musical	.12	.08	.08	.11	.07	-.06
MAQ						
Linguistic	.54**	.18*	-.03	.27**	.03	-.03
Spatial	.19**	.30**	.01	.11	-.14	-.17*
Logical/Mathematical	.15*	.32**	.03	.06	-.23**	-.12
Interpersonal	.32**	.26**	.38**	.33**	.19**	.17*
Intrapersonal	.18**	.34**	-.05	.24**	-.05	.00
Naturalistic	.09	.07	.10	.02	-.07	.02
Bodily-Kinesthetic	.18*	.25**	.16*	.28**	-.09	-.12
Musical	.22**	.13	.18*	.13	.28**	.10

Note. Openness = Openness to Experience; Conscient. = Conscientiousness; Agree. = Agreeableness; Extra. = Extraversion; Emotion. = Emotionality; Hon.-Hum. = Honesty-Humility.

N = 200.

p* < .05. *p* < .01, two-tailed.

Because the HEXACO personality factors were used, rather than the Big Five, there was no measure that was entirely equivalent to Emotional Stability. The HEXACO Emotionality factor is most similar, but it was not correlated with either Intrapersonal measure.

Correlations with Self-Estimated Multiple Intelligences. The correlations of the six personality factors with both direct self-ratings and indirect MAQ self-estimates are presented in Table 19. Openness to Experience and Conscientiousness were generally related to self-estimated Intrapersonal ability as well as the cognitive abilities (i.e., Linguistic, Spatial, Logical/Mathematical, Interpersonal, and Naturalistic) with the exception of Naturalistic. Openness to Experience was uncorrelated with self-rated Logical/Mathematical ability, however, and both Openness to Experience and Conscientiousness were significantly and positively correlated with MAQ Bodily-Kinesthetic ability.

Honesty-Humility was generally uncorrelated with self-estimated multiple intelligences, with the exception of its significant positive correlation with MAQ Interpersonal and its significant negative correlation with MAQ Spatial. Emotionality was positively correlated with MAQ Musical and Interpersonal abilities, but negatively correlated with both MAQ and self-rated Logical/Mathematical ability and self-rated Bodily-Kinesthetic ability. Agreeableness was significantly and positively correlated with self-rated and MAQ Bodily-Kinesthetic ability, as well as MAQ Interpersonal and Musical abilities. Extraversion was positively correlated with both self-rated and MAQ self-estimates of Linguistic, Interpersonal, Intrapersonal, and Bodily-Kinesthetic abilities.

Sex Differences

Measured Multiple Intelligences (Table 20). Women outperformed men in Making Groups (Naturalistic) and Cartoon Predictions (Interpersonal). Men performed better than women in Necessary Arithmetic Operations (Logical/Mathematical). These differences, while significant, were quite small. The differences in Interpersonal and mathematical reasoning tests were hypothesized, whereas women's slightly greater

Table 20

Means and Standard Deviations for Each Ability Test by Sex, and Results of t-tests of Sex Differences for Each Variable

Measure	Mean (Men) <i>n</i> = 84	<i>SD</i> (Men)	Mean (Women) <i>n</i> = 116	<i>SD</i> (Women)	<i>t</i>	<i>df</i>	<i>p</i>
Opposites	11.99	2.83	12.00	2.92	-0.03	198	.99
Vocabulary	9.62	2.88	9.71	2.50	0.24	198	.81
Map Plan.	11.97	3.68	11.08	3.69	1.68	198	.09
Paper Fold.	5.42	1.94	5.14	2.05	0.98	198	.33
Subt/Mult	22.08	9.22	19.88	6.70	1.86	143	.07
Nec. Arith.	7.65	2.94	6.80	2.94	2.00*	198	.05
Cart. Pred.	10.08	1.66	10.69	1.35	-2.75*	156	.01
Soc. Tran.	8.58	2.08	8.76	1.82	-0.66	198	.51
Accuracy	0.26	0.41	0.20	0.41	0.92	198	.36
Consist.	-1.81	0.37	-1.74	0.38	1.35	198	.18
Diag. Rel.	8.79	3.50	8.23	3.45	1.13	198	.26
Mak. Grps.	8.97	1.95	9.76	2.46	-2.53*	196	.01
Stork Stand	17.55	16.52	15.66	14.98	0.84	198	.40
Mark Mak.	38.11	5.48	39.50	5.05	-1.86	198	.07
Rhythm	13.54	2.36	13.97	2.27	-1.32	198	.19
Tonal	12.15	2.47	11.51	2.30	1.89	198	.06

Note. Map Plan. = Map Planning; Paper Fold. = Paper Folding; Subt/Mult = Subtraction and Multiplication; Nec. Arith. = Necessary Arithmetic Operations; Cart. Pred. = Cartoon Predictions; Soc. Tran. = Social Translations; Consist. = Consistency; Diag. Rel. = Diagramming Relationships; Mak. Grps = Making Groups; Mark Mak. = Mark Making.

**p* < .05, two-tailed.

performance in the Naturalistic measure was unanticipated. Hypothesized sex differences in arithmetic computation, fine motor ability, generating opposites, and visualization did not emerge in this study, nor was there a sex difference in the second Interpersonal test.

Table 21

Means and Standard Deviations for Direct Self-Ratings and Indirect MAQ Self-Estimates by Sex, and Results of t-tests of Sex differences for Each Variable

	Mean (Men) <i>n</i> = 84	<i>SD</i> (Men)	Mean (Women) <i>n</i> = 116	<i>SD</i> (Women)	<i>t</i>	<i>df</i>	<i>p</i>
Self-Rating							
Linguistic	67.74	19.80	67.51	18.22	0.09	198	.93
Spatial	66.47	20.73	57.61	20.78	2.98***	198	.00
Logical/Math	62.89	23.25	49.88	23.44	3.89***	198	.00
Interpersonal	67.55	20.10	71.88	14.81	-1.67	145	.10
Intrapersonal	70.83	20.60	72.42	17.04	-0.59	198	.55
Naturalistic	61.09	20.41	60.92	20.21	0.06	198	.95
Bodily-Kin	69.76	19.15	60.19	22.46	3.16**	198	.00
Musical	57.12	22.32	52.40	25.82	1.38	192	.17
Overall	72.71	17.17	67.64	16.81	2.09*	198	.04
MAQ							
Linguistic	3.88	.63	3.98	.61	-1.20	198	.23
Spatial	3.62	.69	3.44	.64	1.97	198	.05
Logical/Math	3.75	.69	3.35	.77	3.79***	198	.00
Interpersonal	3.82	.64	4.01	.51	-2.23*	154	.03
Intrapersonal	3.39	.64	3.66	.56	-3.12**	198	.00
Naturalistic	3.06	.78	2.79	.75	2.46*	198	.02
Bodily-Kin	3.61	.70	3.49	.77	1.08	198	.28
Musical	3.61	.80	3.68	.93	-0.55	198	.58

Note. Logical/Math = Logical/Mathematical; Bodily-Kin. = Bodily-Kinesthetic; Overall = Overall Intelligence.

p* < .05, *p* < .01, ****p* < .001, two-tailed.

Self-Estimated Multiple Intelligences (Table 21). All sex differences in direct self-ratings were in the direction of men having higher self-rated abilities than had women. Men's self-ratings of Logical/Mathematical, Spatial, Bodily-Kinesthetic, and overall intelligence were significantly greater than were women's. All but the Bodily-Kinesthetic sex difference were hypothesized.

When MAQ self-estimates were examined for sex differences, the results were somewhat different from those of the direct self-ratings. Logical/Mathematical and Naturalistic were the only domain in which men's MAQ self-estimates were significantly higher than women's. Women's MAQ self-estimates were higher than men's in the Intrapersonal and Interpersonal domains. The sex differences in Logical/Mathematical and Interpersonal intelligences were hypothesized, whereas women's higher self-estimated Intrapersonal intelligence and men's higher self-estimated Naturalistic intelligence were not.

Table 22

Means and Standard Deviations for Self-Reported Personality by Sex and Results of t-tests of Sex Differences for Each Variable

Personality Scale	Mean (Men) <i>n</i> = 84	<i>SD</i> (Men)	Mean (Women) <i>n</i> = 116	<i>SD</i> (Women)	<i>t</i>	<i>df</i>	<i>p</i>
Openness	6.68	1.00	6.80	0.96	-.81	198	.42
Conscientiousness	6.49	1.20	6.67	1.21	-1.07	198	.29
Agreeableness	5.79	1.16	5.89	1.23	-.57	198	.57
Extraversion	5.97	1.34	6.36	1.38	-1.98*	198	.05
Emotionality	5.40	1.06	6.32	1.10	-5.90***	198	.00
Honesty-Humility	5.49	1.51	6.18	1.38	-3.38**	198	.00

Note. **p* < .05, ***p* < .01, ****p* < .001, two-tailed.

Personality (Table 22). There were sex differences in three of the six personality factors. Women reported significantly higher levels of Extraversion, Emotionality, and Honesty-Humility. Only the difference in Extraversion was unhypothesized.

General Discussion

Measured Multiple Intelligences

The results of this study provide support for hierarchical theories of intelligence, in that the cognitive tests were all correlated with the measure of general intelligence and all loaded together on a single factor. Tests with the highest saturations with general intelligence, as established both through correlations with a measure of general cognitive ability and through *g*-loadings, were Diagramming Relationships (Naturalistic) and Necessary Arithmetic Operations (Logical/Mathematical). The high *g*-loading of Necessary Arithmetic Operations was expected, because mathematical reasoning is generally assumed to be highly associated with general intelligence. Additionally, it was unsurprising that this test, of all ability tests, had the highest correlation with the WPT, since there was overlap in content, with mathematical reasoning items represented on the WPT.

The emergence of Diagramming Relationships, a Naturalistic test, as having the highest *g*-loading and the second-highest correlation with the WPT was perhaps slightly surprising. This test's correlation with the WPT cannot be explained by overlapping content, as there are no tasks resembling Diagramming Relationships on the WPT. Although this test closely matches Gardner's description of Naturalistic Intelligence, its high *g*-loading can likely be explained by its demand for logical reasoning. In fact, Diagramming Relationships would seem to be a sound measure of general intelligence in that it was substantially and significantly correlated with every cognitive test, as well as having small but significant correlations with Consistency (Intrapersonal) and Rhythm (Musical). This finding suggests not only that logical reasoning is a good measure of *g*,

but also that reasoning might be a cross-cutting ability which is utilized within each of Gardner's cognitive intelligence domains, counter to Gardner's model which assumes there are no such horizontal abilities.

Other tests with high *g*-loadings (greater than .45) are the remaining highly cognitive tests—that is, both Linguistic tests (Opposites and Vocabulary), both Spatial tests (Map Planning and Paper Folding), the second Naturalistic test (Making Groups), and Social Translations (Interpersonal). The high *g*-loadings of the Spatial tests would seem to contradict Gardner's notion that observed correlations among ability tests occur due to a common verbal element. However, some support for the independence of Gardner's intelligences, at least in regards to a distinction between Linguistic and Spatial intelligences, was found in that neither Spatial test was significantly correlated with either of the Linguistic tests. This result might be due to the nature of the sample in the current study. The overwhelming majority of participants were undergraduate or graduate students, a group who have possibly been admitted to university partly on the basis of their abilities in Spatial or Linguistic areas. In addition, some restriction in range of intelligence of the sample might have attenuated correlations in the current study. On average, WPT scores in this study corresponded to a full-scale WAIS score of 114—almost one standard deviation above the WAIS mean of 100. Also, the WPT standard deviation in the current study (5.79) is smaller than that reported for the WAIS normative study (7.12), indicating that participants in this study did not vary in cognitive ability as widely as in the general population.

Vocabulary was not as highly correlated with the WPT as previous research would have suggested (e.g., Wechsler, 1997), despite the fact that vocabulary items are

represented on the WPT. However, the modest reliability of the Vocabulary test used in the current study attenuated somewhat its association with the WPT and its *g*-loading. This weaker than expected reliability might have been a function of the difficulty of the vocabulary test employed in the current study. The test included relatively uncommon words (e.g., salaam, perambulator, yawl, hauteur). Ullstadius, Gustafsson, and Carlstedt (2002) found that *g*-loadings for easier words were greater than for difficult words. In addition, they reported that not only were infrequent words poor measures of *g*, but also of verbal ability. In addition, there were only two Linguistic measures, whereas studies that find high vocabulary loadings typically have several verbal measures, potentially resulting in an overweighting of the verbal *g*-loading. However, Jensen (1991) noted that the WAIS vocabulary subscale has a loading of .82 when factor analyzed with only the six nonverbal subtests.

Tests with moderate *g*-loadings (.24 to .36) included Cartoon Predictions (Interpersonal), Consistency (Intrapersonal), and Subtraction and Multiplication (Logical/Mathematical). These tests are all somewhat less cognitive in nature than the highly *g*-loaded tests, which all contained an element of reasoning. Particularly noteworthy is the fact that Subtraction and Multiplication had a *g*-loading that was much smaller than that of the other Logical/Mathematical task, Necessary Arithmetic Operations. While both tests have some element of mathematical ability, the reasoning task was far more saturated with general intelligence than performing rote mathematical computations.

It is possible that Cartoon Predictions might have been more *g*-loaded had it been more reliable. Future investigations into the structure of mental abilities would do well to

employ a more reliable nonverbal measure of Interpersonal ability.

Three of the remaining tests, Rhythm, Tonal, and Accuracy, had small *g*-loadings between .10 and .20. It is difficult to make any conclusions about these tests, given that two had extremely poor reliability, and the other was a first attempt at measuring Intrapersonal intelligence. It seems possible that these three tests have somewhat greater cognitive content than originally hypothesized, and that their *g*-loadings have been attenuated by reliability and measurement issues. The small loadings of the Musical tests and the fact that they were not significantly correlated with the WPT contrasts with some research findings (e.g., Lynn, Wilson, & Gault, 1989). In addition, whereas studies have found AMMA Tonal and Rhythm to be correlated with spatial tests (Gromko, 2004), no such association appeared in the current study. Whether the unreliability of the Musical tests was due to the test itself or factors specific to this study, such as acoustics in the testing rooms or participant fatigue, it is almost impossible to draw conclusions from the current study about how musical ability fits into the structure of mental abilities. Future research with refined Intrapersonal measures and reliable Musical testing could result in very different findings regarding these domains as they relate to other abilities and to achievement.

The Bodily-Kinesthetic tests, on the other hand, were purely non-cognitive in nature, and their near-zero *g*-loadings and correlations with the WPT were not surprising.

The tests within each ability domain tended to share more variance than they did with tests in other domains, suggesting that the intelligences can be seen as second-order ability factors. The Linguistic tests, for example, had considerable shared variance beyond their substantial *g*-loadings. This finding fits with both hierarchical models and

with Gardner's conceptualization of independent intelligence domains. The two Naturalistic tests, on the other hand, shared little variance beyond that attributable to *g*, suggesting that Naturalistic intelligence is not a coherent domain. The two Bodily-Kinesthetic tests were not significantly correlated, suggesting that they, also, do not form a coherent ability domain.

While it is difficult to imagine any pattern of results which would refute Multiple Intelligences theory, the results of the current study support hierarchical models of intelligence. The substantial *g*-loadings of Naturalistic and Spatial tests suggest that *g* is far more pervasive than Gardner would seem to suggest. Gardner has stated that intelligence tests tend to largely measure Linguistic and Logical/Mathematical abilities, but the WPT was positively and significantly correlated with all cognitive tests, both verbal and nonverbal, and all cognitive measures clearly loaded onto a single factor. Multiple Intelligences theory would seem to predict more coherent ability domains than those which emerged in the current study. In addition, Gardner's theory would seem to predict that the ability domains should have been relatively independent of each other.

Self-Estimated Multiple Intelligences

The weak correlations between self-estimated and actual abilities indicated that neither direct self-ratings nor the indirect MAQ was an adequate substitute for ability tests. Participants were better able to estimate their ability in domains that are frequently observed in people's lives (gross motor, linguistic, mathematical), but even then, correlations were only moderate. Neither Naturalistic test was correlated with self-rated Naturalistic ability, for example. The ability to classify objects is not greatly discussed or

measured in western cultures, and participants might have been relatively unaware of strengths or deficits in this area.

It is surprising, however, that neither Spatial measure was significantly correlated with self-rated Spatial intelligence, because this ability would seem to be quite important in people's lives if they are to understand diagrams and graphs or even find their way around without getting lost. It is also somewhat surprising that Interpersonal measures were not more highly correlated with self-rated Interpersonal ability, because people might be expected to have become aware of strengths and shortcomings in an ability that is so frequently employed. However, the poor reliability of the Cartoon Predictions test attenuated its correlation with self-rated Interpersonal ability.

Self-rated Bodily-Kinesthetic ability was significantly correlated with the Stork Stand but not with Mark Making. The fact that the two MAQ Bodily-Kinesthetic items that specifically referred to fine motor abilities were uncorrelated with Mark Making is troublesome, suggesting validity issues with either these two items or with Mark Making or with self-estimates in general.

Consistency (Intrapersonal) showed a clear pattern of negative correlations with both self-ratings and MAQ self-estimates, whereas correlations between Consistency and measured abilities were generally positive. One explanation for this effect is that participants who were consistent in describing their personalities were also more realistic in describing their own abilities. Participants typically described themselves as above average in every ability domain. Perhaps participants who were high in Consistency rated their own abilities at a more realistic and, thus, lower level than participants who received low scores in Consistency. This increased accuracy would not be reflected in the

Accuracy score of these participants, since general over- or under-estimation of one's own abilities was controlled in the calculation of the Accuracy score.

Paulhus et al. (1998) found that self-report measures, even with aggregation, did not exceed .30 in correlations with measured intelligence. In the current study, the MAQ's indirectly self-estimated Logical/Mathematical ability had the highest correlation with the WPT, at .32, a greater correlation with measured ability than Paulhus et al. achieved in associating self-estimates with measured ability. Within the direct self-ratings, self-rated Logical/Mathematical ability also had the highest correlation with the WPT. In fact, self-rated Logical/Mathematical ability was more highly correlated with the WPT than was self-rated overall intelligence.

In the current study, several correlations between self-estimated abilities and specific tests slightly exceeded Paulhus et al.'s (1998) .30 maximum: self-rated Logical/Mathematical intelligence and Necessary Arithmetic, a mathematical reasoning test; self-rated Bodily-Kinesthetic intelligence and the Stork Stand, a balance test; indirectly self-estimated Linguistic ability and Opposites, an associational fluency task; and indirectly self-estimated Linguistic ability and Vocabulary. Paulhus et al. reported that direct items about broad cognitive abilities and items with content related to enjoyment and frequency of reading were most highly associated with measured intelligence, but that was not the case in the current study.

Neither form of self-assessment used in this study provided an accurate assessment of all the intelligence domains; in fact, neither instrument yielded any correlation with actual performance above .36. While the MAQ was somewhat superior to direct self-ratings in terms of its correlations with measured abilities, it would seem

that people are not accurate in predicting their own performance in most domains, particular those in which they would not have received a great deal of clear feedback about their performance, such as Naturalistic and Intrapersonal. This finding casts some doubt as to the validity of the numerous self-report questionnaires being used to assess Multiple Intelligences. It seems possible that these questionnaires could be measuring interest, self-esteem, or some other construct altogether.

Personality

As hypothesized, self-reported Openness to Experience was correlated with both Linguistic tests, which are measures of crystallized intelligence. In addition, Openness to Experience had a smaller but significant correlation with Cartoon Predictions and Social Translations, which also have a strong element of crystallized knowledge. Correlations with all other tests were non-significant and near-zero, suggesting that, as hypothesized, Openness to Experience is not associated with fluid intelligence.

The significant correlation between Vocabulary and Honesty-Humility could potentially be related to the content of the Honesty-Humility items (*sly, sneaky, tricky, undevious, pretentious, pompous, cunning, and undeceptive*). Because there are fewer positive adjectives in the Honesty-Humility domain, this personality factor was the only one which did not have equal numbers of positive and negative words. While *pretentious* was defined for participants, the other words were not. It is possible that participants with weaker vocabularies did not understand the negative connotations of some words and were, thus, more likely to agree with them.

Openness to Experience and Conscientiousness were both correlated with many

of the directly self-rated and indirectly self-estimated abilities. It could be that academically-oriented individuals who see themselves as highly conscientious and open-minded, two characteristics highly valued in a university environment, might also see themselves as being highly able in other school-related abilities. Perhaps the correlations between the two personality factors and cognitive abilities would be somewhat smaller in a community sample.

Sex Differences

Sex differences in this study were generally unremarkable. Differences between the sexes in measured multiple intelligences were small in number and in magnitude. Men's direct self-ratings of their abilities were significantly higher than women's in half of the rated abilities. Given that all the self-ratings tended to be inflated, it seems possible that women's self-ratings might be more realistic and accurate than men's in these four domains. In the Multiple Ability Questionnaire, the sex differences were more evenly distributed in that men estimated their Logical/Mathematical and Naturalistic abilities higher than did women, and women estimated their Interpersonal and Intrapersonal abilities higher than did men. Sex differences in personality were as hypothesized in that women reported greater Emotionality and Honesty-Humility. Women's greater self-reported Extraversion was unhypothesized, but the effect size was quite small.

Conclusions

The substantial *g*-loadings of all cognitive tests in the current study contradict Gardner's assertion that there are at least eight independent intelligence domains. While Gardner has acknowledged the existence of *g* and conceded that the eight intelligences might not be entirely independent, his contention that positive correlation between various cognitive tasks are largely due to verbal demands was clearly not supported in this study. That finding that the non-cognitive tests in this study were not *g*-loaded is unsurprising, and suggests that Gardner is likely correct in claiming that Bodily-Kinesthetic ability is quite different from the various cognitive abilities.

The coherence of some of the ability domains, as shown by significant residual correlations after the extraction of *g*, tends to lend partial support to Multiple Intelligence theory and to hierarchical models. Tests within certain domains, such as Linguistic, clearly had much in common over and above any relationship with *g*. Tests within the Intrapersonal domain, however, were uncorrelated, and the small correlation between the two Bodily-Kinesthetic tests did not reach statistical significance, suggesting that the intelligence domains do not all represent coherent constructs.

Without further research, it seems premature for schools and parents to wholeheartedly embrace Multiple Intelligences curriculums. A cautious approach to Multiple Intelligences theory does not imply that students should not be treated as individuals with strengths to be nurtured and weakness to be remediated. Instead, the current study would seem to suggest that Multiple Intelligences theory does not explain the data with sufficient precision for this model to be the basis of classroom planning.

Future explorations of Gardner's model could focus on refining assessment

instruments, particularly Intrapersonal measures. Gardner's approach to assessment, which would have people tested with intelligence-fair but familiar materials in the performance of culturally-valued tasks, does not easily lend itself to experimental standardization. Other future investigations of mental abilities could explore the existence and nature of cross-cutting, horizontal abilities, such as memory, creativity, and/or reasoning.

References

- Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, 121, 219-245.
- Alix, N. M. (2000). A theory of multiple intelligences: A case of missing cognitive matter. *Australian Journal of Education*, 44, 272-293.
- Ashton, M. C., & Lee, K. (2005). Problems with the method of correlated vectors. *Intelligence*, 33, 431-444.
- Ashton, M. C., Lee, K., & Goldber, L. R. (2004). A hierarchical analysis of 1,710 English personality-descriptive adjectives. *Journal of Personality & Social Science*, 87, 707-721.
- Ashton, M. C., Lee, K., Perugini, M., Szarota, P., De Vries, R. E., Di Blas, L., et al. (2004). A six factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, 86, 356-366.
- Ashton, M. C., Lee, K., & Vernon, P. A., & Jang, K. L. (2000). Fluid intelligence, crystallized intelligence, and the openness/intellect factor. *Journal of Research in Personality*, 34, 198-207.
- Bates, T. C., & Shieles, A. (2003). Crystallized intelligence as a product of speed for experience: the relationship of inspection time and openness to g and Gc. *Intelligence*, 31, 275-287.
- Bennett, M. (1996). Men's and women's self-estimates of intelligence. *Journal of Social Psychology*, 136, 411-4112.
- Borkenau, P., & Liebler, A. (1993). Convergence of stranger ratings of personality and intelligence with self-ratings, partner ratings, and measured intelligence. *Journal of Personality & Social Psychology*, 65, 546-553.
- Brandler, S., & Rammsayer, T. H. (2003). Differences in mental abilities between musicians and non-musicians. *Psychology of Music*, 31, 123-138.
- Budaev, S. V. (1999). Sex differences in the Big Five personality factors: Testing an evolutionary hypothesis. *Personality & Individual Differences*, 26, 801-813.
- Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Personality & Individual Differences*, 29, 1057-1068.

- Campbell, J. D. (1990). Self-esteem and clarity of the self-concept. *Journal of Personality & Social Psychology*, 59, 538-549.
- Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. New York: Cambridge University Press.
- Chamorro-Premuzic, T. & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality*, 37, 319-338.
- Chamorro-Premuzic, T., Furnham, A., & Moutafi, J. (2004). The relationship between estimated and psychometric personality and intelligence scores. *Journal of Research in Personality*, 38, 505-513.
- Chen, J. (2004). Theory of multiple intelligences: Is it a scientific theory? *Teachers College Record*, 106, 17-23.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences*. Hillsdale, NJ: Erlbaum.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd ed. Hillsdale, NJ: Erlbaum.
- Colom, R., & Garcia-Lopez, O. (2002). Sex differences in fluid intelligence among high school graduates. *Personality & Individual Differences*, 32, 445-451.
- Colom, R., Juan-Espinosa, M., Abad, F. J., & Gardia, L. F. (2000). Negligible sex differences in general intelligence. *Intelligence*, 28, 57-68.
- Costa Jr., P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality & Social Psychology*, 81, 322-331.
- Cronbach, L. J. (1960). *Essentials of Psychological Testing* (2nd ed.). New York: Harper & Rowe.
- Davies, M., Stankov, L., & Roberts, R. D. (1998). Emotional intelligence: In search of the elusive construct. *Journal of Personality and Social Psychology*, 75, 989-1015.
- De Fruyt, F., & Mervielde, I. (1996). Personality and interests as predictors of educational streaming and achievement. *European Journal of Personality*, 10, 405-425.

- DeVellis R F. (1991). *Scale Development: Theory and Application*. Newbury Park, CA, Sage Publications.
- DiGiammarino, M. (1990). Functional music skills of persons with mental retardation. *Journal of Music Therapy*, 27, 209-220.
- Dodrill, C. B. (1983). Long-term reliability of the Wonderlic Personnel Test. *Journal of Consulting and Clinical Psychology*, 51, 316-317.
- Ekstrom, R. B., French, J. W., Harman, H. H., & Derman, D. (1999). *Manual for kit of factor-referenced cognitive tests 1976*. Princeton: Educational Testing Service.
- Enders, C. K. (2001). The performance of the full information maximum likelihood estimator in multiple regression models with missing data. *Educational and Psychological Measurement*, 61, 713-740.
- Farsides, T., & Woodfield, R. (2003). Individual differences and undergraduate academic success: the roles of personality, intelligence, and application. *Personality & Individual Differences*, 34, 1225-1243.
- Furnham, A. (2001). Self-estimates of intelligence: Culture and gender difference in self and other estimates of both general (g) and multiple intelligences. *Personality and Individual Differences*, 31, 1381-1405.
- Furnham, A., & Fong, G. (2000). Self-estimated and psychometrically measured intelligence: A cross-cultural and sex differences study of British and Singaporean students. *North American Journal of Psychology*, 2, 191-200.
- Furnham, A., & Rawles, R. (1995). Sex differences in the estimation of intelligence. *Journal of Social Behavior & Personality*, 10, 741-745.
- Furnham, A., & Ward, C. (2001). Sex differences, test experience, and self-estimation of multiple intelligences. *New Zealand Journal of Psychology*, 30, 52-59.
- Furnham, A., Tang, T. L., Lester, D., O'Connor, R. & Montgomery, R. (2002). Estimates of ten multiple intelligences. *European Psychologist*, 7(4), 245-255.
- Furnham, A., Clark, K. & Bailey, K. (1999). Sex differences in estimates of multiple intelligences. *European Journal of Personality*, 13, 247-259.
- Gardner, H. (1983). *Frames of mind*. New York: BasicBooks.
- Gardner, H. (1993). *Multiple intelligences*. New York: BasicBooks.
- Gardner, H. (1999). *Intelligence reframed*. New York: BasicBooks.

- Gardner, H., & Connell, M. (2000). Response to Nicholas Alix. *Australian Journal of Education*, 44, 288-293.
- George, D., & Mallery, P. (1999). *SPSS for Windows step by step: A simple guide and reference*. Boston: Allyn & Bacon.
- Gordon, E. (1982). *Intermediate Measures of Music Audiation: A music aptitude test for first, second, third, and fourth grade children*. Chicago: G.I.A. Publications.
- Gordon, E. (1989). *Advanced Measures of Music Audiation*. Chicago: GIA Publications.
- Gottfredson, L. S. (2002). *g*: Highly general and highly practical. In R. J. Sternberg & E. L. Grigorenko (Eds.), *The General Factor of Intelligence: How General is It?* (pp. 331-380). Mahwah, NJ: Erlbaum.
- Gromko, J. E. (2004). Predictors of music sight-reading ability in high school wind players. *Journal of Research in Music Education*, 52, 6-16.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Hair, E. C., & Graziano, W. G. (2003). Self-esteem, personality and achievement in high school: A prospective longitudinal study in Texas. *Journal of Personality*, 71, 971-994.
- Halpern, D. F. (1997). Sex differences in intelligence. *American Psychologist*, 52, 1091 - 1102.
- Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, 32, 175-191.
- Hogan, H. (1978). IQ self-estimates of males and females. *Journal of Social Psychology*, 106, 137-138.
- Holland, D. C., Dollinger, S. J., Holland, C. J., & MacDonald, D. A. (1995). The relationship between psychometric intelligence and the five-factor model of personality in a rehabilitation sample. *Journal of Clinical Psychology*, 51, 79-88.
- Humphreys, L. G. (1986). Commentary. *Journal of Vocational Behavior*, 29, 421-437.
- Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107, 139-155.
- Jensen, A. R. (1991). Vocabulary and general intelligence. *Behavioral & Brain Sciences*, 24, 1109-1110.
- Jensen, A. (1998). *The g factor*. London: Praeger.

- Johnson, B. L., & Nelson, J. K. (1986). *Practical measurements for evaluation in physical education*, 4th ed. New York: MacMillan.
- Klein, P. D. (1998). A response to Howard Gardner: falsifiability, empirical evidence and pedagogical usefulness in educational psychologies. *Canadian Journal of Education*, 23, 103-112.
- Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO Personality Inventory. *Multivariate Behavioral Research*, 39, 329-358.
- Loehlin, J. C., McCrae, R. R., & Costa, P. T. (1998). Heritabilities of common and measure-specific components of the Big Five personality factors. *Journal of Research in Personality*, 32, 431-453.
- Lynn, R. (1999). Sex differences in intelligence and brain size: a developmental theory. *Intelligence*, 27, 1-12.
- Lynn, R., Wilson, G., & Gault, A. (1989). Simple musical tests as measures of Spearman's g . *Personality and Individual Differences*, 10, 25-28.
- McCrae, R. R., Zonderman, A. B., Costa Jr., P. T., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the revised NEO personality inventory: Confirmatory factor analysis versus procrustes rotation. *Personality Processes & Individual Differences*, 70, 552-556.
- McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60, 175-215.
- McMahon, S. D., Rose, D. S., & Parks, M. (2004). Multiple intelligences and reading achievement: An examination of the Teele Inventory of Multiple Intelligences. *Journal of Experimental Education*, 73, 41-52.
- Messick, S. (1992). Multiple intelligences or multilevel intelligence? Selective emphasis on distinctive properties of hierarchy: On Gardner's *Frames of Mind* and Sternberg's *Beyond IQ* in the context of theory and research on the structure of human abilities. *Psychological Inquiry*, 3, 365-384.
- Moutafi, J., Furnham, A., & Crump, J. (2003). Demographic and personality predictors of intelligence: A study using the Neo Personality Inventory and the Myers-Briggs type indicator. *European Journal of Personality*, 17, 79-94.
- Moutafi, J., Furnham, A., & Paltiel, L. (2004). Why is conscientiousness negatively correlated with intelligence? *Personality & Individual Differences*, 37, 1013-1022.

- O'Connor, Jr., R. M., & Little, I. S. (2003). Revisiting the predictive validity of emotional intelligence: self-report versus ability-based measures. *Personality and Individual Differences*, 35, 1893-1902.
- O'Sullivan, M. (1973). Black-white differences of tests of semantic, figural and behavioral-cognition abilities. *Unpublished manuscript*, University of San Francisco.
- O'Sullivan, M., & Guilford, J. P. (1976). *Four factor tests of social intelligence (behavioral cognition) manual of instructions and interpretations*. CA: Sheridan Psychological Services.
- O'Sullivan, M., Guilford, J. P., & deMille, R. (1965). The measurement of social intelligence. *Reports from the Psychological Laboratory*, 34.
- Paulhus, D. L., Lysy, D. C., & Yik, M. S. M. (1998). Self-report measures of intelligence: Are they useful as proxy IQ tests? *Journal of Personality*, 66, 525-554.
- Rammsayer, T. H., & Rammstedt, B. (2000). Sex related differences in time estimation: the role of personality. *Personality & Individual Differences*, 29, 301-312.
- Rammstedt, B., & Rammsayer, T. H. (2000). Sex differences in self-estimates of different aspects of intelligence. *Personality and Individual Differences*, 29 (2000), 869-880.
- Reilly, J., & Mulhern, G. (1995). Gender-differences in self-estimated IQ: the need for care in interpreting group data. *Personality & Individual Differences*, 18, 189-192.
- Root-Berstein, R. S. (2001). Music, creativity, and scientific thinking. *Leonardo*, 34, 63-68.
- Ruthsatz, J., & Detterman, D. K. (2003). An extraordinary memory: The case study of a musical prodigy. *Intelligence*, 31, 509-518.
- Schulte, M. J., Ree, M. J., & Carretta, T. R. (2004). Emotional intelligence: not much more than g and personality. *Personality & Individual Differences*, 37, 1059-1068.
- Shearer, C. B. (1991). An investigation into the validity, reliability and clinical utility of the Hillside Assessment of Perceived Intelligences. (Docotral dissertation, Union Institute, 1991). *Dissertation Abstracts International*, 52, 6647B.
- Shearer, C. B. (1994). *The validation of the Hillside Assessment of Perceived Intelligences (HAPI): A measure of Howard Gardner's theory of multiple intelligences*. Paper presented at the annual meeting of the American Educational

Research Association, New Orleans, LA.

Shearer, C. B. (1996). *The MIDAS: A guide to assessment and education for the multiple intelligences*. Columbus, OH: Greyden Press.

Shearer, C. B. (1997, August). *Reliability, validity and utility of a multiple intelligences assessment for career planning*. Paper presented at the annual convention of the American Psychological Association, Chicago, IL.

Spearman, C. (1904). General intelligence objectively determined and measured. *American Journal of Psychology*, 5, 201-293.

Stankov, L. (2000). Structural extensions of a hierarchical view on human cognitive abilities. *Learning and Individual Differences*, 12, 35-51.

Stollery, P., & McPhee, A. (2002). Some perspectives on musical gift and musical intelligence. *British Journal of Music Education*, 19(1), 89-102.

Stewart, L, Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003). Brain changes after learning to read and play music. *NeuroImage*, 20, 71-83.

Teele, S. (1992). *Teele Inventory of Multiple Intelligences*. Redlands, CA: Sue Teele and Associates.

Thurston, L. L. (1938). *Primary Mental Abilities*. Chicago: University of Chicago Press.

Torff, B., & Gardner, H. (1999). The vertical mind—the case for multiple intelligences. In M. Anderson (Ed.), *The Development of Intelligence* (pp. 139-159). East Sussex, UK: Psychology Press.

Ullstadius, E., Gustafsson, J. E., & Carlstedt, B. (2002). Influence of general and crystallized intelligence on vocabulary test performance. *European Journal of Psychological Assessment*, 18, 78-84.

United States Department of Labor (1970). *General Aptitude Test Battery*. Washington, D.C.: U. S. Government Printing Office.

Van der Zee, K., Thijs, M., & Schakel, L. (2002). The relationship of emotional intelligence with academic intelligence and the Big Five. *European Journal of Personality*, 16, 103-125.

Vernon, P. E. (1961). *The structure of human abilities (Rev. ed.)*. London: Methuen.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale—third edition administration and scoring manual*. San Antonio, TX: The Psychological Corporation.

Wonderlic (2002). *Wonderlic Personnel Test & Scholastic Level Exam user's manual*.
Libertyville, IL: Wonderlic.

APPENDIX A

Brock University

Senate Research Ethics Board**Extensions 3943/3035, Room AS 302**

DATE: March 10, 2004

FROM: Joe Engemann, Chair
Senate Research Ethics Board (REB)

TO: Dr. Michael Ashton, Psychology
Beth Visser

FILE: 03-317 Visser

TITLE: Beyond g: Putting Multiple Intelligence Theory to the Test

Chair
Research Ethics Board

The Brock University Research Ethics Board has reviewed the above research proposal.

DECISION: Accepted as is. However, please remove guarantee of anonymity in the Information Letter. (Anonymity means you cannot associate the data to participants). Guarantees of confidentiality are sufficient.

This project has been approved for the period of **March 10, 2004** to **April 30, 2005** subject to full REB ratification at the Research Ethics Board's next scheduled meeting. The approval may be extended upon request. *The study may now proceed.*

Please note that the Research Ethics Board (REB) requires that you adhere to the protocol as last reviewed and approved by the REB. The Board must approve any modifications before they can be implemented. If you wish to modify your research project, please refer to www.BrockU.CA/researchservices/forms.html to complete the appropriate form **REB-03 (2001) Request for Clearance of a Revision or Modification to an Ongoing Application**.

Adverse or unexpected events must be reported to the REB as soon as possible with an indication of how these events affect, in the view of the Principal Investigator, the safety of the participants and the continuation of the protocol.

If research participants are in the care of a health facility, at a school, or other institution or community organization, it is the responsibility of the Principal Investigator to ensure that the ethical guidelines and approvals of those facilities or institutions are obtained and filed with the REB prior to the initiation of any research protocols.

The Tri-Council Policy Statement requires that ongoing research be monitored. A Final Report is required for all projects, with the exception of undergraduate projects, upon completion of the project. Researchers with projects lasting more than one year are required to submit a Continuing Review Report annually. The Office of Research Services will contact you when this form **REB-02 (2001) Continuing Review/Final Report** is required.

Please quote your REB file number on all future correspondence.

APPENDIX B

MULTIPLE ABILITY QUESTIONNAIRE

Sex: M F Faculty: _____

Year of Study: _____ Major: _____

Age: _____ Previous year's average mark: _____ %
(Estimate if necessary)

Think about each statement, and indicate how accurately it describes your abilities, using the following scale:

- 1 = very inaccurate**
2 = somewhat inaccurate
3 = neither accurate nor inaccurate
4 = somewhat accurate
5 = very accurate

I'd be good at figuring out how to design a scientific experiment.	1	2	3	4	5
I am comfortable with writing letters when I need to.	1	2	3	4	5
I have difficulty doing math calculations in my head.	1	2	3	4	5
I'm pretty clumsy.	1	2	3	4	5
I notice when a singer's pitch is off.	1	2	3	4	5
All types of clouds look pretty much the same to me.	1	2	3	4	5
I'm good at figuring out how things work.	1	2	3	4	5
I am able to get the hang of a new sport pretty quickly.	1	2	3	4	5
I think I have a pretty good understanding of what my strengths and weaknesses are.	1	2	3	4	5
When I walk into a building, I can usually figure out the layout pretty quickly.	1	2	3	4	5
I can identify some birds from their calls.	1	2	3	4	5
I can usually tell when someone is lying to me.	1	2	3	4	5
I can usually remember a tune after I've heard it once or twice.	1	2	3	4	5
I enjoy taking tests and quizzes that help me learn more about myself.	1	2	3	4	5
I have some difficulty shuffling a deck of cards.	1	2	3	4	5
I can estimate distances pretty accurately.	1	2	3	4	5
I have good balance.	1	2	3	4	5
I don't have a very scientific mind.	1	2	3	4	5
I'm frequently off-key when I sing.	1	2	3	4	5
I can picture how a room will look with different furniture arrangements.	1	2	3	4	5
I notice when the signs of fall are earlier or later than usual.	1	2	3	4	5

Think about each statement, and indicate how accurately it describes your abilities, using the following scale:

- 1 = very inaccurate**
2 = somewhat inaccurate
3 = neither accurate nor inaccurate
4 = somewhat accurate
5 = very accurate

I don't care for reading.	1	2	3	4	5
When I've made a mistake solving a math problem, I can usually go back and figure out where I went wrong.	1	2	3	4	5
I'm usually the one who smooths things over when my friends or family get into disagreements .	1	2	3	4	5
I sometimes choose a job or course then discover it doesn't suit me at all.	1	2	3	4	5
I can speak very persuasively when I need to.	1	2	3	4	5
I have trouble differentiating between various makes and models of cars.	1	2	3	4	5
I'm good at making things with my hands.	1	2	3	4	5
I find it pretty easy to learn new dance steps.	1	2	3	4	5
I often tap my foot in perfect rhythm while listening to music.	1	2	3	4	5
I have difficulty understanding information presented in diagrams.	1	2	3	4	5
Sometimes I have no idea why I do the things I do.	1	2	3	4	5
I sometimes plan out what my life will be like in five years.	1	2	3	4	5
I'm a very logical person	1	2	3	4	5
I don't have a very good sense of direction.	1	2	3	4	5
People often use words I don't understand.	1	2	3	4	5
I'm good at completing jigsaw puzzles.	1	2	3	4	5
I have difficulty clapping to a beat.	1	2	3	4	5
I often find myself in conflict with my friends or family.	1	2	3	4	5
When I'm in a bad mood, I can usually pinpoint exactly what's bothering me.	1	2	3	4	5
I sometimes get annoyed when people tell me their problems.	1	2	3	4	5
I think of myself as a musical person.	1	2	3	4	5
When I look in the night sky, I can usually distinguish stars from planets.	1	2	3	4	5
I sometimes find it difficult to understand articles in the newspaper.	1	2	3	4	5
I'm pretty good at learning foreign languages.	1	2	3	4	5
I can usually tell someone's upset before they tell me.	1	2	3	4	5
I pride myself on being a good friend.	1	2	3	4	5
I can identify most species of trees.	1	2	3	4	5

APPENDIX C

SELF-ESTIMATED ABILITY INVENTORY

For each ability below, indicate how you think you compare with other Brock University students by placing an "x" on the line to indicate what percentage of these students would have a lower level of ability than you do.

1. Using and understanding language (both in writing and speaking).

2. Analyzing problems logically, and working things out mathematically and/or scientifically.

3. Natural sense of musical pitch and rhythm.

4. Spatial ability (for example, visualizing objects from different angles, orienting yourself in an unfamiliar place, recognizing visual patterns).

5. Coordinating your physical movements (for example, in sports, dancing, crafts, technical work).

6. Understanding yourself, your strengths and weaknesses, and making good decisions about your life.

7. Working with other people and understanding why they do the things they do.

8. Classifying things (especially living things) and making distinctions between similar objects.

9. Overall intelligence.

APPENDIX D

Personality Self-Rating Form

Please consider each adjective below, and indicate how accurately that adjective describes your typical personality, using the following scale:

- 1 = extremely inaccurate
- 2 = very inaccurate
- 3 = somewhat inaccurate
- 4 = slightly inaccurate
- 5 = neither accurate nor inaccurate
- 6 = slightly accurate
- 7 = somewhat accurate
- 8 = very accurate
- 9 = extremely accurate

Creative	1 2 3 4 5 6 7 8 9	Talkative	1 2 3 4 5 6 7 8 9
Unimaginative	1 2 3 4 5 6 7 8 9	Quiet	1 2 3 4 5 6 7 8 9
Philosophical	1 2 3 4 5 6 7 8 9	Outgoing	1 2 3 4 5 6 7 8 9
Narrow-minded	1 2 3 4 5 6 7 8 9	Shy	1 2 3 4 5 6 7 8 9
Original	1 2 3 4 5 6 7 8 9	Lively	1 2 3 4 5 6 7 8 9
Conventional	1 2 3 4 5 6 7 8 9	Passive	1 2 3 4 5 6 7 8 9
Complex	1 2 3 4 5 6 7 8 9	Cheerful	1 2 3 4 5 6 7 8 9
Shallow	1 2 3 4 5 6 7 8 9	Restrained	1 2 3 4 5 6 7 8 9
Orderly	1 2 3 4 5 6 7 8 9	Anxious	1 2 3 4 5 6 7 8 9
Disorganized	1 2 3 4 5 6 7 8 9	Fearless	1 2 3 4 5 6 7 8 9
Thorough	1 2 3 4 5 6 7 8 9	Worrying	1 2 3 4 5 6 7 8 9
Undisciplined	1 2 3 4 5 6 7 8 9	Insensitive	1 2 3 4 5 6 7 8 9
Diligent*	1 2 3 4 5 6 7 8 9	Sentimental	1 2 3 4 5 6 7 8 9
Inefficient	1 2 3 4 5 6 7 8 9	Tough	1 2 3 4 5 6 7 8 9
Precise	1 2 3 4 5 6 7 8 9	Emotional	1 2 3 4 5 6 7 8 9
Irresponsible	1 2 3 4 5 6 7 8 9	Feelingless	1 2 3 4 5 6 7 8 9
Patient	1 2 3 4 5 6 7 8 9	Sly	1 2 3 4 5 6 7 8 9
Quick-Tempered	1 2 3 4 5 6 7 8 9	Sneaky	1 2 3 4 5 6 7 8 9
Tolerant	1 2 3 4 5 6 7 8 9	Tricky	1 2 3 4 5 6 7 8 9
Argumentative	1 2 3 4 5 6 7 8 9	Undevious	1 2 3 4 5 6 7 8 9
Forgiving	1 2 3 4 5 6 7 8 9	Pretentious**	1 2 3 4 5 6 7 8 9
Irritable	1 2 3 4 5 6 7 8 9	Pompous	1 2 3 4 5 6 7 8 9
Gentle	1 2 3 4 5 6 7 8 9	Cunning	1 2 3 4 5 6 7 8 9
Critical	1 2 3 4 5 6 7 8 9	Undeceptive	1 2 3 4 5 6 7 8 9

* Diligent means hard-working

** Pretentious means self-important

APPENDIX E